Mental health risk factors for criminal recidivism and mortality in individuals given community sentences

A thesis presented for the degree of Doctor of Philosophy

Denis Yukhnenko

Department of Psychiatry

University College, University of Oxford

Word count: 39,500

Michaelmas 2021

To those who speak up against injustice and war

ACKNOWLEDGEMENTS

I would like to thank colleagues, friends, and family, who have supported me throughout my studies and contributed to my work.

I am deeply grateful to my supervisors, Seena Fazel and Nigel Blackwood, for their guidance, support, and patience. I learned a lot while studying under your supervision. I wanted to thank my colleagues Yasmina Molero Samuelson, Rongqin Yu, Howard Ryland, Tyra Lagerberg, Daniel Whiting, and Achim Wolf for their support, advice, and contribution to my research work. You are all wonderful to be around. I am also thankful to my colleagues from Karolinska Institutet and Swedish Prison and Probation Service for their advice. I am especially grateful to Thomas Fanshawe, Andrey Kormilitzin, and Audinga-Dea Hazewinkel for their advice on statistical methods.

My time in Oxford would not be complete without my friends from Magpie Lane, London Place, and the Exchange. It is a long list, but you know who you are. You are all amazing people, and you will go on to do amazing things.

I wanted to thank my mother, Zoya, and my grandmother, Maria, for their support and encouragement throughout my life full of seemingly never-ending studies. I am grateful to my close friend, Axioma, who is no longer with us, for her continuous unconditional support during the last 13 years. I miss you.

Last but not least, I would like to thank my girlfriend, Vita Spivak, for her support, understanding, and ability to inspire. You are the best!

Publication arising from the thesis

Yukhnenko, D., Wolf, A., Blackwood, N., & Fazel, S. (2019). Recidivism rates in individuals receiving community sentences: A systematic review. *PLOS ONE*, 14(9), e0222495. Retrieved from https://doi.org/10.1371/journal.pone.0222495

Yukhnenko, D., Blackwood, N., & Fazel, S. (2020). Risk factors for recidivism in individuals receiving community sentences: a systematic review and meta-analysis. *CNS Spectrums*, 25(2), 252–263. https://doi.org/DOI: 10.1017/S1092852919001056

Yukhnenko, D., Blackwood, N., & Fazel, S. (2020). Risk factors for recidivism in individuals receiving community sentences: a systematic review and meta-analysis. In Warburton, K., and Stahl, S. M. (Eds.). *Decriminalizing Mental Illness* (Chapter 30). Cambridge University Press.

Related publications

Yukhnenko, D., Sridhar, S., & Fazel, S. (2019). A systematic review of criminal recidivism rates worldwide: 3-year update. *Wellcome Open Research*, *4*.

CONTENTS

List of Tables

LIST OF FIGURES

LIST OF APPENDICES

ABSTRACT

Individuals with psychiatric disorders are overrepresented in correctional populations. Given their high prevalence, understanding the role that psychiatric disorders play in criminal behaviour is essential for achieving the primary goal of correctional systems - preventing future crime. Understanding psychiatric disorders is also relevant for mitigating potential health risks, mortality being the most serious, associated with criminal behaviour and sentencing. Despite the apparent need, very few studies have examined the association of psychiatric disorders with criminal recidivism and mortality in non-custodial populations, i.e., individuals given community sentences. The present thesis addresses this research gap and provides recommendations for evidence-based risk management of individuals with psychiatric disorders undergoing community supervision.

The thesis aimed to examine the role of psychiatric disorders as factors for criminal recidivism and mortality in individuals given community sentences and to create a simple, scalable tool for criminal recidivism risk monitoring in this population. To achieve these aims, I conducted five separate studies.

The first study was a systematic review of the recidivism rates in individuals given community sentences internationally. The review covers 28 studies with data from 19 countries. Based on this systematic review, the two-year reconviction was selected as the primary criminal recidivism outcome in the present thesis. The second study reports the meta-analysis of existing research on risk factors for criminal recidivism in individuals given community sentences. The meta-analysis results further highlighted the lack of published research into psychiatric disorders and criminal recidivism in community sentenced populations. The strength of reported associations between mental health risk factors and criminal recidivism was comparable to that of non-modifiable risk factors, such as age, gender, and criminal history. However, the studies reporting these associations had major limitations, which I addressed in the third study.

The third study examined the association between psychiatric disorders and criminal recidivism in a Swedish population cohort of adult individuals given community sentences (N = 82,415) using register data. Psychiatric disorders were associated with an increased risk of both general and violent reoffending. This association persisted in

individuals matched with their sentenced siblings without a known psychiatric diagnosis. Schizophrenia spectrum disorders, personality disorders and substance use disorders had stronger effects on violent reoffending than other psychiatric disorders. Comorbid substance use fully mediated the association between psychiatric disorders and general reoffending. However, comorbid substance use only partially mediated the association between psychiatric disorders and violent reoffending. In addition, first-time psychiatric diagnosis during the follow-up period was associated with a higher risk of general and violent reoffending.

The fourth study additionally examined the association between psychiatric disorders and mortality in a Swedish total population cohort of adult individuals given community sentences (N = 109,751) using register data. The leading cause of death in the cohort was suicide. Having substance use or any other psychiatric disorder at the time of a sentence or receiving a new diagnosis during the follow-up was associated with an increased risk of all-cause and external-cause mortality in the study cohort. Comorbid substance use partially mediated the association of psychiatric disorders with mortality.

The fifth study used pre-specified criminal, sociodemographic, and clinical risk factors to develop a dynamic prediction model for criminal recidivism in individuals under community supervision. The model was deployed as an online dynamic risk assessment tool OxMore with good calibration and discrimination performance (c-index = 0.74 for violent reoffending, c-index = 0.69 for general reoffending). As an important secondary outcome, the study demonstrated that actuarial recidivism risk assessment tools, which have not been developed as risk monitoring instruments but are used as such, will systematically overestimate the risk of recidivism over time.

The thesis emphasises the role of substance use and other psychiatric disorders as important targets for intervention in individuals given community sentences. Addressing psychiatric disorders in this population could potentially prevent many cases of reoffending and improve the long-term health outcomes decreasing the burden on correctional and healthcare systems. The period of community supervision should be viewed as a major opportunity window for rehabilitation and reintegration of sentenced individuals with a psychiatric disorder. Other recommendations include an increase in integrative mental health research in community sentenced populations and a more extensive application of dynamic prediction modelling for actuarial recidivism prediction.

Introduction

Community sentences are currently used in many countries as a key alternative to imprisonment (Porporino, 2018). Community sentences form a heterogeneous group of criminal justice disposals, including suspended custodial sentences, probation with supervision, electronic monitoring, mandated community service, mandated treatment or training programmes, and other measures. As the name suggests, the key feature of community sentences is that a sentenced individual serves their sentence in a community and not in prison. A community sentenced individual can participate in social life, has the opportunity to maintain employment, and, to some extent, retains freedom of movement. Individuals given community sentences also have access to public education, healthcare and welfare systems, which is often not the case for prisoners.

Governments often expand the use of community sentences in an attempt to lower their country's crime rates and prison population (Heard, 2015). Community sentences are also viewed as a more humane approach to sentencing that allows for better rehabilitation and reintegration of offenders (Tonry, 2017). Compared to imprisonment, community sentences have been associated with lower criminal recidivism (Mews et al., 2015) and lower costs for the criminal justice system (Mann & Bermingham, 2020). Community sentences have also been recommended for specific subgroups, such as individuals with psychiatric disorders, who are more vulnerable to the effects of imprisonment (Canada et al., 2020; Ministry of Justice, 2018b; Sentencing Council, 2020).

However, the positive effect of community sentences is not automatic, it highly depends on their implementation in a particular jurisdiction. In many countries, the expanded use of community sentences did not achieve its declared aims (Commons Chamber, 2020; Heard, 2015). Prior investigations have identified several potential causes for the suboptimal results of community sentences (Heard, 2016). These causes include moving away from rehabilitation towards punishment and using imprisonment as an automatic sanction for community sentence violations. Potential remedies have also been proposed. One of these potential remedies, the improvement of community sentencing outcomes for sentenced individuals with psychiatric disorders, is the ultimate goal of the current thesis. Criminal recidivism after community sentences is the primary outcome examined in the thesis, and mortality is a secondary outcome.

Community sentences in Sweden

Sweden is a country that provides a study setting for the current thesis. In Sweden, most imposed sentences are community sanctions. In 2019, 11,300 individuals were sentenced to probation or received a suspended custodial sentence (Swedish National Council for Crime Prevention, 2021b). In the same year, 5,450 individuals received a prison sentence.

Probation is the most common community sentence in Sweden (Swedish Prison and Probation Service, 2021). An individual sentenced to probation usually gets a trial period of three years, the first year of which is supervised. Probation can also be coupled with conditions, such as required treatment (including substance use treatment), vocational training, and community service (Bosly et al., 2012). Sentenced individuals may also be required to keep the same place of residence. Committing a new crime can result in revocation of probation and imposition of another sentence.

A conditional sentence allows a sentenced individual to avoid the sanction for a committed crime on the condition that they live an orderly life during the next two years (Bosly et al., 2012). Conditional sentences are not supervised. A conditional sentence can also be combined with day-fines and include community service. Committing a new crime can result in the revocation of a conditional sentence and imposition of another sentence.

Community sentences in England and Wales

The description of community sentences in England and Wales is provided for comparison. In 2019, 127,213 individuals were sentenced to probation or received a suspended custodial sentence (Ministry of Justice, 2020a). In the same year, 75,971 individuals received a prison sentence.

Suspended prison sentences (or suspended sentence orders) in England and Wales are similar to conditional sentences in Sweden. The individual serves their sentence in the community instead of going to prison. Despite this, they are legally considered a type of custodial sentence. A suspended sentence can replace a regular custodial sentence, which intended length is no longer than two years (Sentencing Academy, 2021). A suspended sentence lasts between six months and two years. Community sentences (called ‘community orders’) are given when an individual does not meet the threshold for a custodial sentence. Community sentences can last for up to three years.

Both suspended sentences and community sentences can include a different combination of conditions that a sentenced individual has to observe. The conditions include regular supervision, community service, recidivism reduction programme, drug and alcohol treatment, mental health treatment, curfew, fixed residence, and prohibition of certain types of activity (Warwickshire and West Mercia CRC, 2021). Suspended and community sentences can be revoked if a sentenced individual does not comply with imposed conditions (Sentencing Council, 2021). From 2007 until 2020, the probation system in England and Wales was largely privatised. Community Rehabilitation Companies (CRCs) were responsible for supervising sentenced individuals in the community. In 2021, the new national probation service was created, and the contracts with CRCs were withdrawn (Ministry of Justice, 2021c).

Criminal recidivism

Recidivism is defined as a 'tendency to relapse into a previous condition or mode of behaviour; especially, relapse into criminal behaviour' (Marriam-Webster, 2021). Criminal recidivism could also be described as 'repeated criminal behaviour.' Although these definitions make intuitive sense, they are too broad for any meaningful practical application and need to be appropriately operationalised. When criminal recidivism is operationalised, it can be assessed quantitatively. In theory, recidivism estimates should serve as performance metrics that enable the comparison of different rehabilitation programmes, court sentencing practices, and correctional systems in general. In practice, the operationalised definitions substantially vary across various jurisdictions and research papers, which complicates comparisons (Andersen & Skardhamar, 2017; Fazel & Wolf, 2015). There are, however, some common approaches to the operationalisation of recidivism that many sources use. As criminal recidivism is a primary outcome examined in this thesis, I had to carefully approach the problem of its definition to ensure maximum generalisability of the results. Thus

Chapter 1 was fully dedicated to the review of criminal recidivism definitions and rates.

Mortality

Mortality is defined as a 'number of deaths in a given group in a certain period of time' (National Cancer Institute, 2021). The crude death rate is an omnibus measure that serves as a generalised proxy for health outcomes in a given population. Cause-specific mortality allows estimating the contribution of individual causes to overall mortality. In 2019, the three leading causes of death in the world were ischaemic heart disease, stroke, and chronic obstructive pulmonary disease (World Health Organization, 2021). Suicide was the 17th leading cause of death.

The mortality rates are high among correction populations, with most death attributable to external causes (Chang, Lichtenstein, et al., 2015; Phillips et al., 2018). External causes of mortality (ICD-10, chapter XX) include homicides, self-harm, complications of medical interventions, traffic and other accidents. Thus, most deaths that occur in released prisoners and individuals given community sentences are potentially preventable. Subsequently, identifying the risk factors for mortality, especially external-cause mortality, and designing targeted interventions will substantially reduce deaths among correctional populations. For prisoners, this apparent practice need was recognised, as there is an extensive body of research examining risk factors for prisoners' mortality (Borschmann et al., 2020; Skinner & Farrington, 2020; Zhong et al., 2021). However, the research examining the risk factors for mortality in individuals given community sentences is limited. The current thesis aimed to bridge this gap and examine the risk factors for all-cause and external cause mortality in individuals given community sentences, focusing on mental health factors.

Psychiatric disorders and crime

Psychiatric disorders are ‘a clinically recognizable set of symptoms or behaviours associated in most cases with distress and with interference with personal functions’ (World Psychiatric Association, 2011). The lifetime prevalence of any psychiatric disorder was estimated at 29.2% (Steel et al., 2014). This estimate implies that around a third of all individuals in the world at some point experienced symptoms that fit the established criteria of some psychiatric disorder. The most prevalent psychiatric disorders are anxiety disorder, mood disorders (depression and bipolar disorder), and substance use disorder. The lifetime prevalence of schizophrenia spectrum disorders was estimated at around 0.4% (Saha et al., 2005). The term ‘severe mental illness’ is often used to denote psychiatric disorders ‘resulting in serious functional impairment, which substantially interferes with or limits one or more major life activities’ (National Institute of Mental Health, 2021). Severe mental illnesses usually include schizophrenia spectrum disorders, bipolar disorder, and severe forms of depression.

Psychiatric disorders, especially substance use, have been systematically linked to violence and general criminal behaviour (Arseneault et al., 2000; Grann et al., 2008). In particular, severe mental illness has been often associated with an increased risk of adverse outcomes. Prior meta-analysis identified that individuals with psychotic disorders were two to three times more likely to commit a criminal offence compared to controls (Yee et al., 2020). Individuals with severe mental illness are also at higher risk of being victims of crimes themselves (Khalifeh et al., 2015). In addition, individuals with severe mental illness, substance use, and depressive disorders were shown to have substantially higher mortality rates than the general population (Chang et al., 2010).

Although the association between psychiatric disorders and criminal behaviour has been well established, the magnitude and nature of this association are a matter of debate. The magnitude of the association can vary between different populations. For example, a particular psychiatric diagnosis may have a stronger association with criminal conduct in sentenced individuals compared to the general population. There is also likely no single universal mechanism that links psychiatric disorders to criminal behaviour. As different psychiatric disorders have been associated with different behavioural tendencies, many potential symptom- and syndrome-specific pathways might exist between psychiatric disorders and criminal behaviour. A related question concerns the direction of causality along these pathways. It is often unclear whether psychiatric disorders have a causal impact on criminal behaviour or serve as indicators of some unobserved risk factors, such as socioeconomic deprivation. The uninformed assumptions about the causal role of psychiatric disorders in criminal behaviour can lead to the unsubstantiated criminalisation of mental illness (Dean et al., 2020). The resulting stigma and self-stigma can prevent an individual with a psychiatric disorder from meaningful participation in social interactions, work, and recreational activities. The resulting deprivation can itself trigger adverse behaviours, including criminality.

Psychiatric disorders are commonly found in community sentenced populations (Sirdifield, 2012). Therefore, understanding the impact of mental health and substance use disorders on reoffending is central to risk assessment and treatment approaches implemented within community sentencing. At the individual level, psychiatric disorders represent modifiable risk factors that can serve as targets for intervention. The qualification and quantification of their impact on reoffending would help to critically inform sentencing decisions and focus supervision support, preventing further criminalisation and imprisonment. At the populational level, decisions concerning the optimal design of community services commissioned to reduce reoffending would be informed by a clearer understanding of the nature of the association between individual psychiatric disorders and reoffending behaviours.

Aetiology

The diathesis-stress model is a psychological theory that has been applied to describe the aetiology of both psychiatric symptoms (Arnau-Soler et al., 2019; Chasiropoulou et al., 2019), delinquent behaviour (Benekos & Merlo, 2014; Schweder, 2003), and self-harm (O'Connor et al., 2010). The diathesis-stress model postulates the existence of a primarily genetic predisposition vulnerability (diathesis). According to the diathesis-stress model, individuals with high vulnerability are more susceptible to the influences of environmental stressors, such as adverse childhood experiences, when compared to resilient individuals. The interaction between the inherited vulnerability and environmental stress of sufficient intensity leads to the manifestation of psychiatric symptoms and delinquent behaviour, including criminal behaviour.

The life-course theory (Sampson & Laub, 2005) is another aetiological approach that can be used along with the diathesis-stress model to explain criminal behaviour. The theory also has been used to explain the aetiology of suicidal behaviour (Fazel & Runeson, 2020). The life-course theory emphasizes the critical role of individual developmental context. The theory postulates that individual susceptibility to environmental triggers is not constant over time; it changes over the life span of an individual. The effect of different environmental factors on criminal behaviour (or other outcomes) also varies with time. The effect of some factors, like adverse childhood experiences, may be present over the lifespan, whereas the effect of other factors is limited to adulthood. Over time, the combination of different effects can lead to cumulative disadvantage associated with persistently increased risk of criminal behaviour or other adverse behaviours (Sampson & Laub, 1997). For example, individuals who had experienced significant childhood adversities had, as adults, a higher frequency of alcohol intake, higher self-reported depression, and increased medication use (Bauermeister & Gallacher, 2018). All these factors have been associated with criminal recidivism and mortality in the correctional population (Chang, Lichtenstein, et al., 2015; Fazel et al., 2016; Molero et al., 2018).

Prior studies have demonstrated that different disorders in adult individuals are associated with different aetiological pathways to violence (Arseneault, Moffitt, Caspi, Taylor, & Silva, 2000; Retz & Rösler, 2009). The adverse behaviour itself can manifest in different ways depending on individual risk factors' effects. For example, in individuals with schizophrenia, the excessive perception of threat and history of the conduct disorder can explain violent behaviour. In individuals with alcohol use problems, intoxication was the most useful predictor of violence.

In observational studies, the implementation of specialised designs allows to partially separate the effect of genetic vulnerability from the effect of the individual risk factors (D'Onofrio et al., 2006). In this thesis, I used sibling comparison in my cohort studies to account for unmeasured factors shared between siblings discordant by a given diagnosis. The resulting estimates partially accounted for shared genetic makeup between siblings and quantified the independent effect of a manifested psychiatric disorder on criminal recidivism and mortality.

Overall, the diathesis-stress model and life-course theory are useful theoretical frameworks for examining the effect of psychiatric disorders on criminal recidivism and mortality. Although the thesis does not explicitly rely on a selected theoretical framework, the two described theories shaped the conceptual basis for the interpretation of the thesis' findings.

Aims and overview

The thesis aimed to examine the role of psychiatric disorders as factors for criminal recidivism and mortality in individuals given community sentences and to create a simple, scalable tool for criminal recidivism risk monitoring in this population. To achieve these aims, I conducted five separate studies. Each chapter presents the result of one study.

In Chapter 1, I explored the definitions of criminal recidivism and conducted a systematic review of its prevalence in community sentenced populations in different countries. This study laid the groundwork for further investigations, as it helped to formulate a more generalizable recidivism outcome measure. It additionally provided the general comparative context to the prevalence of recidivism in other studies.

In Chapter 2, I conducted a meta-analysis of risk factors for criminal recidivism and examined the reported associations between mental health risk factors and recidivism. This study outlined the current research context of psychiatric disorders as modifiable risk factors for criminal recidivism. This study overviewed common risk factors of recidivism that can be used in prediction models. The study also provided comparative context for the magnitude of association between psychiatric disorders and criminal recidivism in the following study cohort.

In Chapter 3, I explored the association between psychiatric disorders and criminal recidivism in a national cohort of individuals given community sentences in Sweden from 1991 until 2013. Using sibling comparison, I also explored the association between psychiatric disorders and criminal recidivism while controlling for unmeasured familial confounding. The study additionally reported the effect of having a new psychiatric diagnosis after a sentence on reoffending.

In Chapter 4, I explored the association between psychiatric disorders and mortality in the same national cohort. This study examined the effect of psychiatric disorders on potentially preventable deaths of individuals given community sentences. It also explored all-cause mortality as a proxy measure for adverse health outcomes in this population.

In Chapter 5, I derived and validated a risk assessment instrument to monitor criminal recidivism risk during the post-sentence community supervision. The tool can be used to assist the risk management of supervised individuals with psychiatric diagnoses. This study also explored the role of proper calibration of dynamic prediction tools, arguing that their calibration over time should become standard practice.

Finally, in the general discussion, I provided an overview of the findings, discussed their implications and relevance, and outlined the directions for future research.

Chapter 1. RECIDIVISM RATES IN INDIVIDUALS RECEIVING COMMUNITY SENTENCES: A SYSTEMATIC REVIEW

1.1 Abstract

I aimed to systematically review recidivism rates in individuals given community sentences internationally. I also explored sources of variation between these rates and reporting practices that limit their comparability across jurisdictions.

I searched MEDLINE, PsycINFO, SAGE and Google Scholar for reports and studies of recidivism rates using non-specific and targeted searches for the 20 countries with the largest prison populations worldwide. I identified 27 studies with data from 19 countries. Of the 20 countries with the largest prison populations, only 2 reported recidivism rates for individuals given community sentences.

The most reported recidivism information between countries was for 2-year reconviction, which ranged widely from 14% to 43% in men, and 9% to 35% in women. Explanations for recidivism rate variations between countries include variables such as when the follow-up period started and whether technical violations were included.

Recidivism rates in individuals receiving community sentences are typically lower in comparison to those reported in released prisoners, although these two populations differ in terms of their baseline characteristics. Direct comparisons of the recidivism rates in community sentenced cohorts across jurisdictions are currently not possible, but simple changes to existing reporting practices can facilitate these. I updated existing recidivism reporting guidelines to include the proposed changes.

1.2 Introduction

This chapter presents the systematic review findings that lay the groundwork for other studies in the thesis. In this systematic review, I explored different approaches to the operationalisation of criminal recidivism, which is the primary outcome examined in the thesis. Understanding how different jurisdictions define and measure criminal recidivism is necessary to formulate the most generalisable outcome measure to use in further studies.

As outlined in the general introduction, community sentences require careful evidence-informed implementation to achieve their goal of reducing recidivism rates in sentenced individuals. This evidence-informed implementation often depends on the ability of researchers and policymakers to compare recidivism rates across jurisdictions. Such comparisons, however, can be problematic. For example, the variability of operational definitions of criminal recidivism is a primary factor preventing the international comparison of recidivism rates in released prisoners (Fazel & Wolf, 2015). Community sentences, being much more heterogeneous than custodial sanctions, could potentially add another layer of complexity to the comparison of recidivism rates. Understanding the variability of the criminal recidivism outcomes in individuals given community sentences is essential for comparative research or correctional programmes.

The most common question that requires comparing recidivism rates is whether community sentences are more effective than imprisonment. To assess the impact of community sentences on recidivism, different methods to account for confounding, including the choice of matched controls, can be employed. This is important for validity since direct comparisons of cohorts receiving different sentences are often complicated by selection biases, such as higher risk individuals' likelihood of receiving longer sentences. To account for potential selection biases, some studies have utilised the randomised allocation of sentenced individuals to an alternative sanction (Pearson et al., 2014), different judges (Harding et al., 2017; Nagin & Snodgrass, 2013), and different training programmes (Lowenkamp et al., 2014). However, conducting randomised trials is often not possible in a criminal justice setting. Another approach to address confounding is to compare matched cohorts of individuals receiving different sentences. Propensity score matching, which is widely used, is based on the predicted probability to reoffend (Bales & Piquero, 2012; Evans et al., 2014; Jolliffe & Hedderman, 2015; Trevena & Weatherburn, 2015), whereas precision/exact matching is based on a set of predetermined individual characteristics (Bales & Piquero, 2012). They may also be used in combination (Wermink et al., 2010). Finally, there is a relatively new approach to using machine learning algorithms to sample stratification and allocation of individuals (Hyatt & Barnes, 2017).

Although randomised controlled trials and matched cohort designs will generate more robust findings, they are often costly, and it is often not possible to utilise these methods when comparing the impact of systemic approaches to community sentencing on recidivism across different jurisdictions (especially between countries) because of different methods of data acquisition and legal thresholds for community sentences. The standardisation of reporting practices will allow the comparisons based on routinely reported data to be as close to matched cohort designs as possible. In the previous review conducted by our research group, the group demonstrated that, for national samples of released prisoners, such comparisons are currently difficult due to differences in reporting practices, definitions of outcomes, and follow-up period length (Fazel & Wolf, 2015). Reporting recidivism in community sentenced populations may present additional challenges because, besides reconviction, rearrest and reimprisonment, recidivism reports for individuals receiving non-custodial measures may utilise other types of outcomes and follow-up schemes.

To provide a current picture of the recidivism information in community sentenced populations and to examine reporting practices, I sought to systematically review studies of recidivism rates in individuals receiving community sentences and examine possible explanations for variation in such rates. Based on my findings, I have updated the previously published recidivism guidelines to improve comparability.

1.3 Approach

Systematic reviews allow for comprehensive description and analysis of publications on a particular topic. In contrast to a scoping literature review, a systematic review is a structured approach to the review process that follows a pre-specified replicable plan. This approach reduces the reporting bias and improves the overall quality of information synthesis. There are published guidelines a researcher can follow to ensure the high quality of the conducted systematic review. Cochrane Book for Systematic Reviews of Interventions (Higgins et al., 2019) is currently the golden standard for systematic reviews in healthcare. Although I did not conduct a systematic review of interventions in this thesis, I still relied on Cochrane Book for the general guidance regarding systematic reviews and meta-analyses.

To ensure comparability across different systematic reviews, the reporting checklists are widely used. A reporting checklist helps to produce high-quality publications that are easily comparable with each other. PRISMA is one of the most commonly used checklists in healthcare research. The release of PRISMA in 2009 led to significant improvement in systematic review quality over the following decade (Sun et al., 2018).

Overall, the systematic review process involves several stages: formulation of a research question, specification of inclusion and exclusion criteria, designing a search strategy, screening relevant publications, quality assessment, and data extraction. If the extracted data from the studies are comparable, then a meta-analysis could be conducted to pool the estimates of many smaller studies together, simulating the effect of a larger study. However, the extracted data in the current systematic review were not directly comparable, and a meaningful meta-analysis was not possible.

1.4 Methods

The review protocol was registered in PROSPERO (CRD42018088156). I followed the design of the previous review of recidivism rates among released prisoners conducted by our research group (Fazel & Wolf, 2015). I searched MEDLINE, PsycINFO and SAGE publication databases using search terms in relation to community offenders and recidivism (see Appendix A1 for search terms). In addition, Google Scholar and Google Web were used for targeted searches for the 20 countries with the largest prison populations worldwide (World Prison Brief, 2018). I also reviewed the reference lists of the included publications.

I included cohort studies of the general population of adult offenders receiving community or suspended sentences with no follow-up period restrictions. If there were multiple reports for one country, I used national data from the most recent report. Regional data (i.e., reports from provinces, states or cities) were used when national data were unavailable. Authors of the studies were contacted when necessary.

I excluded studies that focused on specific subpopulations (e.g., sex offenders, mentally disordered offenders). I also excluded studies of individuals on parole after serving a prison sentence and of individuals specifically sentenced to undergo mental health and/or substance abuse treatment. Heterogeneous samples that also contained released prisoners or adolescents were excluded. In addition, I excluded matched cohort studies that compared community sentenced individuals with individuals released from prison, as matched cohorts were not representative of the general population.

I extracted information on the rates of general recidivism, violent recidivism, non-violent recidivism, and violation of probation conditions after the imposition of a community sentence. No outcome measurement restrictions were applied. Data extracted included country or territory, year of selection, index offence disposal, follow-up period, reported outcomes, and base rates of reported outcomes. Three researchers were involved in the screening process: Seena Fazel, Achim Wolf, and myself. I conducted the initial search on the date specified, using the databases and search strategy listed above, and screened titles and abstracts. In addition, Achim Wolf and I carried out a targeted search of recidivism reports using governmental websites of the countries of interest. Any uncertainties were resolved by discussion between the three researchers. Potential differences between the recidivism rates of men and women were examined using relative risk ratios, calculated according to Altman (1990). No meta-analysis was conducted because of high heterogeneity in sample compositions and outcome definitions across included studies.

PRISMA guidelines (Moher et al., 2009) were followed (Appendix A2).

1.5 Results

I identified 28 studies reporting recidivism rates in individuals receiving a community sentence from 19 countries (Figure 1-1, Appendix A3). Of the 20 countries with the largest prison populations, recidivism reports were identified for the USA and UK. The data were mostly reported by governmental agencies; however, five identified papers were published in journals (Bartels, 2009; Flores et al., 2017; Harding et al., 2017; Ķipēna et al., 2013; Leonardi, 2007). The results are provided separately for reconviction (Figure 1-2; Table 1-1 and Table 1-3) and rearrest rates (Table 1-2 and Table 1-3) for follow-up periods of one, two, three, five, and seven years. Only 11 identified reports and studies reported recidivism rates separately for men and women (Table 1-3).

The composition of the identified cohorts varied in terms of the sentencing and demographic characteristics of included offenders. Many published reports were excluded because they used heterogeneous samples of offenders such as released prisoners and probationers, or adolescents and adults, without providing recidivism rates for the subgroups (e.g., Federal Statistical Office (2015)). Equally, we excluded reports that provided recidivism statistics using cross-sectional data (e.g., Federal Penitentiary Service of Russia (2018)). Several reports included fines as a community sanction, and their inclusion influenced reported recidivism rates (Figure 1-3).

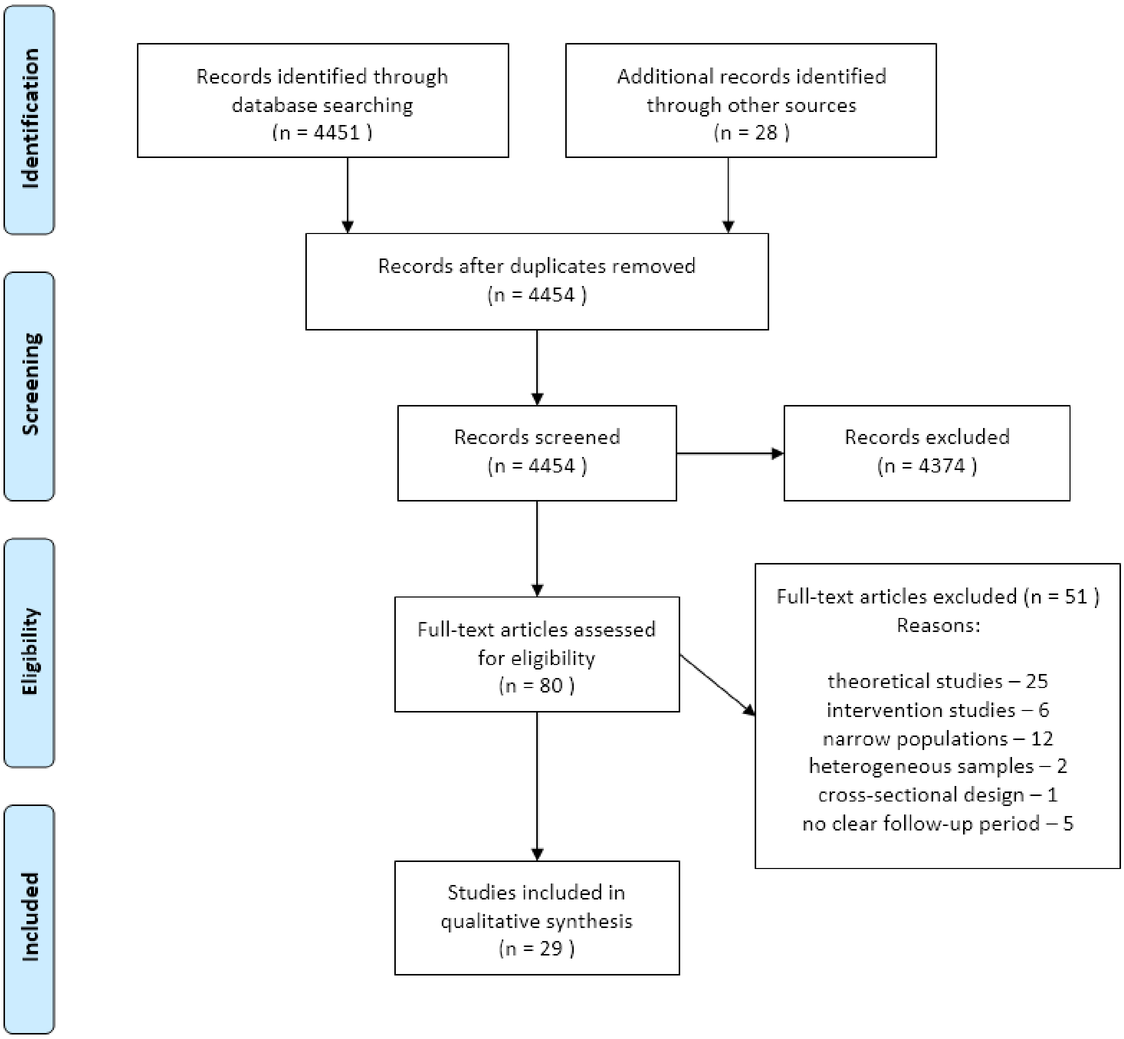


Figure 1-1. PRISMA 2009 Flow Diagram

Most of the included studies used representative cohorts of adult offenders (aged 18 and older) receiving a community sentence. However, the included Danish data were for offenders aged 20 and older (since the lower reported age tier was 15-19 years old), and the Republic of Ireland data did not include sex offenders (see Appendix A3). Several included reports and studies (from Latvia, New Zealand and Australia) did not provide cohort sizes, and they could not be estimated from other sources.

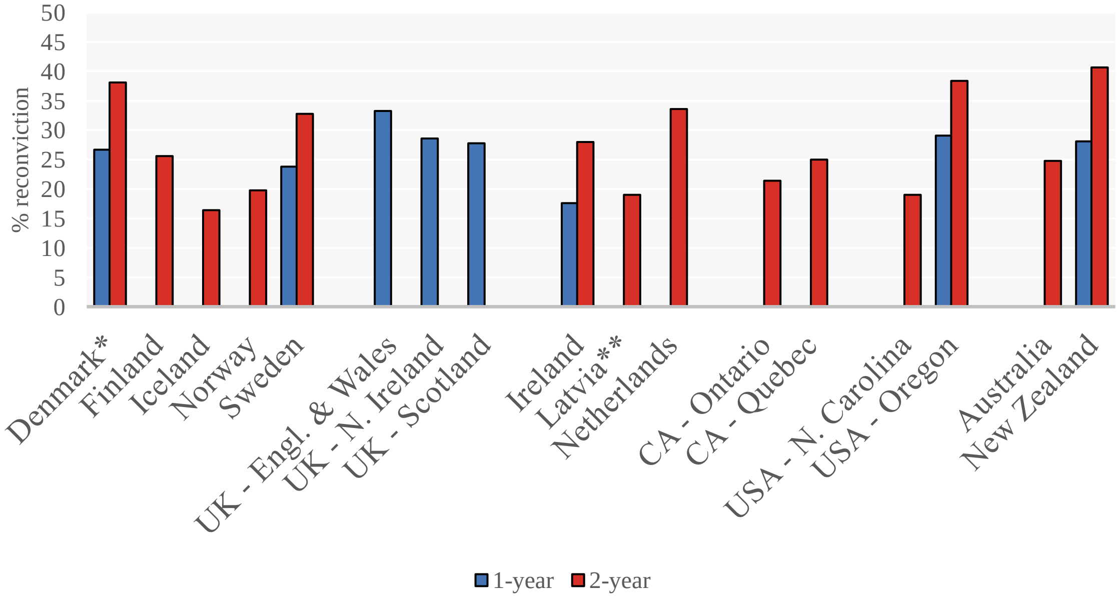


Figure 1-2. Reconvictions rates in adult individuals receiving community sentences for 1-year and 2-year follow-up periods.

For sources, refer to Appendix 1. *The reconviction rate for Denmark was reported for individuals aged 20 and older. All other cohorts include individuals aged 18 and older. **Follow-up for Latvia was 29 months.

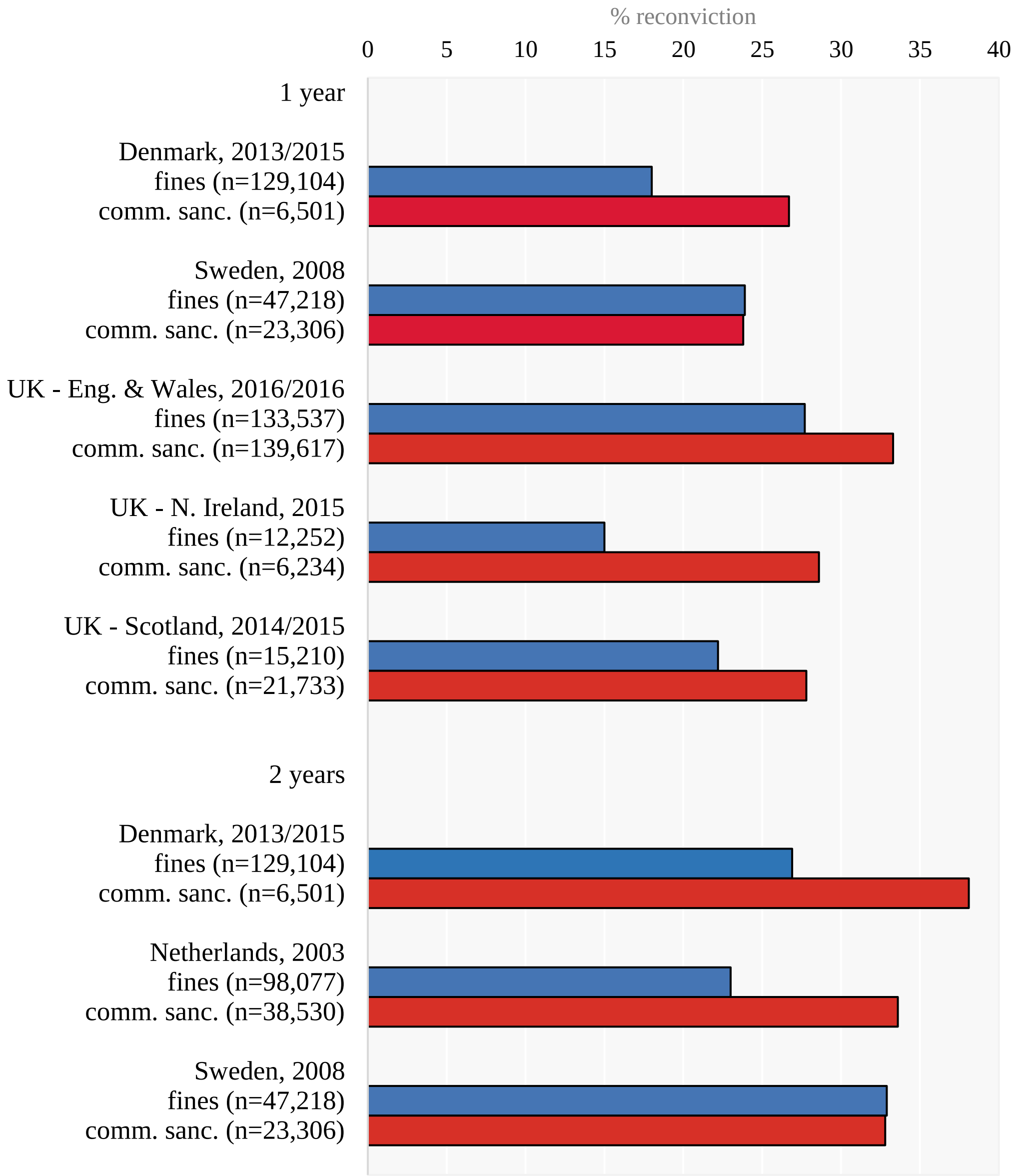


Figure 1-3. Comparison of reconviction rates in adult individuals receiving fines or sentenced to other community sanctions for 1-year and 2-year follow-up periods.

For Denmark, the data reported only for fines larger than DKK 2,500 for road traffic offences and DKK 1,000 for most other offences. In Scotland when an individual receives several sentences (i.e., community supervision and fine) only the most serious sentence is accounted for, which is not the case in other countries.

Table 1-1. Reported reconviction rates (%) for cohorts of adult individuals receiving community sentences

*The reconviction rate for Denmark was reported for individuals aged 20 and older. All other cohorts include individuals aged 18 and older. **Follow-up for Latvia was 29 months.

			Length of follow-up period (years)				
Country	**Selection period**	**Cohort size**	**1**	**2**	**3**	**5**	**7**
Europe							
Nordic countries							
Denmark	2013	6,501	26.7	38.1			
Finland	2005	3,767		25.6			
Iceland	2005	73		16.4			
Norway	2005	2,839		19.8			
Sweden	2008	22,306	23.8	32.8	38.1		
The United Kingdom							
England and Wales	2015/2016	139,617	33.3				
Northern Ireland	2014/2015	6,234	28.6				
Northern Ireland	2005	4,425		26.1			
Scotland	2014/2015	21,733	27.8				
Other							
Austria	2004	93,073			38.1		
Germany	2007	96,521			39.0		
Italy	1998	8,817					19.0
Ireland, Republic of	2010	3,698	17.6	28.0	36.4		
Latvia	2009	n/a		19.0			
Netherlands	2003	38,530		33.6			
North America							
Canada							
Ontario	2013/2014	35,561		21.4			
Quebec	2007/2008	4,851		25.0			
USA							
North Carolina	2013	35,103		19.0			
New York State	2002	31,267				33.0	
Oregon	2014	4,403	29.1	38.4	43.8		
South America							
Chile	2007	23,736			27.7		
Oceania							
Australia (federal)	2012/2013	n/a		24.8			
New Zealand	2013/2015	n/a	28.1	40.7			

Table 1-2. Reported rearrest rates (%) for cohorts aged 18 and older

Country	Selection period	Cohort size	Length of follow-up period (years) 1	2	3	4	5	7
<u>North America</u>								
USA								
USA (federal)	2004/2005	13,504	13.1	22.2	29.0	34.6	38.8	44.9
Illinois	2006	2,770					54.0	
North Carolina	2013	35,103		38.0				
Oregon	2014	4,403	34.8	44.7	50.4			
<u>South America</u>								
Chile	2007	23,736			40.6			

Table 1-3. Reconviction and rearrest rates in adult men and women receiving community sentences

Relative risk ratios were calculated according to Altman (1991). Reported data for New Zealand, Australia (federal, New South West and Western Australia) did not allow for relative risk calculation.

Country	Selection period	Follow-up		Sample size	%	Relative risk (95% CI)
<u>Reconviction rates</u>						
Australia (federal)	2012/2013	2 year	Men Women	n/a	43.4 31.1	n/a
AUS – New South Wales	2015	1 year	Men Women	13,744 3,163	20.9 18.1	1.15 (1.06-1.25)
AUS – Western Australia	2012/2013	2 years	Men Women	n/a	13.9 9.3	n/a
CA - Quebec	2007/2008	2 years	Men Women	4,010 830	26.0 21.0	1.24 (1.08-1.43)
Chile	2007	3 years	Men Women	20,399 3,337	27.4 27.8	0.99 (0.93-1.05)
Denmark	2013	1 year	Men Women	5,413 1,088	32.6 18.2	1.80 (1.57-2.04)
Denmark	2013	2 years	Men Women	5,413 1,088	40.7 25.2	1.61 (1.45-1.80)
Ireland, Republic of	2010	1 year	Men Women	3,241 457	18.0 14.7	1.23 (0.97-1.55)
New Zealand	2014/2015	1 year	Men Women	n/a	30.0 21.4	n/a

UK - Northern Ireland	2005	1 year	Men Women	16,233 2,814	20.8 10.4	2.00 (1.78-2.24)
USA – North Carolina	2013	2 years	Men Women	25,850 9,253	21.0 14.0	1.50 (1.42-1.59)
USA – Oregon	2015	1 year	Men Women	3,102 1,231	27.6 27.0	1.02 (0.92-1.14)
USA – Oregon	2014	2 years	Men Women	3,200 1,203	39.8 34.6	1.15 (1.05-1.26)
USA – Oregon	2014	3 years	Men Women	3,200 1,203	45.5 39.3	1.16 (1.07-1.25)
Rearrest rates						
USA – North Carolina	2013	2 years	Men Women	25,850 9,253	41.0 29.0	1.41 (1.37-1.46)
USA – Oregon	2015	1 year	Men Women	3,102 1,231	34.2 34.2	1.00 (0.91-1.10)
USA – Oregon	2014	2 years	Men Women	3,200 1,203	46.7 39.3	1.19 (1.10-1.29)
USA – Oregon	2014	3 years	Men Women	3,200 1,203	52.7 44.2	1.19 (1.11-1.28)

The recidivism outcome and its operationalisation varied significantly between countries and even between territories within one country (see Appendix A4). The most commonly used outcome was reconviction with a follow-up period of one or two years. To register a reconviction, some countries allowed additional time after a declared follow-up period for the offence to be proven in court (e.g., England and Wales, Northern Ireland, Sweden, and others). Other countries used the initiation of legal proceedings, if there was no acquittal, as an outcome event (e.g., Latvia, the Netherlands). In addition, a follow-up period could start at the beginning or the end of a community sentence. In the latter case, a violation of probation conditions was not counted as a recidivism event. It was often unclear how offenders that had received multiple sanctions for one offence were counted in the reports. Some reports considered each sanction separately (thus, one offender could be counted several times), while others only included the most serious sanction for one offender. Other reports did not clarify this matter. Several reports on probation outcomes (from Australia, the USA, and Singapore) were excluded from the analysis, as they did not have fixed follow-up periods and only followed offenders during the length of the sentence (which may have varied).

The reported general recidivism rates varied significantly between countries and regions (Figure 1-2; Table 1-1 and Table 1-2). For example, the one-year reconviction rate ranged from 5.5% in Michigan, USA to 33.3% in England and Wales. Two-year reconviction rates varied from 16.4% in Iceland to 40.7% in New Zealand. Rearrest rates were reported by several US states with notable differences (22% for federal probationers to 45% in Oregon during 2-year follow-up).

In addition to general recidivism, several countries reported specific recidivism (committing a new crime similar to an index offence), and only two (Denmark and the Republic of Ireland) provided data that allowed for the calculation of different types of recidivism for community sanctioned populations based on the type of new offence. For example, in the Republic of Ireland, reconviction for violent crimes (which included homicides, sexual offences, attempts/threats of murder, assaults, harassment, kidnapping, robbery/extortions) accounted for only 5% of all recidivism cases. In Denmark, violent crimes also included crimes against public order and constituted 11% of all reconviction cases.

Of all the included reports and studies, only reports from Quebec (Ministère de la Sécurité publique, 2015), North Carolina (Flinchum et al., 2016) and Michigan (Harding et al., 2017) separately reported data on technical violations of community sentences. In North Carolina, technical violation automatically led to up to 90 days of imprisonment, whereas in Quebec a breach did not necessarily lead to incarceration. In Michigan, reimprisonment after technical violations was examined as a separate outcome.

Only 10 identified reports and studies provided recidivism data separately for men and women (Table 1-3). Most of them indicated that female offenders had a lower risk of recidivism than male offenders, but with noticeable variability. For example, the absolute risk difference for one-year reconviction rates varied between 0.6% in Oregon, USA and 17.4% in Denmark.

1.6 Discussion

In the present study, I identified 28 reports from 19 countries that reported reoffending rates in individuals given community sentences at the point of court disposal. Only two of these were from the 20 countries with the largest prison populations. Overall, reported 1-year reoffending rates varied between 5% and 33%, and 2-year rates ranged from 16% to 41%. Such recidivism rates are lower than those observed in released prisoners (Fazel & Wolf, 2015). Women, in general, have a lower risk of recidivism when serving a community sentence than men. This is consistent with findings in released prisoners (Spjeldnes & Goodkind, 2009). However, there are exceptions. Many studies were excluded from the review because they reported recidivism rates only for released prisoners or for a mixed sample of individuals who received community and custodial sentences. Although community sentenced individuals constitute the largest population of those receiving criminal sanctions, this review suggests that reporting practices concerning their outcomes need improvement in terms of coverage and detail.

To examine the findings further, I have undertaken a structured search in the last five years of studies in offenders under community supervision with randomised or matched control groups (Appendix A5). I found that community sentences were typically associated with lower recidivism rates than custodial sentences, particularly in those individuals with low to moderate risk levels. In addition, this overview showed that specialised community programmes were more likely to benefit individuals in certain offender groups (such as DWI offenders) in relation to reoffending outcomes. Another finding from these research studies is that they were influenced by the type of model used for matching and the definition of outcome (re-arrest vs reconviction vs reimprisonment).

Overall, the wide variation in recidivism rates between countries was expected. The same factors play a role with individuals given community sentences as with released prisoners, such as differences in judicial practices, and definitions and operationalisations of recidivism measures. Variations in laws and precedent practices across jurisdictions result in cohorts with different proportions of individuals convicted for different types of offences. For example, a conviction for the same drug-trafficking offence may lead to a suspended sentence or probation in one jurisdiction or to 5–10 years in prison in another (European Monitoring Centre for Drugs and Drug Addiction, 2017). Because of this variability, sentenced cohorts may not be equivalent in terms of their initial recidivism risk.

There are, however, some additional issues that contribute to recidivism rate variability in community offenders. First, the set of actual sanctions that comprise community sentences is different depending on the legal system. There are many different combinations of community service work, probation with supervision, treatment orders, suspended sentences, and fines. Community sentences may also include house arrest or night-time curfews. Different combinations of sanctions will lead to variations in recidivism rates. One notable example is the inclusion of fines

(Figure 1-3). In Northern Ireland, the inclusion of fines as an index sanction decreased the reported 2-year reconviction rate from 26% to 19%. Moreover, countries use different rules for the inclusion of fines. In Sweden and the Netherlands, fines and compensatory orders issued by courts and the prosecution office, but not the police, are reported. In Northern Ireland, England and Wales, fines that result from court convictions are reported. In Scotland, the data for “monetary disposals” are reported, which only includes fines issued by courts. Often, the operational definitions of sanctions are not given, and recidivism data are not reported by separate types of sanctions. Providing recidivism data by sanction type may allow comparison between groups of similarly sentenced individuals and in turn substantially increase overall comparability between reports.

Second, there are many different approaches to defining the starting point of a follow-up period in individuals given community sentences. In many reports, the distinction between a follow-up period and a period of supervision was not always clearly described and accounted for. For example, a follow-up period might start after the completion of community service, i.e., after the end of supervision. This approach ignores a significant amount of time when a sentenced individual is at risk of committing a new offence. This may also lead to underreporting of recidivism because offenders who have violated the conditions of community sentences during the time of supervision and subsequently been incarcerated, are excluded from observation. Another definition for the time at risk could be a follow-up period that matches a period of supervision. Some countries and regions (e.g., New York State) may have a fixed period of supervision (5 years) that matches the period of follow-up. In this case, reported recidivism rates are virtually the same as sentence completion rates. If there is no set period for supervision and only completion rates are reported, then the follow-up time period is unclear. In addition, the follow-up period may start with the beginning of supervision, partially overlap with it and extend beyond it. The clear separation of sentence completion rates and post-supervision recidivism rates will enable improved between-jurisdiction comparisons. The report from Quebec (Ministère de la Sécurité publique, 2015) illustrates the importance of distinguishing between sentence completion and reconviction during a follow-up since reported recidivism rates vary from 25% to 41% depending on the starting point employed. There is also a qualitative distinction between recidivism under supervision and recidivism during unsupervised community living, which is important to take into account, especially given that supervision periods may vary between offenders. In the context of community sentences, additional reporting of the rate of technical violations and reconvictions based on technical violations is important for international comparisons since the handling of such violations by courts and probation systems has a significant impact on outcomes.

Finally, there may be differences in the quality of supervision, availability and effectiveness of rehabilitation programmes, availability of vocational training, and access to healthcare services. These are factors that researchers and practitioners are keen to examine; however, because of differences in reporting practices, it is difficult to compare these factors internationally.

Reporting practices can be improved to facilitate international comparisons and to help to inform sentencing decisions. The general recommendation for recidivism reporting in any population is to provide a detailed breakdown of the sample by index offence type, socio-demographic characteristics and outcomes, including general and violent recidivism. We additionally propose two main recommendations specific to reporting recidivism in community sentenced populations. First, recidivism data should be additionally reported by types of sanctions (i.e., community service, electronic monitoring or mandatory treatment) with clear operational definitions of sanction types provided. Second, the completion rate and recidivism rate after the completion of supervision should be reported separately with a clear indication of supervision and follow-up lengths (this could also be applied to released prisoners on parole). Following these recommendations will enable flexible extraction of recidivism data for more meaningful comparisons across jurisdictions.

1.7 Strengths and limitations

To my knowledge, this is the first study to systematically review recidivism rates in a general adult population of individuals receiving community sentences. Studies included in the review were generally of high quality (as assessed by the NIH Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies) and were conducted using large samples.

The heterogeneity of sample compositions and outcome definitions did not allow for direct quantitative comparisons of recidivism rates between different studies. However, I was able to examine how the implementation of certain reporting practices influences the reported recidivism rates and what changes could be made to ensure that recidivism data is more comparable internationally in the future.

1.8 Conclusion

I conclude that recidivism rates in the community sentenced population vary considerably between countries, and, in most cases, different criteria for the time at risk and outcome are used. The most common recidivism outcome reported in individuals given community sentences was two-year reconviction. The recidivism rates are lower than those observed in released prisoners (Fazel & Wolf, 2015). The comparability of recidivism rates can be improved if more detailed information is provided and completion rates and recidivism after the end of supervision are reported separately. The identified methodological problems specific to recidivism reporting in community sentenced populations can be addressed by adjustments to existing reporting practices. I have published an updated version of the recidivism reporting checklist (Fazel, Wolf, & Yukhnenko, 2019), which builds on the previous review of general criminal recidivism (Fazel & Wolf, 2015) to enable consistency and transparency in recidivism rate comparison between countries. I hope that this will enable the development of more informed policy and judicial decision-making in the criminal justice system.

Chapter 2. Risk factor for recidivism in individuals receiving community sentences: A systematic review and meta-analysis

2.1 Abstract

I aimed to systematically review risk factors for criminal recidivism in individuals given community sentences. I searched seven bibliographic databases and conducted targeted searches for studies that investigated risk factors for any repeat offending in individuals who had received community (non-custodial) sentences. I identified 15 studies from 5 countries, which reported data on 14 independent samples and 246,608 individuals. Several dynamic (modifiable) risk factors were significantly associated with criminal recidivism in community sentenced populations, including mental health needs (OR=1.4, 95% CI: 1.2-1.6), substance misuse (OR=2.3, 95% CI: 1.1-4.9), association with antisocial peers (OR=2.2, 95% CI: 1.3-3.7), employment problems (OR=1.8, 95% CI: 1.3-2.5), marital status (OR=1.6, 95% CI: 1.4-1.8), and low income (OR=2.0, 95% CI: 1.1-3.4). The strength of these associations is comparable to that of static (non-modifiable) risk factors, such as age, gender and criminal history.

Assessing dynamic risk factors should be considered in all individuals given community sentences. The further integration of mental health, substance misuse, and criminal justice agencies may reduce reoffending risk in community sentenced populations.

2.2 Introduction

The previous chapter provided an overview of recidivism rates worldwide and explored the reasons for their variability. The goal of this chapter was to provide a comprehensive overview of the published studies of risk factors of recidivism with a focus on mental health risk factors.

Community sentences are the commonest type of court sanction in many countries (Ministry of Justice, 2018a; Pew Center, 2011; Statistics Denmark, 2018). They often employ offender management and rehabilitation programmes to prevent recidivism and the further criminalisation of individuals receiving community sentences (Landenberger & Lipsey, 2005; Visher et al., 2005). The ultimate goal of these programmes is to ensure public safety and to ease the economic burden on justice systems, they may be based on different conceptualisations of repeated criminal behaviours and employ different methods. The criminogenic needs of individuals (the characteristics of an individual that directly relate to the likelihood of recidivism) are typically broken down into non-modifiable (static) and modifiable (dynamic) risk factors Non-modifiable risk factors are unchanging characteristics of an individual and include gender, age, prior criminal history, and others. Modifiable risk factors are items that can be influenced or changed during the process of rehabilitation such as employment problems or substance misuse.

Both non-modifiable and modifiable risk factors are considered during risk assessment and intervention planning (Bonta & Andrews, 2007). Non-modifiable risk factors are strong predictors of future offending behaviour but are, by definition, poor targets for intervention. Moreover, a common issue with many risk assessment approaches is their overreliance on non-modifiable risk factors, and a failure to take time and change into account may result in biased estimations of an offender's recidivism risk (Hanson, 2018). Taking into account modifiable risk factors and their change over time may improve the accuracy of risk assessment (Clarke et al., 2017). It is also important to study modifiable risk factors for recidivism in community-sentenced populations separately from released prisoner populations. For individuals serving a community sentence, certain risk factors may be more or less predictive than in released prisoners or operate through different pathways.

However, many individual studies that examine risk factors for recidivism in community-sentenced populations, especially on a national level, focus exclusively on non-modifiable risk factors, typically offenders' demographics and prior contact with the justice system (Central Statistics Office, 2016; Ministry of Justice, 2018a; Swedish National Council for Crime Prevention, 2017). This may be especially problematic given that, when assessed using standardized diagnostic tools, community sentenced populations show a higher prevalence of modifiable risk factors such as psychiatric disorders and misuse of illicit substances (Lurigio et al., 2003) in comparison to the general population. In addition, prior meta-analyses that have investigated risk factors in community-sentenced populations either examined mixed samples of released prisoners and community-sentenced individuals (Gendreau et al., 1996; Olver et al., 2014) or looked into narrow subpopulations of community-sentenced individuals, e.g., sexual offenders (Hanson & Morton-Bourgon, 2005) or offenders in forensic psychiatric treatment (Bonta et al., 2014).

The present study examined both non-modifiable and modifiable risk factors for recidivism in 225,261 individuals receiving community sentences worldwide.

To my knowledge, this study is the first meta-analysis that examines risk factors for criminal recidivism in the general adult community-sentenced population.

2.3 Approach

This chapter followed the general approach outlined in Chapter 1. However, in this study, it was possible to compare the results across identified studies using meta-analysis quantitatively. Meta-analysis is conducted after the estimates from studies identified in a systematic review have been extracted (Higgins et al., 2019). To ensure comparability, all estimates are converted to the same metric. The heterogeneity (i.e., variation in study outcomes) is then accessed. If measured variation exceeds what is expected from the measurement error, the estimates have high heterogeneity.

Then a weighted average is computed across all estimates from identified studies. The weight assigned to individual study estimates depends on the corresponding study's sample size and the model used to pool them (Liu, Tian, Lee, & Xie, 2018). If heterogeneity between the individual studies' estimates is low, then the fixed effects model is usually used to pool them. The fixed model assumes that the individual studies' estimates have been drawn from the same distribution. If heterogeneity is high, the estimates are pooled with a random-effects model that assumes that each estimate has been drawn from a different distribution. The random effects model usually provides more conservative pooled estimates of the effect of interest.

Meta-analyses also allow identification of publication bias for a given pooled effect, i.e., selective reporting of 'positive' findings for the smaller studies while 'negative' findings remain unpublished. The presence of publication bias could lead to overestimating the effect of interest by researchers and practitioners, so it is important to test for its presence. Pooling data across the set of studies with significant publication bias will likely result in a biased estimate of the effect of interest unless the correction is applied (Andrews & Kasy, 2019).

The simple way to test for possible publication bias is to plot the results on a funnel plot. The asymmetric distribution of the plotted effects will indicate the potential publication bias. Egger's test is then applied to test the significance of the observed asymmetry (Egger et al., 1997).

2.4 Methods

The systematic review protocol was pre-registered in PROSPERO (CRD42018099606) and PRISMA guidelines (Shamseer et al., 2015) were followed (Figure 2-1; Appendix B1).

2.4.1 Search strategy

Publication search with no time or language restrictions used the following databases: MEDLINE, SAGE, JSTOR, PsycINFO, PsycARTICLE, EMBASE, Global Health. Search terms consisted of (recidivism OR "re-offending" OR reoffending OR rearrest OR "re-arrest") AND (risk OR predictor OR need) AND (criminogenic OR modifiable OR dynamic) AND ("community service" OR probation OR "community sentence"). We scanned the reference lists of the screened-in articles to identify new studies. In addition, the Google Scholar "cited by" tool was used to identify additional studies. Key investigators with relevant publications were contacted to determine if they had any new or missed studies.

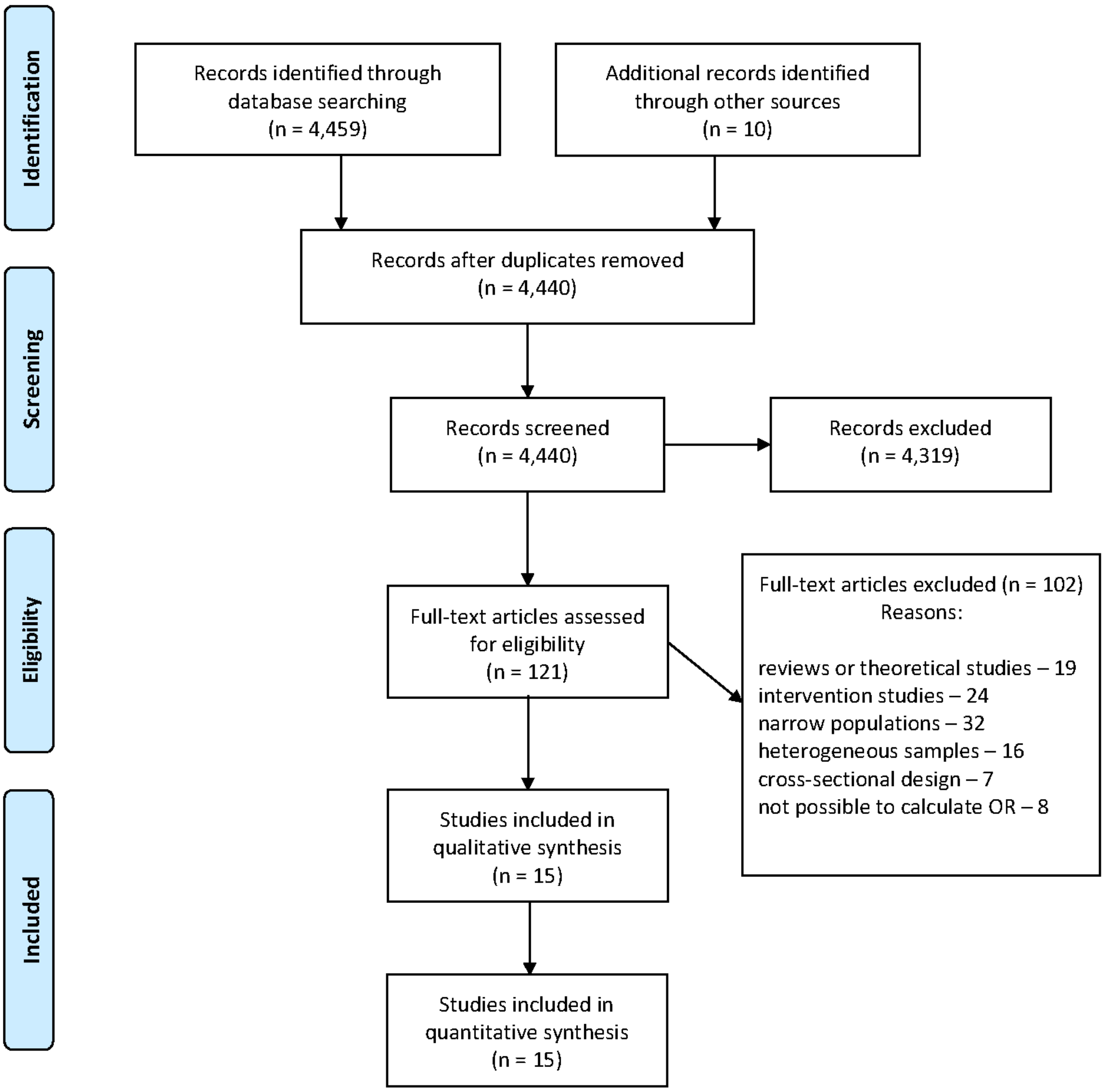


Figure 2-1. PRISMA flowchart

2.4.2 Study eligibility

I included studies of individuals from the general adult (≥ 18 y.o.) population given community sentences. Individuals released to community supervision after serving a prison sentence (parolees) were excluded. To be included, a study had to contain data that enabled the estimation of odds ratios for at least one risk factor. We excluded studies conducted in narrow subpopulations of individuals given community sentences (e.g., only adolescents, only women, only people with psychiatric disorders), cross-sectional studies, studies of interventions, and validation studies for risk assessment tools. There were no exclusions based on the reported recidivism outcomes, which might include any reoffending, violent and non-violent reoffending, re-arrest, revocation of probation, or technical violations.

2.4.3 Data extraction

The data extraction process happened in two stages. Standardised forms were used for each stage, and several variables were pre-specified for later subgroup analysis. First, for each study, I extracted the year of publication, study design, geographical region, coverage (province, state, country), sample characteristics (number of individuals, selection year, reported outcomes, number of people with reported outcomes, type of follow-up, the length of the follow-up period, gender composition, mean age), and the list of all risk factors. Second, if at least three studies examined a risk factor, the following data were extracted: number of individuals in the exposed and comparison groups, operationalisation of risk factor in a particular study, description of a comparison group, source of information (records, risk assessment instrument). Risk factors judged to be similar by their descriptions in papers were collapsed in risk factor domains. If a study reported multiple outcomes, the most prevalent outcome for a particular risk factor was extracted to enhance comparability. The most serious outcome was used when the prevalence for two outcomes was the same in a group (in order of priority: reconviction, probation failure, re-arrest, technical violation). When a study used the same dataset as another study for a given risk factor, the data from the most recent study were extracted. I extracted the data, then my colleague, Howard Ryland, independently repeated data extraction (thank you, Howard!). Any disagreements were resolved by discussion with Seena Fazel.

Several studies that explored data on substance misuse reported it separately for alcohol and drug abuse without providing joint data for any substance misuse. Taking this into account, to avoid duplicating samples, we analysed individuals with substance abuse problems by three subgroups (problems with alcohol, problems with drugs, and problems with substance misuse in general). In addition, different studies reported risk for ethnicity domain in different ways. Race and ethnicity might be defined as one or as two separate categories. We used data comparing white and non-white individuals, which was the most common way of reporting risk for this domain.

The data were converted to odds ratios for pooling. If a study reported frequencies or proportions, crude odds ratios were calculated directly with corresponding 95% confidence intervals. If no such data had been reported, we used other metrics that allowed the estimation of odds ratios. If crude odds ratio estimation was not possible, adjusted odds ratios were extracted. Reported chi-square values were converted into Cohen's d and, consequently, into log-transformed odds ratios (Rosenthal & DiMatteo, 2001). All odds ratios were reported to one decimal place.

Quality assessment was performed using the Newcastle-Ottawa Quality Assessment Scale for Cohort Studies (Peterson et al., 2011). This scale evaluates cohort selection, exposure ascertainment, comparability between cohorts, and the quality of outcome measurement. For each item on the scale, the study can be awarded one or two stars, with a maximum score of 9 stars. Any uncertainties about quality rating were resolved by discussion between authors. Egger's tests were used to access possible publication bias for each risk factor.

2.4.4 Statistical analyses

All statistical analyses were done in STATA Version 15 for Windows (StataCorp, 2017), using *admetan* package (Fisher, 2019). To access heterogeneity across studies, we used **I^2** statistics, which estimates the percentage of variance due to differences between studies. Random effect models were used to produce near to equal weighting between studies. Subgroup analysis was then performed to investigate potential sources of heterogeneity using pre-defined subgroups.

2.5 Results

2.5.1 Study characteristics

I identified 15 studies from 5 countries, which reported data on 246,608 (82% male) individuals from 14 independent samples (Table 2-1). The included studies were published between 1997 and 2018. The majority (11) of the studies were from the USA. Four included papers were reports by governmental agencies (Adams et al., 2011; Department of Justice, 2011; North Carolina Sentencing & Advisory Commission, 2018; Peillard et al., 2012), one was a published thesis (Maliek, 2017), and the rest were published in peer-reviewed journals.

The participants were either representative samples or full cohorts of individuals who received community sentences in a given country or province during a selection period. All included studies utilised a cohort (either prospective or retrospective) design. Re-arrest was the most frequently reported outcome (8 studies) and the mean reported follow-up period was 3.5 years (3 studies did not provide information on the mean follow-up length). The identified risk factor domains that were examined in three or more studies were gender, income, ethnicity, criminal history, marital status, substance misuse problems, mental health needs, educational problems, employment problems, and association with antisocial peers (Table 2-1). The definitions of exact outcomes included in these domains are reported in Appendix B2.

Table 2-1. Description of studies included in the meta-analysis

Study	Country	Cohort (selection years, n, % male, mean age)	Outcomes (type and length of follow-up)	Extracted risk factors	NOQAS score
(Adams et al., 2011)	USA	Individuals sentenced to probation sampled from several counties of Illinois. One individual can be sentenced multiple times (2006, 2770, 82%, 31)	Rearrest (during supervision, mean 19.4 months) Revocation (during supervision, mean 19.4 months)	Gender, age, ethnicity, criminal history, marital status, substance abuse, mental health, education, employment, income, negative peers association	7
(Caudy, Tillyer, & Tillyer, 2018)	USA	Individuals sentenced to probation from one urban county of an unnamed southwestern state (2011-2013, 10642, 76%, 34)	Rearrest (fixed end date, mean 1 year)	Gender	7
(Department of Justice, 2011)	UK – N. Ireland	National cohort of individuals receiving non-custodial sentences (2005, 19047, 85%, ≈33)	Reconviction (starts with a sentence, 2 years)	Gender, age	7
(Grann et al., 2008)	Sweden	National cohort of individuals receiving non-custodial sentences (1993-2001, 4828, 91%, 36)	Reconviction for a violent crime (starts with a sentence + fixed end date, average 4.8 years)	Substance abuse, mental health	8
(Harris, 2011)	USA	A cohort of felony probationers from an unnamed south central state (1993, 3598, 78%, 29)	Rearrest, excluding arrests for technical violations (starts with a sentence, 3 years)	Gender, age, substance abuse, mental health, education, employment, negative peer association	7
(Huebner & Cobbina, 2007)	USA	Individuals sentenced to probation sampled from several counties of Illinois (2000, 3017, 80%, 31) *same dataset as in Olson, 2003	Rearrest (starts with an end of a sentence, 4 years)	Gender, substance abuse, employment	7
(Humphrey, Burford, & Dye, 2012)	USA	Individuals sampled from cohort of standard and reparative probationers from Vermont (1998-2000, 4792, 73%, 28)	Reconviction (starts with a sentence, 5 years)	Gender, criminal history	7

Study	Country	Cohort (selection years, n, % male, mean age)	Outcomes (type and length of follow-up)	Extracted risk factors	NOQAS score
(Maliek, 2017)	USA	Individuals sentenced to probation from Texas (2014-2017, 10127, 73%, 35)	Probation revocation (during supervision, unspecified)	Gender, ethnicity, substance abuse, mental health, education, employment, income, marital status, negative peer association	7
(Minor, Wells, & Sims, 2003)	USA	Individuals sentenced to probation from Kentucky (1996-1999, 200, 68%, median 40)	Probation violation (starts with a sentence, 2 years)	Gender, ethnicity, criminal history, substance abuse, mental health, education, employment	7
(North Carolina Sentencing & Advisory Commission, 2018)	USA	Individuals sentenced to probation from N. Carolina (2015, 32537, 72%, 32)	Rearrest (starts with a sentence, 2 years)	Gender, ethnicity, criminal history, substance abuse, mental health, education, employment, marital status	7
(Olson & Lurigio, 2000)	USA	Individuals sentenced to probation sampled from several counties of Illinois (1997, 2438, 80%, ≈30)	Rearrest (during supervision, unspecified)	Gender, ethnicity, criminal history, substance abuse, education, income	7
(Olson, Alderden, & Lurigio, 2003)	USA	Individuals sentenced to probation from Illinois (2000, 3325, 79%, 31) *same dataset as in Huebner, 2007	Rearrest (during supervision, unspecified)	Gender, ethnicity, age, criminal history, substance abuse, education, income,	7
(Peillard et al., 2012)	Chile	National cohort on individuals receiving non-custodial sentences (2007, 23736, 86%, ≈33)	Rearrest (starts with a sentence, 3 years)	Gender	7
(Sims & Jones, 1997)	USA	Individuals sentenced to probation from North Carolina (1993, 2850, 83%, 27)	Probation failure/revocation (during supervision, mean 30 months) *cohort is selected based on release date	Gender, ethnicity, age, criminal history, substance abuse, education, employment, income, marital status, negative peer association	7
(Wood et al., 2015)	UK – England & Wales	National cohort of individuals sampled from different probation trusts, excluding Tier 1 (low risk) probationers (2009-2010, 125718, 84%, ≈32).	Proven reoffending (starts with a sentence, 1 year for an offence to happen + 6 months for conviction)	Gender, criminal history, substance abuse, mental health, employment, income, negative peer association	8

Table 2-2. Summary of the meta-analysis results

Risk factor domain	Number of studies (k)	Number of individuals (n)	Pooled OR	95% CI	I^2
Non-modifiable					
Gender (male)	13	241,481	1.4	1.2 – 1.6	94%
Age (<21)	5	160,728	1.9	1.6 – 2.3	96%
Ethnicity (non-white)	7	53,248	1.7	1.3 – 2.3	97%
Educational problems (not graduating high school or having identified ed. needs)	9	58,342	1.6	1.3 – 1.9	94%
Criminal history (prior arrest or convictions)	9	185,491	3.0	1.9 – 4.5	99%
Modifiable					
Low income (as specified in jurisdiction)	4	10,302	2.0	1.1 – 3.4	97%
Marital status (single or divorced)	4	40,483	1.6	1.4 – 1.8	42%
Employment problems (unemployed)	8	56,604	1.8	1.3 – 2.5	98%
Substance misuse					
- unspecified	3	47,492	2.3	1.1 – 4.9	98%
- drug misuse	5	13.408	1.7	1.2 – 2.6	97%
- alcohol misuse	3	7,953	1.1	1.0 – 1.2	19%
Association with antisocial peers	6	24,175	2.2	1.3 – 3.7	97%
Mental health needs (diagnosed disorder or symptoms that limit functioning)	4	20,049	1.4	1.2 – 1.6	46%

2.5.2 Quality assessment

Among identified studies, 2 received a score of 8 on the Newcastle-Ottawa Scale, 13 received a score of 7. The commonest identified limitation was missing data for a number of individuals in a cohort or failure to report a non-response rate.

2.5.3 Recidivism risk and static risk factors

The most commonly reported non-modifiable risk factor domains were gender, age, marital status, ethnicity, criminal history and educational problems (Table 2-2, Appendix B3).

In the criminal history domain, I included individuals with arrests or convictions that pre-dated their index crime. Having a prior criminal history was significantly associated with recidivism (k = 9, n = 185,505, pooled OR = 3.1 (CI 95% [2.0, 4.7]; I^2 = 99%). No predefined subgroups explained the observed heterogeneity. To determine the association between age and recidivism, I compared adult individuals younger than 21 years old at their index conviction with older offenders, as this was the most commonly reported age grouping across the studies. Younger age was associated with recidivism (k = 5, n = 160,728, pooled OR = 1.9 (CI 95% [1.6, 2.3]; I^2 = 96%). Heterogeneity somewhat decreases in studies when individuals were followed during their supervision. No other subgroups were relevant for heterogeneity.

Several other non-modifiable risk factor domains had significant associations with recidivism. Those included educational problems, i.e. not having high school diploma or having high educational needs indicated by standardised assessment tools (k = 9, n = 58,342, pooled OR = 1.6 (CI 95% [1.3, 1.9]; I^2 = 94%), being single or divorced (marital status domain; k = 4, n = 40,483, pooled OR = 1.6 (CI 95% [1.4, 1.8]); I^2 = 42%), and being male (gender domain; k = 13, n = 241,481, pooled OR = 1.4 (CI 95% [1.2, 1.6]; I^2 = 94%). In addition, having non-white ethnicity was associated with recidivism (ethnicity domain; k = 7, n = 53,248, pooled OR = 1.7 (CI 95% [1.3, 2.3]; I^2 = 97%). No subgroups explained heterogeneity in these risk factor domains. Data for ethnicity domain were reported only by the studies conducted on US samples.

No significant publication bias was identified for any non-modifiable risk factor.

2.5.4 Recidivism risk and dynamic risk factors

The most commonly reported dynamic risk factor domains were substance misuse (Figure 2-2), mental health needs (Figure 2-3), association with antisocial peers (Figure 2-4), employment problems (Figure 2-5), low income (Figure 2-6), and marital status (Figure 2-7).

Substance misuse as a risk factor was reported differently. A standardised diagnosis was only used in one study (Grann et al., 2008). Instead, problems with alcohol or drugs were typically reported based on interviews and assessments conducted by a probation officer or on record analysis. Recidivism was significantly associated with unspecified substance misuse (k = 3, n = 47,492, pooled OR = 2.3 (CI 95% [1.1, 4.9]); I^2 = 98%) and drug misuse (k = 5, n = 13,408, pooled OR = 1.7 (CI 95% [1.2, 2.6]); I^2 = 96%). There was a weak association with alcohol misuse (k = 3, n = 7,953, pooled OR = 1.1 (CI 95% [1.0, 1.2]); I^2 = 19%). The studies that reported a referral to substance misuse treatment programmes as a measure of this risk factor were excluded. We considered the referral an unsatisfactory proxy for diagnosis since the referral process is, in many cases, voluntary or may not be a part of a sentence at all, even if an offender has known substance misuse problems Data from (Adams et al., 2011; Minor et al., 2003) two studies were excluded for this reason. Further subgroup analysis did not identify factors associated with heterogeneity for substance misuse.

Mental health needs (excluding substance misuse) were associated with increased risk of recidivism (k = 4, n = 20,049, pooled OR = 1.4 (CI 95% [1.2, 1.6]); I2 = 46%). As was the case with substance misuse, medically diagnosed disorders were almost never used as a predictor with the exception of one study (Grann et al., 2008). This domain also included presenting with symptoms that limit functioning or having unspecified mental health needs, an assessment of which was often conducted by a probation officer or was not described. Applying a similar rationale to our approach to substance misuse reporting, we excluded data from one study (Adams et al., 2011) that used a mental health treatment referral as a measure of this risk factor. No pre-identified subgroups explained heterogeneity for mental health needs.

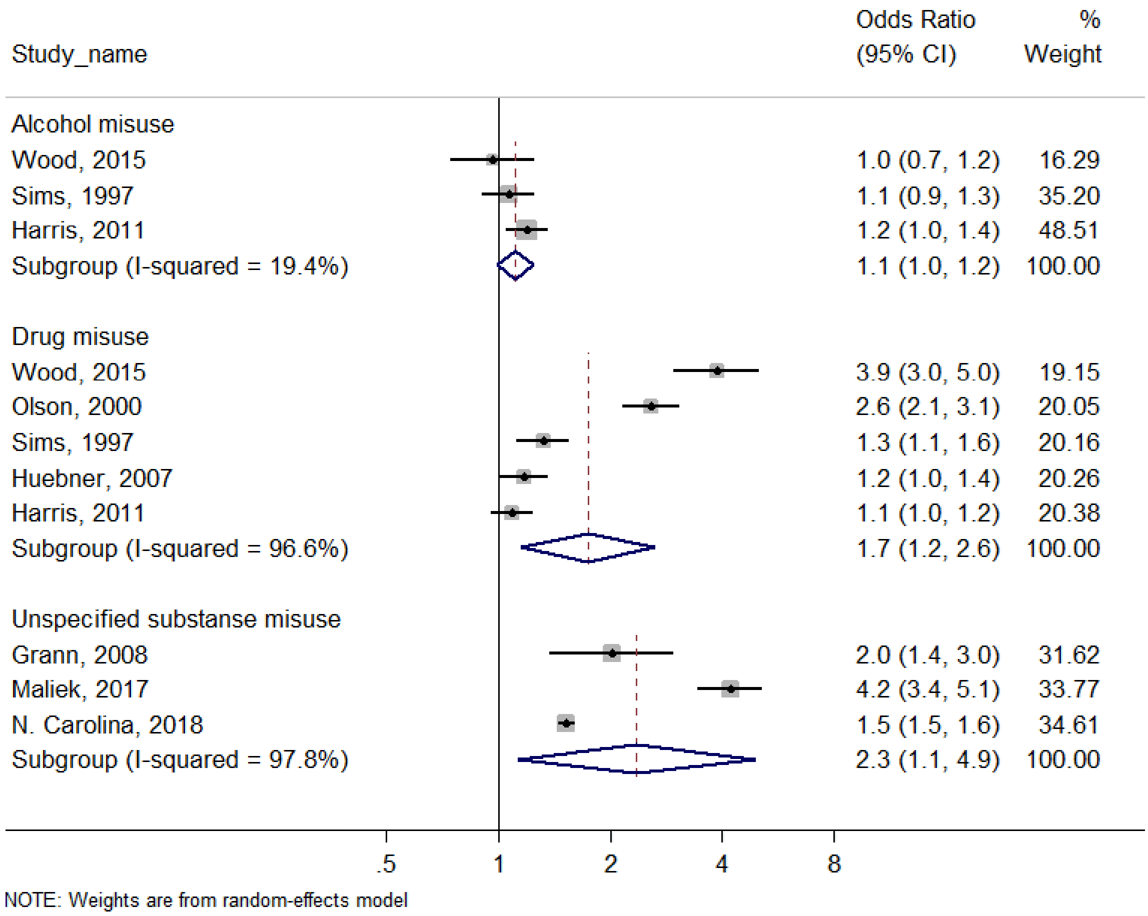


Figure 2-2. ORs for the association between substance misuse and the risk of recidivism in community-sentenced populations by type of misuse

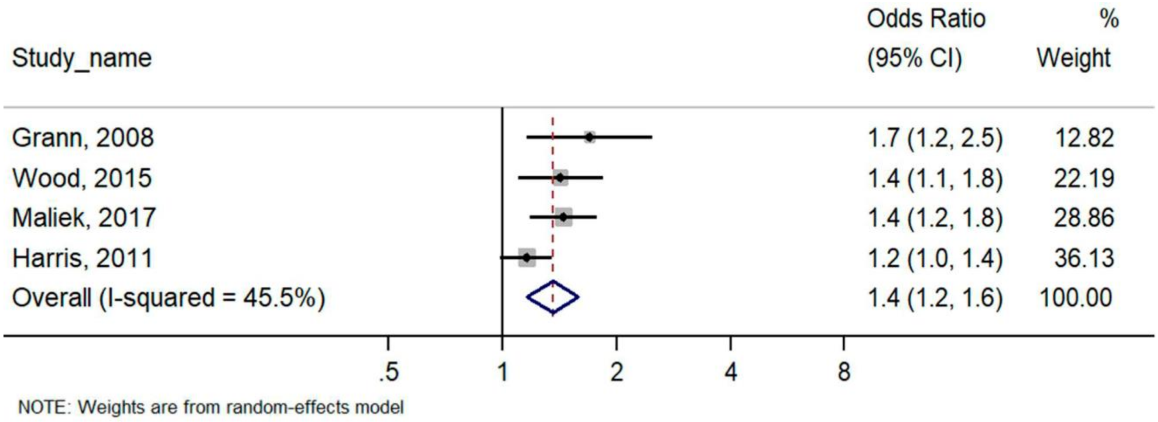


Figure 2-3. Odds ratios (ORs) for the association between mental health needs and the risk of recidivism in community-sentenced populations

Having antisocial peers was also associated with recidivism (k = 6, n = 24,175, pooled OR = 2.2 (CI 95% [1.3, 3.7]); I2 = 97%). This domain included individuals with known gang affiliations, antisocial friends or lack of prosocial friends. The assessment of this factor was performed mostly by a probation officer or through an analysis of records. Heterogeneity was partially explained by lower risk estimates in those below 30 years old compared to 30-35 years old. No other subgroups were associated with lower heterogeneity.

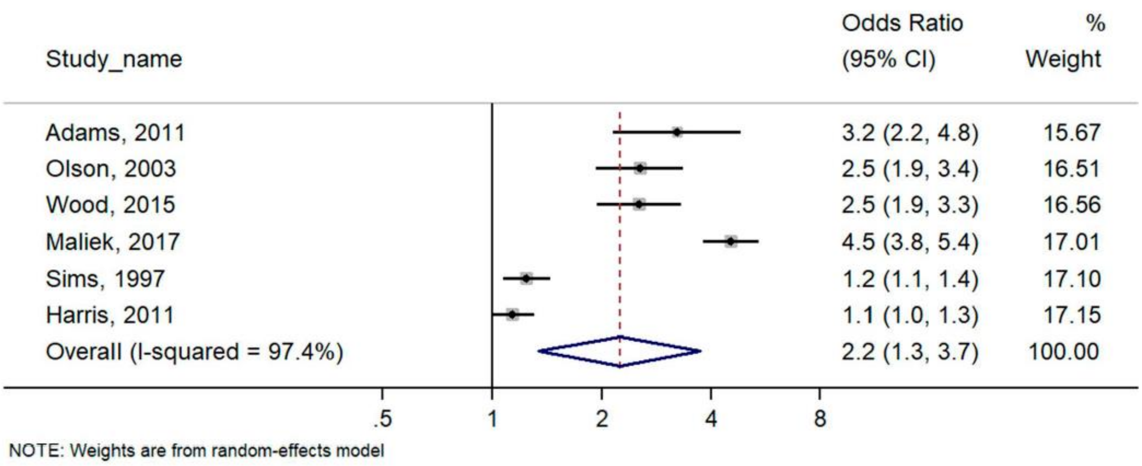


Figure 2-4. Odds ratios (ORs) for the association between association with antisocial peers and the risk of recidivism in community sentenced populations

Being unemployed at the time of the conviction was associated with the increased risk of recidivism (k = 8, n = 56,604, pooled OR = 1.8 (CI 95% [1.3, 2.5]); I2 = 98%) as well as having low income (k = 4, n = 10,302, pooled OR = 2.0 (CI 95% [1.1, 3.4); I2 = 97%] and being single or divorced (marital status domain; k = 4, n = 40,483, pooled OR = 1.6 (CI 95% [1.4, 1.8]); I2 = 42%). No predefined subgroups explained heterogeneity for low income, unemployment or marital status.

No significant publication bias was identified for any dynamic risk.

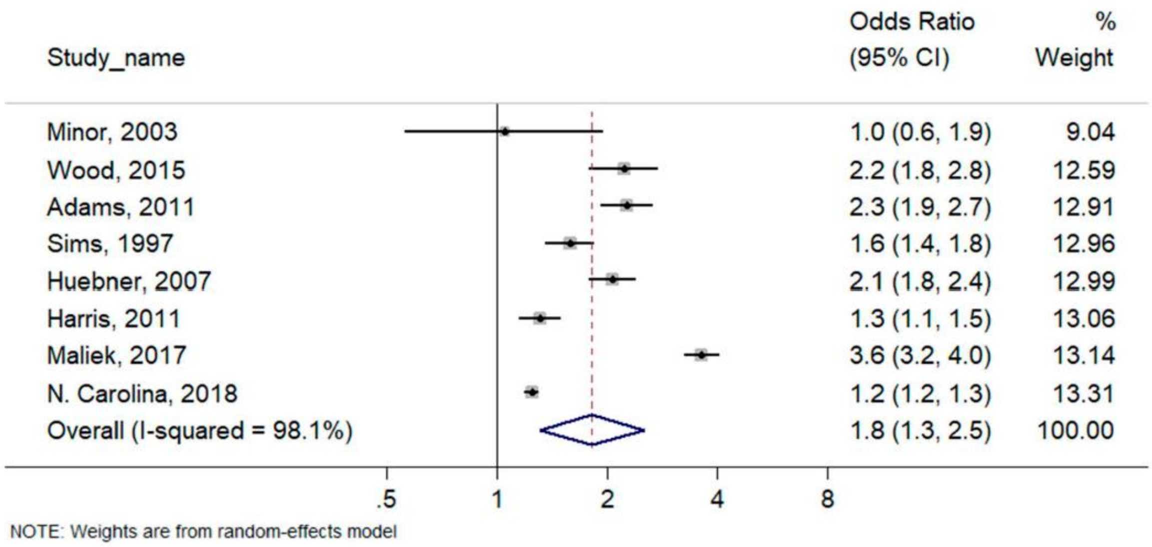


Figure 2-5. Odds ratios (ORs) for the association between employment problems and the risk of recidivism in community sentenced populations

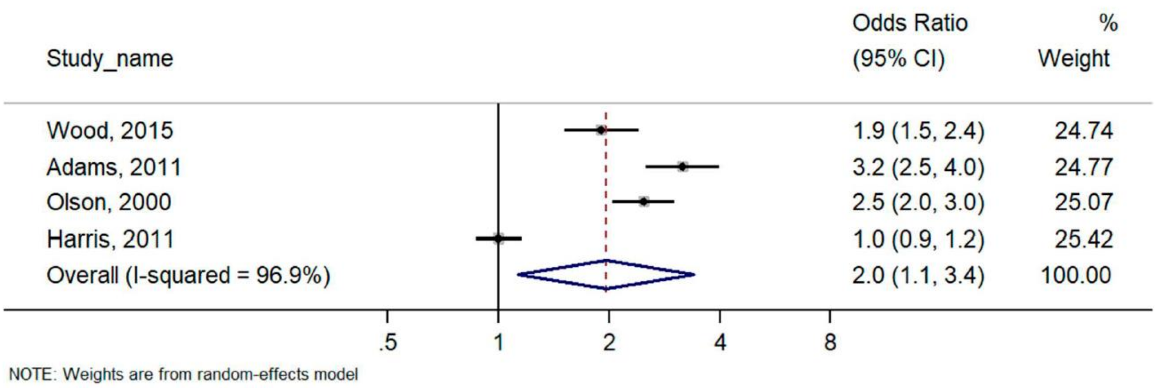


Figure 2-6. Odds ratios (ORs) for the association between having low income and the risk of recidivism in community sentenced populations

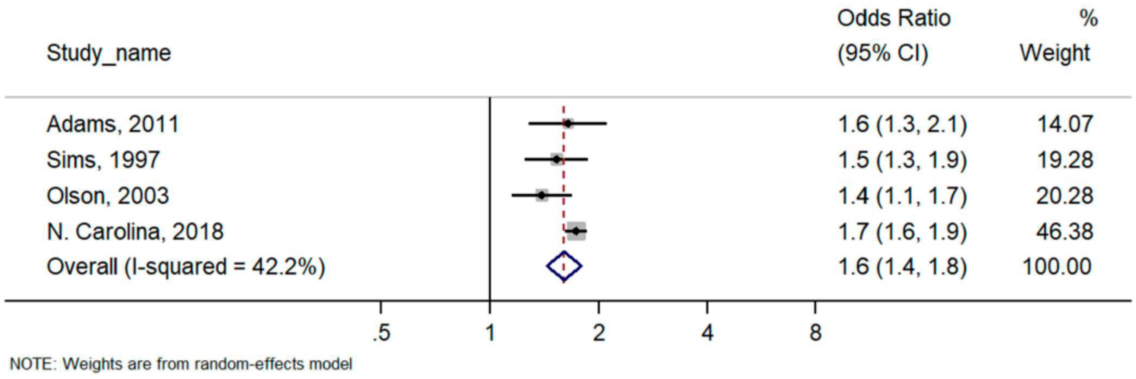


Figure 2-7. Odds ratios (ORs) for the association between marital status (being single or divorced) and the risk of recidivism in community sentenced populations

2.6 Discussion

This meta-analysis examined the most reported risk factors for recidivism in community-sentenced populations and identified 15 studies involving 246,608 individuals. Three main findings emerge. The first is that dynamic risk factors such as mental health needs, substance misuse, association with antisocial peers and employment problems increased the risk of recidivism in community sentenced populations. The second is that the strength of these associations was comparable with static risk factors, such as age, gender and criminal history. The third is a relative dearth of published studies on dynamic risk factors that specifically examine individuals receiving community sentences.

Among static risk factors, younger age and prior criminal history had the strongest association with recidivism. The strength of this association may be considered moderate. Those factors, along with sex, are the common predictors of recidivism in different populations (Fazel et al., 2016). The frequency of criminal behaviour peaks in adolescence and early adulthood, and having a prior criminal history may reflect a life-span persistent criminal career, which also often begins during adolescence. Educational problems such as not completing high school education may also reflect the early adolescent onset of criminal behaviour and be related to persistent problems with social adjustment, which could make the successful reintegration of an offender challenging.

I identified several commonly reported dynamic risk factors that were associated with recidivism in community sentenced populations, including substance misuse, association with antisocial peers, mental health needs, low income and problems with employment. The association between substance misuse and recidivism is a common finding in studies on violence and recidivism among released prisoners (Z. Chang, Larsson, et al., 2015; Fazel et al., 2016; Stahler et al., 2013). The association may reflect core endophenotypes for substance misuse, such as poor inhibitory control (Ersche et al., 2012). Drug or alcohol intake may have a disinhibiting effect on an individual, thus increasing the risk of committing an impulsive crime. Chronic consumption may lead to long-term neurological deficits that are also associated with decreased self-control and increased risk for violence (Arseneault et al., 2000; Sinha, 2008). Moreover, drugs may serve as a direct motive for a crime, and illegal possession of drugs may itself be considered a crime. The risk may also depend on the type of drugs used (Hendricks et al., 2017). When alcohol misuse was examined separately, the association was not as strong. Studies in released prisoners have previously shown that diagnosis of alcohol use disorder increased the risk (to the same level as drug use disorder) of reoffending (Z. Chang, Larsson, et al., 2015). Most likely, there were not enough identified studies that examined alcohol abuse as a stand-alone risk factor in the meta-analysis. Another possible reason for this finding was how alcohol or drug misuse were measured, typically based on self-report or poorly defined criteria.

Mental health needs were significantly associated with recidivism, which is an important finding. Prior meta-analyses have also found that mental health disorders in general and forensic populations increased the risk of violence (Fazel et al., 2009; Oram et al., 2014). However, very few identified studies have investigated the mental health of general community-sentenced populations. Also, using this broad category of mental health needs as a risk factor may not be that practically meaningful for the prediction of repeated criminal behaviour since different types of disorders may have different associations with recidivism (Grann et al., 2008). The existing standardised tools used by probation officers have a ‘mental health needs/problems’ code, but often do not account for any specific diagnosis that an individual may have. Although not always possible, using more precise diagnostic categories may be more helpful since this requires professional assessment or access to medical records. Overall, further analysis is required to assess the usefulness of psychiatric diagnoses and their comorbidity in community sentenced populations.

To determine precise mechanisms of recidivism, it may be informative to examine potential interactions between dynamic and static risk factors, as some factors become more informative for specific subgroups of offenders. For example, Harris (2011) compared risk factors among offenders with different criminal career trajectories. Familial problems predicted future re-arrest among first-time adult offenders, but they were not a predictor for sentenced offenders with known criminal histories. Many of the factors that may have associations with repeat offending, such as childhood adverse experiences and history of victimisation (Fitton et al., 2020), were examined in the context of general violent behaviour, but not in the context of recidivism studies.

2.7 Strengths and limitations

This is the first meta-analysis, to my knowledge, to investigate risk factors for criminal recidivism in the general population of individuals receiving community sentences. Studies included in the final analysis were of high quality (as assessed by Newcastle-Ottawa Scale) and were conducted using large samples.

The small number of published studies limits the generalisability of the results and leads to several additional limitations. First, it was often not possible to reliably estimate potential sources of heterogeneity, which was high for almost every included risk factor domain. Given the variety of ways in which one risk factor may be defined and measured across different studies, the conclusions should be viewed with caution. Second, it was not possible to separately analyse distal (i.e., prior history of substance abuse/mental health problems) and proximal risk factors (ongoing problems with substance misuse/mental health at the time of the conviction). For this reason, they were combined under their respective domains. Third, I was not able to compare the effects across different outcomes (re-arrest, reconviction, technical violation with/without termination of a sentence, reimprisonment) and different follow-up models (recidivism while serving a sentence vs. recidivism after the completion of a sentence). Also, I did not have enough data to compare violent and non-violent recidivism outcomes.

Another variable that might have contributed to heterogeneity was the difference in sentencing practices among jurisdictions, which this study does not account for. In particular, in jurisdictions where prison sentences are more common, community sentenced cohorts may be comprised of lower-risk individuals when compared to jurisdictions where prison sentences are less common. Differences in sentencing practices result in cohorts with varying compositions that render direct comparisons problematic.

Although there was no identified publication bias, there is still a possibility that some studies may not have provided data for risk factors in cases of non-significant findings. Finally, in terms of geographical generalisability, the included studies were limited to Europe and the USA. The USA studies were overrepresented in the meta-analysis. All studies that examined ethnicity were from the USA, and this risk factor is not generalisable to other countries.

2.8 Conclusion

Modifiable risk factors such as mental health needs, substance misuse, association with antisocial peers, low income, employment problems and marital status were associated with the risk of recidivism in individuals receiving community sentences. Further integration of mental health services within criminal justice community supervision agencies requires careful thought. It should be based upon the understanding of the treatment needs and recidivism mechanisms of these specific populations. In addition, overreliance on static (non-modifiable) risk factors and underplaying dynamic (modifiable) mental health needs during risk assessment should be avoided. It may lead to less effective rehabilitation practices considering the high prevalence of mental health problems in general community-sentenced populations.

When reporting data for mental health risk factors, diagnostic categories should be provided when the medical records are available, and comorbidity with substance misuse should be documented. When reporting mental health and substance misuse problems as risk factors, the differences between ongoing problems at the time of a conviction (proximal factors) and problems in the past (distal factors) should be clearly indicated. In addition, researchers and agencies should explore other types of predictors identified in the literature, such as a history of maltreatment and victimisation, since chronic or ongoing psychological trauma may be an important therapeutic target during rehabilitation. Some of these factors have been extensively studied in other contexts (as predictors of violent behaviour and well-being), but not within the context of recidivism. Exploring the association of particular symptoms of mental disorders with a plausible connection to recidivism may also be useful. Finally, to make comparisons between studies more meaningful, recidivism data should be reported across different outcomes, including violent and non-violent recidivism. The use of common reporting guidelines (Fazel, Wolf, & Yukhnenko, 2019) may facilitate this process.

Chapter 3. Psychiatric disorders and reoffending: a national cohort study of individuals given community sentences in Sweden

3.1 Abstract

I examined the association between psychiatric disorder and criminal recidivism in a large nationwide cohort of individuals given community sentences in Sweden, employing a sibling design to account for potential unmeasured familial confounding. Criminal recidivism was operationalised as 2-year general and violent reoffending.

The study cohort consisted of 82,415 individuals given community sentences between 1991 and 2013 in Sweden. Psychiatric disorders were associated with an increased risk of criminal recidivism in male and female individuals given community sentences. In men, this association remained substantial after adjustment for unmeasured familial confounding. The association between psychiatric disorders and general reoffending was fully explained by comorbid substance use disorder. The association between psychiatric disorders and violent reoffending was only partially explained by comorbid substance use disorder. Schizophrenia spectrum disorders, personality disorders and substance use disorders had stronger effects on violent reoffending than other psychiatric disorders.

Psychiatric disorders significantly were associated with an increased risk of general and violent reoffending in the community sentenced population. Policymakers and probation agencies should consider expanding the use of mandated mental health and substance use treatment programmes to reduce reoffending in this important population.

3.2 Introduction

Chapter 2, based on several identified studies, reported that dynamic risk factors such as mental health needs and substance misuse increased the risk of recidivism in community sentenced populations. This chapter aimed to test this conclusion and add to existing evidence on the association between psychiatric disorders and criminal recidivism. The chapter also provided additional justification regarding the relevance of mental health research in individuals given community sentences.

As discussed in the general introduction to the thesis, community sentences provide a key alternative to custodial sentences and are more frequently imposed (Aebi et al., 2019). In England and Wales in 2020, only 7% of all sentences were immediate custodial sentences (Ministry of Justice, 2021a). The remaining 93% consisted of community sentences, suspended custodial sentences, and, most frequently, fines. This community sentenced population represents the bulk of the England and Wales probation service's supervisory workload: 93,600 individuals compared to the 63,800 individuals on post-custodial supervision (Ministry of Justice, 2021b). In the same year in Sweden, the caseload distribution was similar, with 5,983 community sentenced individuals and 2,645 individuals under post-custodial supervision (Kriminalvården, 2020). The proportion of individuals serving a community sentence in the US is even higher. In 2019, there were 3,492,900 community sentenced individuals under probation supervision compared to 878,900 individuals under post-custodial supervision.

Research examining the effect of psychiatric disorders on reoffending in community sentenced populations is limited. The systematic review presented in the previous chapter identified very few studies that specifically examined the role of psychiatric disorders in community sentenced individuals. The identified studies reported a small to moderate increase in the risk of reoffending in individuals with mental health and substance use disorders. These studies, however, had major limitations. Most of them lacked diagnostic specificity and did not provide information about individual psychiatric diagnoses. Instead, they relied on general categories of 'mental impairment' (Harris, 2011) or 'identified mental health need' (Wood et al., 2015). Some relied on self-report measures and not on medical records (Wood et al., 2015). The single identified study that provided information on specific diagnostic categories (Grann et al., 2008) was underpowered, did not examine non-violent outcomes, and did not consider the potential impact of familial confounding. The current study aimed to address these limitations.

In this population-based retrospective cohort study of individuals given community sentences, I examined the association of psychiatric disorders with general and violent reoffending. I used official medical records obtained from population registers to establish a history of psychiatric diagnoses. In addition, I utilised a sibling design to control for potential familial confounding. We addressed three questions. First, whether being diagnosed with a psychiatric disorder is independently associated with general and violent reoffending in community sentenced individuals. Second, whether this association differs by psychiatric diagnosis. Third, the extent to which any observed associations are explained by comorbid substance use disorder.

3.3 Approach

Cox proportional hazard regression was the main statistical method used in the current study. It is an epidemiological model that have been widely used in medical and criminological research (Flores, Holsinger, Lowenkamp, & Cohen, 2016; Grann et al., 2008; Olsson, Öjehagen, Brådvik, & Håkansson, 2015). Cox regression ‘produces a survival function that predicts the probability that the event of interest has occurred at a given time for given values of the predictor variables’ (IBM, 2021). The main outputs of the Cox regression model are the baseline hazard and covariate coefficients. Covariate coefficients are the weights assigned to predictors (covariates) entered in the model. The baseline hazard represents the probability of an outcome prior to a given time if all covariate values in the model are set to zero. Covariate coefficients modify the baseline hazard depending on their sign, increasing or decreasing the probability of an outcome occurrence at a given time. The effects of the covariates on the baseline hazard are assumed to be constant over time (the proportionality assumption). Cox proportional hazard regression also accounts for censoring, which is the loss of individuals during the follow-up period for reasons other than the outcome of interest.

To test the potential effect of unmeasured familial confounding, I also employed a fixed-effect sibling comparison based on Cox regression (Petersen & Lange, 2020). In this analysis, siblings from one family were assumed to have the same baseline hazard. This baseline hazard absorbs the effect of all unmeasured factors shared between siblings. Thus, the effect of measured covariates is estimated relative to the familial baseline hazard. The covariates’ coefficients are computed as a weighted average of coefficients across families.

To test the effect of the first-time diagnosis during the follow-up, I additionally fitted a Cox regression model with time-varying covariates (Zhang et al., 2018).

All analyses were done on data extracted from Swedish population registers. The accuracy of the Swedish national registers was estimated to be 85-95% (depending on diagnosis) for medical data (Ludvigsson et al., 2011) and 85% for sociodemographic information (Ludvigsson et al., 2019).

3.4 Methods

I followed STROBE guidelines (von Elm et al., 2007) for the reporting of observational studies (see Appendix C1 for the checklist).

3.4.1 Study setting

I linked several longitudinal, nationwide Swedish registers: the National Crime Register, containing information about criminal offences and convictions since 1973; the National Patient Register, providing information about diagnoses for individuals admitted to inpatient hospitals (since 1973) and outpatient care (since 2001); the Migration Register, containing dates of migration to and from Sweden; the Cause of Death Register, containing information about deaths and causes of deaths since 1958; the Multi-Generation Register, containing information about biological relationships for all individuals living in Sweden since 1933; the Longitudinal Integration Database for Health Insurance and Labour Market Studies, containing yearly estimations of income, marital and employment status, and education since 1990. In Sweden, all residents (including immigrants) have a unique personal identifier used in all national registers, thus enabling data linkage (Ludvigsson et al., 2009).

This study was approved by the Regional Ethics Committee at the Karolinska Institutet (Stockholm, Sweden).

3.4.2 Participants

I included all adult (18 years or older) Swedish residents who received any community sentence at any point from Nov 1, 1991, to Dec 31, 2013. We chose the starting point to ensure the full availability of sociodemographic information for all selected individuals as this information was only available in registers from 1991. Community sentences included conditional sentences with community service, probation with community service, and probation with contracted treatment. I only selected individuals whose sentences came into legal force and were not appealed or dismissed. For each individual, I used the date when a community sentence came into force as the start of the follow-up period. If an individual received multiple community sentences, the index sentence was selected at random.

I excluded individuals who were born before 1958 because these individuals would not have a continuous criminal record on the National Crime Register. I also excluded individuals who committed a crime before the start of the follow-up period but were not sentenced for it by that time – this is referred to as a pseudo-reconviction. The inclusion of such pseudo-reconvictions could have resulted in an overestimation of recidivism risk. Therefore, if an individual had a prior recorded crime without a corresponding sentence at their selected starting time, the starting point was re-selected to reflect this. If the individual again had a pseudo-reconviction outcome on the second attempt, they were then excluded from the analysis. We additionally identified same-sex full siblings within the cohort using the Multi-Generation Register. The flowchart for the selection process is included in Appendix C2.

3.4.3 Measures

I extracted sociodemographic, criminal and medical history information at the start of the community sentence. The sociodemographic information included sex, age, years of education, marital status, employment information, and receipt of income support.

Criminal history included dates of all prior convictions that were legally enforced and their corresponding crime codes. I separately recorded if the index sentence was an individual's first entry in the Crime Register or if they had been previously sentenced. I also recorded if an individual was sentenced for a violent offence in the past and whether their index offence was a violent offence. The violent offence was defined as homicide, assault, robbery, arson, any sexual offence, illegal threats, or intimidation.

Medical history included a history of any psychiatric diagnosis received before the index sentence. In line with previous research using Swedish national registers (Z. Chang, Larsson, et al., 2015), I utilised a hierarchical approach to coding main diagnostic categories. The hierarchy was as follows: schizophrenia spectrum disorders, bipolar disorder, depression, and anxiety disorder. If an individual had a diagnosis of schizophrenia and any other diagnoses, we classified that individual as having schizophrenia. If an individual did not have schizophrenia but had bipolar disorder and depression or anxiety, we classified that individual as having bipolar disorder, and so on.

To explore the effects of comorbidity between psychiatric disorders, I also investigated alcohol use disorder, drug use disorder, personality disorder, attention-deficit hyperactivity disorder, and other developmental or childhood disorders. We did not use a hierarchical approach for these comorbidities but examined whether they were present or not. ICD codes for the psychiatric diagnoses are listed in Appendix C3. I additionally coded the substance use disorder category as having either alcohol use disorder, drug use disorder, or both. I recorded the number of distinct diagnostic categories that an individual belonged to assess the cumulative effect of multiple diagnoses.

Medical data from outpatient visits were only available from 2001 in the National Patient register. To check the possible effect of the information source on the estimated association between psychiatric diagnoses and the outcomes, I conducted a sensitivity analysis by separately analysing individuals sentenced before 2001 and after 2001.

3.4.4 Missing data

0.7% of individuals within the cohort did not have demographic information, and 4.0% did not have education data at baseline. We did not replace missing data by imputation or other methods because the number of individuals with missing values was quite small. However, in a sensitivity analysis, we recalculated the results with missing values imputed.

3.4.5 Outcomes and censoring

General reoffending

I defined general reoffending as committing any offence after the start of the index sentence until Dec 31, 2013. The dates of crimes are recorded in registers retroactively after the circumstances of the crime have been established by a court. If no date for the offence had been recorded, I used the date of the court sentence. This approach allows for less conservative estimates of risk. All individuals were followed up until their first new offence, death, permanent emigration from Sweden, or the end of the available register records (Dec 31, 2013).

Violent reoffending

I defined violent reoffending as committing any violent crime after the start of the index sentence but before Dec 31, 2013. Violent crime was defined as homicide, assault, robbery, arson, any sexual offence (rape, sexual coercion, child molestation, indecent exposure, or sexual harassment), illegal threats, or intimidation, which is the same operational definition used in previous studies (Z. Chang, Larsson, et al., 2015; Fazel et al., 2016). All individuals were followed up until their first new violent offence, imprisonment for any crime, death, permanent emigration from Sweden, or the end of the available register records (Dec 31, 2013).

3.4.6 Statistical analysis

To examine the association between psychiatric disorders and risks of general and violent reoffending, I compared community sentenced individuals with and without psychiatric disorders. I used the Cox proportional hazard model as the method of quantifying this association and produced Kaplan-Meier survival curves. I tested proportional hazards assumptions by visually examining the Kaplan-Meier curves and Schoenfeld residuals diagrams. All analyses were stratified by sex. If the date of a new offence was the same as the date of the start of the sentence, we changed the end time of 0 to the end time of 0.5 (Therneau & Grambsch, 2000). The analyses were completed in R using the *survival* package (Therneau, 2021). The visualizations were created using the *survminer* package (Kassambara et al., 2017) and Tableau Desktop (Tableau Software, 2021).

Psychiatric disorders

I estimated the association between psychiatric disorders and reoffending by fitting two Cox regression models. In the first model, I adjusted only for age at the time of the sentence. In the second model, I fitted a fixed-effect Cox regression model (Allison, 2009) to the subsamples of same-sex full siblings receiving a community sentence but discordant for a given psychiatric diagnosis. The fixed-effect model adjusts for all unmeasured genetic and environmental factors that are shared between siblings, provided their effect remains constant over time. To test the effect of first-time diagnosis during the follow-up, I additionally fitted a Cox regression model with time-varying covariates in the subsample of individuals without a psychiatric diagnosis at sentence (Zhang et al., 2018).

To further assess the effect of psychiatric disorders on criminal reoffending behaviours, I calculated the population attributable fraction (PAF). The PAF estimates the proportion of new offences that can be attributed to a risk factor, with the assumption that a causal association exists. To calculate PAF, I used the model-based adjusted attributable fraction function for Cox proportional hazard models in the *AF* package for R (Dahlqwist & Sjolander, 2019).

Comorbidity

To further explore the association between individual psychiatric disorders and the risk of reoffending, I compared individuals diagnosed with schizophrenia, bipolar disorder, depression, or anxiety disorder with and without comorbid substance use to individuals without a psychiatric history. In this case, I fitted Cox regression models adjusted by age and stratified by sex. In addition, I estimated the effect of having multiple psychiatric diagnoses on reoffending, stratifying on comorbid substance use diagnosis.

3.5 Results

The study cohort consisted of 82,415 individuals (70,643 men and 11,772 women), who received at least one community sentence in Sweden during the period from Nov 1, 1991, to Dec 31, 2013 (see Appendix C2 for the selection process). These persons were followed for up to 10 years after their index sentence with a mean follow-up period of 2 years and 10 months. In men, 33,652 (47.8%) of 70,593 committed a new offence during follow-up, and 10,577 (15.0%)

committed a new violent offence. In women, 4,392 (37.5%) of 11,793 committed a new offence during follow-up, and 863 (7.3%) committed a new violent offence (see Appendix C4 for overall survival estimates).

At baseline, a higher proportion of women in the cohort (7,094 of 11,793 [60.2%]) had been diagnosed with a psychiatric disorder compared to men (27,110 of 70,593 [38.4%]). The baseline sociodemographic and criminological information, psychiatric diagnoses, and follow-up data are presented in Table 3-1. The univariate association of baseline characteristics with general and violent reoffending is presented in Appendix C5. Many psychiatric diagnoses were associated with others, reflecting psychiatric comorbidity (see Appendix C6 for collinearity tests).

3.5.1 Psychiatric diagnoses and general reoffending

Having any prior psychiatric diagnosis at the start of a community sentence was associated with an increased risk of general reoffending (Figure 3-1, Figure 3-2, Figure 3-3; Table 3-2). All individual psychiatric diagnoses, except for bipolar disorder and depression, were associated with an increased risk of general reoffending in both men and women.

In men, hazard ratios for individual disorders ranged from 1.02 (95% CI 0.91-1.15) for bipolar disorder to 2.27 (95% CI 2.21-2.33) for drug use disorder. Overall, in the male cohort, 24,654 cases of general reoffending occurred during the first two years of the follow-up period, 3,349 of which were potentially attributable to psychiatric disorders. This corresponds to a PAF of 13.6% (95% CI 12.9-14.2%). For PAFs estimated from the models adjusted for measured sociodemographic covariates and criminal history, refer to Appendix C7.

Table 3-1. Baseline characteristics and follow-up data of adult individuals receiving community sentences from November 1, 1991, to December 31, 2013

510 men and 58 women have missing values for marital status, employment, and income support. 2,912 men and 423 women have missing values for education.

	Men	Women	Total
Number of individuals	70,643	11,772	82,415
Baseline			
Median age	27 (IQR: 22-38)	31 (IQR: 23-41)	28 (IQR: 22-38)
Married or in a registered partnership	7,453 (10.6%)	1,642 (13.9%)	9,095 (11.0%)
Employed	31,153 (44.1%)	3,909 (33.2%)	35,062 (42.5%)
Highest level of education			
< 9 yr	2,322 (3.3%)	500 (4.2%)	2,822 (3.4%)
9-11 yr	60,649 (85.9%)	9,705 (82.4%)	70,354 (85.4%)
≥ 12 yr	4,760 (6.7%)	1,144 (9.7%)	5,904 (7.2%)
Recipient of income support	24,367 (34.5%)	5,634 (47.9%)	30,001 (36.4%)
Prior criminal history	54,395 (76.9%)	7,832 (66.5%)	62,227 (75.5%)
Prior violent crime	27,222 (38.5%)	2,373 (20.2%)	29,595 (35.9%)
Prior imprisonment	15,755 (22.3%)	1,392 (11.8%)	17,147 (20.8%)
Index violent offence	32,941 (46.6%)	4,013 (34.1%)	36,954 (44.8%)
Any psychiatric disorder	27,138 (38.4%)	7,062 (60.0%)	34,200 (41.5%)
Any psychiatric disorder (other than substance use)	18,047 (25.5%)	5,486 (46.6%)	23,533 (28.6%)
Schizophrenia spectrum disorder	2,032 (2.9%)	563 (4.8%)	2,595 (3.1%)
Bipolar disorder	690 (1.0%)	340 (2.9%)	1,030 (1.2%)
Depression	5,447 (7.7%)	2,037 (17.3%)	7,484 (9.1%)
Anxiety disorder	5,604 (7.9%)	1,869 (15.9%)	7,473 (9.1%)
Personality disorder	2,671 (3.8%)	1,324 (11.2%)	3,995 (4.8%)
Attention-deficit hyperactivity disorder	3,370 (4.8%)	608 (5.2%)	3,978 (4.8%)
Other developmental or childhood disorder	3,246 (4.6%)	777 (6.6%)	4,023 (4.9%)
Substance (drug or alcohol) use disorder	18,680 (26.4%)	4,825 (41.0%)	23,505 (28.5%)
Alcohol use disorder	11,569 (16.4%)	2,961 (25.2%)	14,530 (17.6%)
Drug use disorder	11,864 (16.8%)	3,345 (28.4%)	15,209 (18.5%)
Follow-up data. General reoffending			
Number of persons-year at risk	201,415.6	38,800.4	240,216
Incidents of reoffending during follow-up	33,774 (47.8%)	4,434 (37.7%)	38,208 (46.4%)
Median time to any new offence (months)			
All individuals	22.4 (IQR: 7.4-53.0)	28.2 (IQR: 9.7-61.6)	23.2 (IQR: 7.6-54.9)

Individuals with psychiatric disorder	15.7 (IQR: 5.1-39.8)	22.8 (IQR: 6.8-52.2)	17.1 (IQR: 5.4-43.2)
Individual without psychiatric disorder	27.4 (IQR: 9.4-60.5)	38.2 (IQR: 13.5-74.0)	28.2 (IQR: 9.7-61.6)
Reoffending rate (cumulative)			
1-year	18,019 (25.5%)	2,427 (20.6%)	20,446 (24.8%)
2-year	24,654 (34.9%)	3,234 (27.5%)	27,888 (33.8%)
3-year	27,992 (39.6%)	3,666 (31.1%)	31,658 (38.4%)
4-year	30,012 (42.5%)	3,927 (33.4%)	33,939 (41.2%)
5-year	31,351 (44.4%)	4,104 (34.9%)	35,455 (43.0%)
Died during follow-up	1,024 (1.4%)	179 (1.5%)	1,203 (1.5%)
Emigrated during follow-up	750 (1.1%)	106 (0.9%)	856 (1.0%)
Follow-up data. Violent reoffending			
Number of persons-year at risk	276,437.0	53,244.5	329,681.5
Incidents of reoffending during follow-up	10,591 (15.0%)	868 (7.4%)	11,459 (13.9%)
Median time to a violent offence (months)			
All individuals	37.7 (IQR: 14.5-72.4)	46.5 (IQR: 20.3-82.9)	38.7 (IQR: 15.1-73.6)
Individuals with psychiatric disorder	29.9 (IQR: 10.8-61.4)	41.9 (IQR: 16.6-75.7)	32.1 (IQR: 11.8-64.1)
Individual without psychiatric disorder	41.9 (IQR: 16.6-75.7)	55.6 (IQR: 25.5-90.3)	44.8 (IQR: 18.8-80.0)
Reoffending rate (cumulative)			
1-year	4,030 (5.7%)	300 (2.5%)	4,330 (5.3%)
2-year	6,219 (8.8%)	467 (4.0%)	6,686 (8.1%)
3-year	7,552 (10.7%)	572 (4.9%)	8,124 (9.9%)
4-year	8,462 (12.0%)	657 (5.6%)	9,119 (11.1%)
5-year	9,077 (12.8%)	711 (6.0%)	9,788 (11.9%)
Imprisoned during follow-up	9,802 (13.9%)	1,082 (9.2%)	10,884 (13.2%)
Died during follow-up	1,655 (2.3%)	324 (2.8%)	1,979 (2.4%)
Emigrated during follow-up	1,027 (1.5%)	152 (1.3%)	1,179 (1.4%)

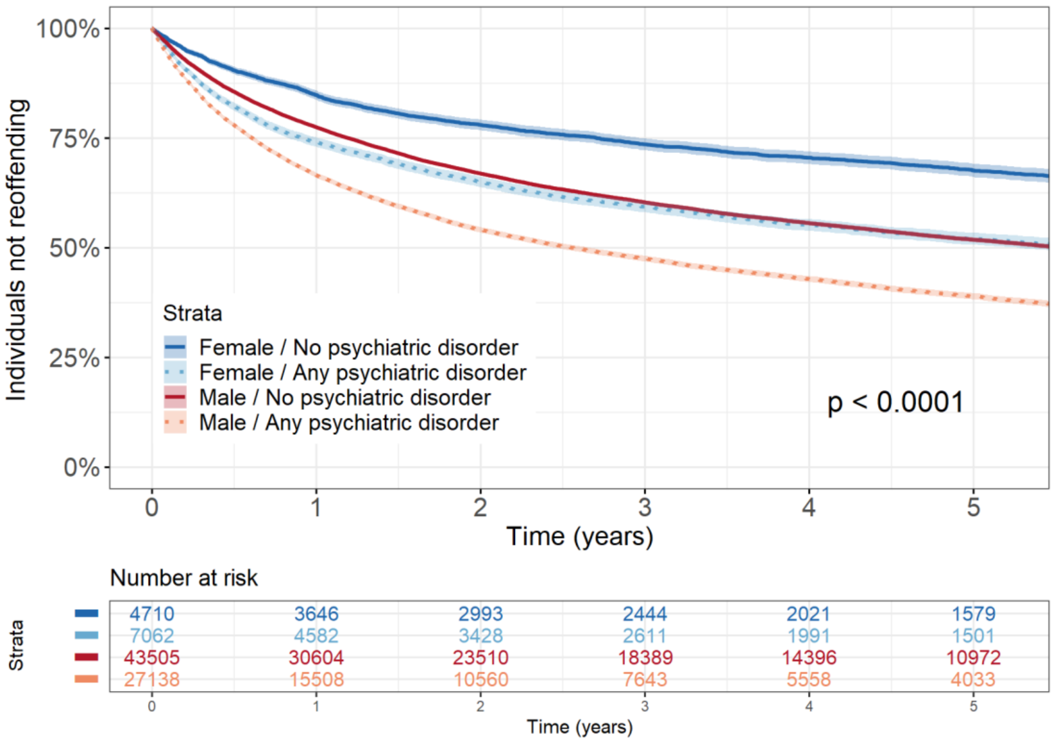


Figure 3-1. Kaplan-Meier curves (unadjusted model) for general reoffending in individuals given community sentences by sex and psychiatric disorder status

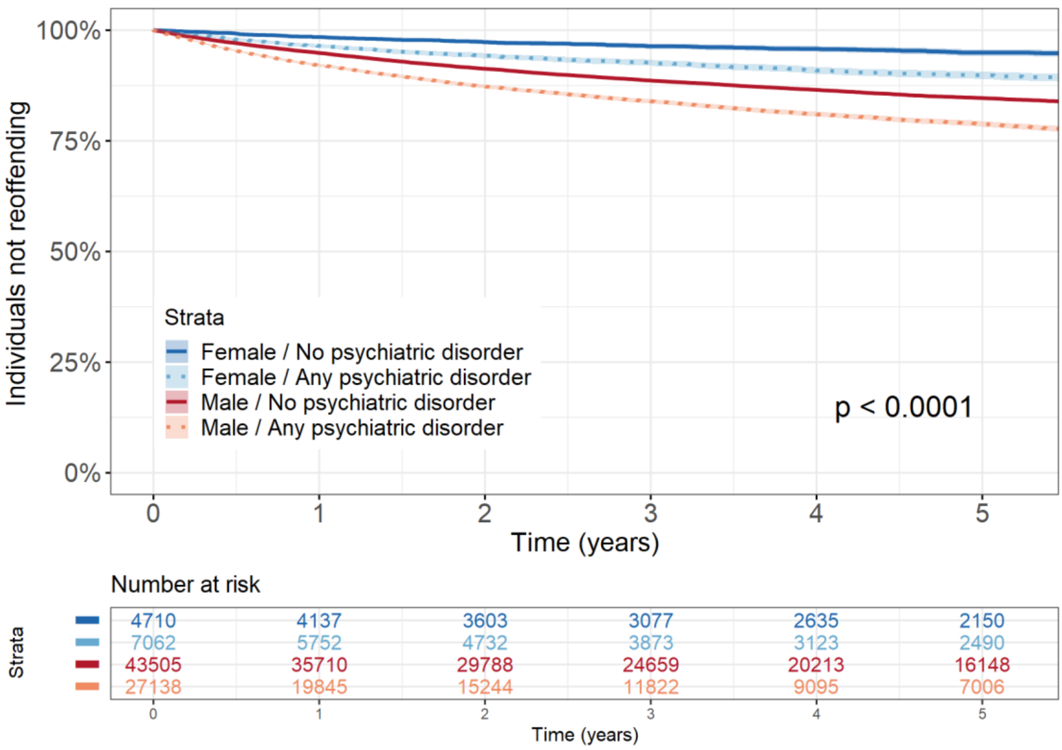
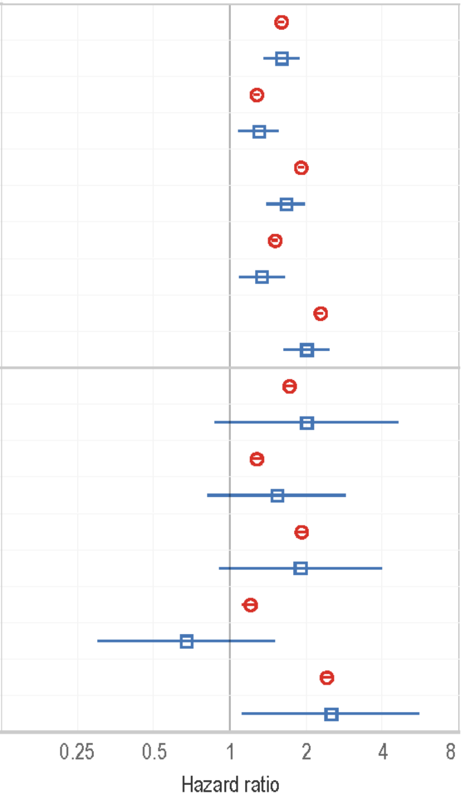


Figure 3-2. Kaplan-Meier curves (unadjusted model) for violent reoffending in individuals given community sentences by sex and psychiatric disorder status

	Diagnosis	N w. diag.	Cases/siblings	HR (95% CI)
Men	Any psychiatric diagnosis	27,138	978 / 1,011	1.59 (1.56-1.63)
				1.59 (1.35-1.87)
	Any psychiatric diagnosis (exl. substance use)	18,047	782 / 815	1.27 (1.24-1.31)
				1.30 (1.08-1.56)
	Substance use	18,680	788 / 821	1.90 (1.86-1.95)
				1.66 (1.39-1.98)
	Alcohol use	11,569	554 / 580	1.50 (1.46-1.54)
				1.33 (1.09-1.64)
	Drug use	11,864	592 / 626	2.27 (2.21-2.33)
				2.00 (1.63-2.46)
Women	Any psychiatric diagnosis	7,062	48 / 48	1.71 (1.61-1.82)
				2.00 (0.87-4.59)
	Any psychiatric diagnosis (exl. substance use)	5,486	65 / 66	1.27 (1.20-1.35)
				1.53 (0.82-2.86)
	Substance use	4,825	53 / 56	1.91 (1.80-2.03)
				1.89 (0.90-3.97)
	Alcohol use	2,961	41 / 40	1.20 (1.12-1.28)
				0.67 (0.30-1.50)
	Drug use	3,345	46 / 49	2.40 (2.26-2.55)
				2.50 (1.12-5.60)



Model
Standard model
Sibling model

Figure 3-3. Association between psychiatric disorders and general reoffending in individuals given community sentences stratified by sex

The hazard ratios were estimated using Cox proportional hazard regression. All risk factors and covariates were recorded at baseline (start of a sentence). Standard Cox regression model is adjusted for age. Sibling model is a fixed-effect model adjusted for age and for any unmeasured covariates shared between siblings discordant by a given risk factor.

Table 3-2. Association between individual psychiatric diagnoses and general reoffending in individuals given community sentences stratified by sex

The hazard ratios were estimated using Cox proportional hazard regression. All risk factors and covariates were recorded at baseline (start of a sentence). Standard Cox regression model is adjusted for age. Sibling model is a fixed-effect model adjusted for age and for any unmeasured covariates shared between siblings discordant by a given risk factor. I reported sibling models only if there were at least 10 discordant siblings in the cohort.

Outcome: **General reoffending**

Men (N = 70,643)	Full cohort		Comparison between same-sex full siblings	
Previous psychiatric disorder	*N with diagnosis*	*Standard model HR (95% CI)*	*Cases / disc. siblings*	*Sibling model HR (95% CI)*
Any psychiatric diagnosis	27,138	1.59 (1.56-1.63)	978 / 1011	1.59 (1.35-1.87)
Any psychiatric diagnosis (other than substance use)	18,047	1.27 (1.24-1.31)	782 / 815	1.30 (1.08-1.56)
Schizophrenia spectrum	2,032	1.37 (1.28-1.45)	112 / 127	1.67 (1.05-2.67)
Bipolar	690	1.02 (0.91-1.15)	39 / 43	1.58 (0.61-4.11)
Depression	5,447	1.04 (0.99-1.08)	278 / 300	0.88 (0.65-1.20)
Anxiety	5,604	1.21 (1.16-1.26)	327 / 353	1.34 (1.02-1.76)
Personality disorder	2,671	1.76 (1.67-1.85)	173 / 190	1.34 (0.92-1.95)
Attention-deficit hyperactivity	3,370	1.32 (1.26-1.39)	202 / 212	1.01 (0.71-1.44)
Other developmental or childhood	3,246	1.27 (1.21-1.34)	192 / 200	0.81 (0.58-1.14)
Substance use	18,680	1.90 (1.86-1.95)	788 / 821	1.66 (1.39-1.98)
Alcohol use	11,569	1.50 (1.46-1.54)	554 / 580	1.33 (1.09-1.64)
Drug use	11,864	2.27 (2.21-2.33)	592 / 626	2.00 (1.63-2.46)
Women (N = 11,772)	Full cohort		Comparison between same-sex full siblings	
Previous psychiatric disorder	*N with diagnosis*	*Standard model HR (95% CI)*	*Cases / disc. siblings*	*Sibling model HR (95% CI)*
Any psychiatric diagnosis	7,062	1.71 (1.61-1.82)	48 / 48	2.00 (0.87-4.59)
Any psychiatric diagnosis (other than substance use)	5,486	1.27 (1.20-1.35)	65 / 66	1.53 (0.82-2.86)
Schizophrenia spectrum	563	1.32 (1.16-1.50)	7 / 7	-
Bipolar	340	0.87 (0.72-1.06)	6 / 7	-
Depression	2,037	1.02 (0.94-1.10)	35 / 37	0.96 (0.39-2.38)
Anxiety	1,869	1.15 (1.06-1.25)	44 / 46	1.41 (0.64-3.08)
Personality disorder	1,324	1.52 (1.40-1.66)	25 / 26	1.39 (0.56-3.50)
Attention-deficit hyperactivity	608	1.29 (1.12-1.48)	8 / 9	-
Other developmental or childhood	777	1.46 (1.31-1.63)	19 / 19	2.14 (0.52-8.73)
Substance use	4,825	1.91 (1.80-2.03)	53 / 56	1.89 (0.90-3.97)
Alcohol use	2,961	1.20 (1.12-1.28)	41 / 40	0.67 (0.30-1.50)
Drug use	3,345	2.40 (2.26-2.55)	46 / 49	2.50 (1.12-5.60)

Sibling analyses showed that men with schizophrenia spectrum disorder, anxiety disorder, alcohol use disorder, and drug use disorder had a significantly higher risk of general reoffending compared to their same-sex full siblings, discordant for the tested diagnosis (Table 3-2: Sibling model). The associations between the other individual psychiatric disorders (depression, personality disorder, ADHD or other developmental disorder) and general reoffending became non-significant in the sibling models when such men were compared to their unaffected same-sex full siblings.

Receiving a first-time psychiatric diagnosis during the follow-up was associated with a higher risk of general reoffending in men (Table 3-6). For new substance use diagnosis, the hazard ratio was 2.02 (95% CI 1.90-2.16). For other new psychiatric diagnoses, the hazard ratio was 1.37 (95% CI 1.20-1.46).

In women, hazard ratios for individual disorders ranged from 0.87 (95% CI 0.72-1.06) for bipolar disorder to 2.40 (95% CI 2.26-2.55) for drug use disorder. Overall, in the female cohort, 3,234 cases of general reoffending occurred during the first two years of the follow-up period, 822 of which were potentially attributable to psychiatric disorders. This corresponds to the PAF of 25.4% (95% CI 22.6-28.2%).

The female cohort contained a relatively small number of same-sex full siblings, discordant by a given diagnosis. Consequently, almost all associations between individual psychiatric disorders and general reoffending estimated using sibling models were non-significant with wide confidence intervals. The one exception was drug use disorder. Women with drug use disorder had a significantly higher risk of general reoffending compared to their same-sex full siblings without a drug use diagnosis.

Receiving a first-time psychiatric diagnosis during the follow-up was associated with a higher risk of general reoffending in women (Table 3-7). For new substance use diagnosis, the hazard ratio was 2.20 (95% CI 1.81-2.67). For other new psychiatric diagnoses, the hazard ratio was 1.57 (95% CI 1.32-1.86).

3.5.2 Comorbidity and general reoffending

Among individuals with psychiatric disorders, individuals with substance use comorbidity had a higher risk of general reoffending compared to individuals without such comorbidity (Table 3-3). This was the case for men and women. In men, hazard ratios for individual disorders with comorbid substance use ranged from 1.53 (95% CI 1.32-1.77) for bipolar disorder to 2.29 (95% CI 2.13-2.46) for schizophrenia spectrum disorder. However, except for anxiety disorder, individual disorders without comorbid substance use were not significantly associated with general reoffending.

In women, hazard ratios for individual disorders in comorbid substance use ranged from 1.52 (95% CI 1.20-1.92) for bipolar disorder to 2.51 (95% CI 2.15-2.93) for schizophrenia spectrum disorder. Similarly to men, except for anxiety disorder, the individual disorders without comorbid substance use were not significantly associated with general reoffending in women.

In men and women, having multiple psychiatric diagnoses (other than drug or alcohol use disorders) was associated with an increased risk of general reoffending (Appendix C8). The risk increased in a stepwise manner with each additional diagnosis. However, when individuals with and without comorbid substance use were analysed separately, the stepwise increase in the risk of general reoffending was no longer present. Overall, in our cohort, the relationship between having multiple disorders and general reoffending was fully mediated by comorbid substance use diagnosis.

3.5.3 Psychiatric diagnoses and violent reoffending

Having any prior psychiatric diagnosis at the start of a community sentence was associated with an increased risk of violent reoffending (Figure 3-2; Figure 3-4; Table 3-4). In men, all individual psychiatric diagnoses, except for bipolar disorder, were associated with an increased risk of violent reoffending. Hazard ratios for individual disorders in men ranged from 1.13 (95% CI 0.91-1.40) for bipolar disorder to 2.18 (95% CI 2.00-2.37) for personality disorder. Overall, in the male cohort, 6,219 cases of violent reoffending occurred during the first two years of the follow-up, 1,039 of which are attributable to psychiatric disorders. This corresponds to the PAF of 16.7% (95% CI 15.3-18.1%).

Sibling analyses showed that men with schizophrenia spectrum disorder, personality disorder, alcohol use disorder, and drug use disorder had a higher risk of violent reoffending compared to their same-sex full siblings, discordant by a given diagnosis (Table 3-4). The associations between the other individual psychiatric disorders (depression, anxiety, ADHD or other developmental disorder) and violent reoffending became non-significant in the sibling models when such men were compared to their unaffected same-sex full siblings.

Receiving a first-time psychiatric diagnosis during the follow-up was associated with a higher risk of general reoffending in men (Table 3-6). For new substance use diagnosis, the hazard ratio was 1.92 (95% CI 1.76-2.10). For any other new psychiatric diagnosis, the hazard ratio was 1.48 (95% CI 1.34-1.62).

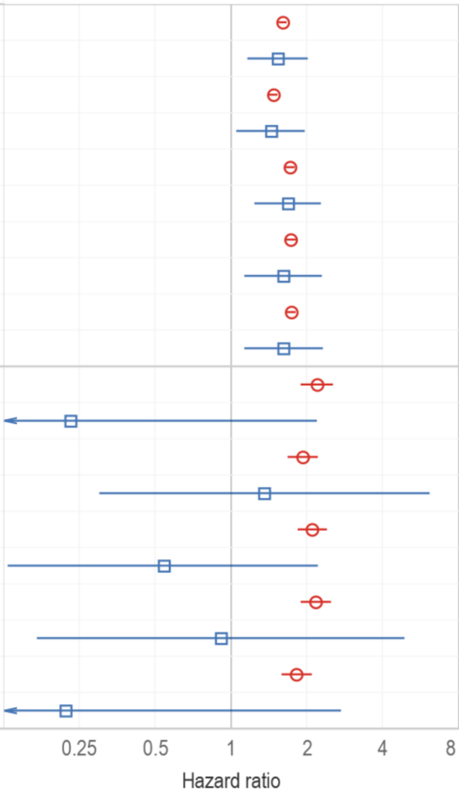
Table 3-3. General reoffending in individuals given community sentences with psychiatric disorders with and without substance use disorder comorbidity

Hazard ratios were estimated by comparing individuals with psychiatric diagnoses to individuals without known psychiatric diagnoses. *Compared with individuals without any psychiatric disorder.

Outcome: **General reoffending**

Men (N = 70,643)				
	Incidence of reoffending Reoffended / N of individuals with disorder (%)		Hazard ratio (95% CI)* Adjusted for age	
Previous psychiatric disorder	*With substance use*	*Without substance use*	*With substance use*	*Without substance use*
Any psychiatric (other than substance use)	10,777 / 18,680 (58%)	3,504 / 8,458 (41%)	1.92 (1.87-1.96)	1.07 (1.03-1.11)
Schizophrenia spectrum	791 / 1,368 (58%)	248 / 664 (37%)	2.29 (2.13-2.46)	0.95 (0.84-1.08)
Bipolar	181 / 451 (40%)	85 / 239 (36%)	1.53 (1.32-1.77)	1.01 (0.82-1.25)
Depression	1,534 / 3,220 (48%)	774 / 2,227 (35%)	1.61 (1.53-1.70)	0.92 (0.85-0.98)
Anxiety	1,573 / 2,841 (55%)	1,126 / 2,763 (41%)	2.03 (1.93-2.14)	1.07 (1.00-1.13)
Women (N = 11,772)				
	Incidence of reoffending Reoffended / N of individuals with disorder (%)		Hazard ratio (95% CI)* Adjusted for age	
Previous psychiatric disorder	*With substance use*	*Without substance use*	*With substance use*	*Without substance use*
Any psychiatric (other than substance use)	2,266 / 4,825 (47%)	720 / 2,237 (32%)	2.01 (1.88-2.14)	1.18 (1.08-1.29)
Schizophrenia spectrum	192 / 369 (52%)	55 / 194 (28%)	2.51 (2.15-2.93)	1.03 (0.79-1.35)
Bipolar	75 / 234 (32%)	25 / 106 (24%)	1.52 (1.20-1.92)	0.86 (0.58-1.28)
Depression	481 / 1,255 (38%)	232 / 782 (30%)	1.69 (1.52-1.87)	1.11 (0.97-1.28)
Anxiety	476 / 1,050 (45%)	269 / 819 (33%)	1.96 (1.76-2.17)	1.20 (1.05-1.37)

	Diagnosis	N w. diag.	Cases/siblings	HR (95% CI)
Men	Any psychiatric diagnosis	27,138	978 / 1,011	1.60 (1.54-1.66)
				1.53 (1.16-2.01)
	Any psychiatric diagnosis (exl. substance use)	18,047	782 / 815	1.47 (1.41-1.53)
				1.44 (1.05-1.96)
	Substance use	18,680	788 / 821	1.71 (1.64-1.78)
				1.68 (1.24-2.27)
	Alcohol use	11,569	554 / 580	1.72 (1.63-1.80)
				1.61 (1.13-2.28)
	Drug use	11,864	592 / 626	1.73 (1.65-1.81)
				1.61 (1.13-2.30)
Women	Any psychiatric diagnosis	7,062	48 / 48	2.19 (1.88-2.54)
				0.23 (0.02-2.18)
	Any psychiatric diagnosis (exl. substance use)	5,486	65 / 66	1.92 (1.68-2.20)
				1.35 (0.30-6.10)
	Substance use	4,825	53 / 56	2.09 (1.83-2.39)
				0.54 (0.13-2.20)
	Alcohol use	2,961	41 / 40	2.16 (1.88-2.49)
				0.91 (0.17-4.87)
	Drug use	3,345	46 / 49	1.81 (1.58-2.08)
				0.22 (0.02-2.73)



Model
Standard model
Sibling model

Figure 3-4. Association between psychiatric disorders and violent reoffending in individuals given community sentences stratified by sex

The hazard ratios were estimated using Cox proportional hazard regression. All risk factors and covariates were recorded at baseline (start of a sentence). Standard mode is adjusted for age. Sibling model is a fixed-effect model adjusted for age and for any unmeasured covariates shared between siblings discordant by a given risk factor.

Table 3-4. Association between individual psychiatric diagnoses and violent reoffending in individuals given community sentences stratified by sex

The hazard ratios were estimated using Cox proportional hazard regression. All risk factors and covariates were recorded at baseline (start of a sentence). Standard Cox regression model is adjusted for age. Sibling model is a fixed-effect model adjusted for age and for any unmeasured covariates shared between siblings discordant by a given risk factor. I reported sibling models only if there were at least 10 discordant siblings in the cohort.

Outcome: **Violent reoffending**

Men (N = 70,643)	Full cohort		Comparison between same-sex full siblings	
Previous psychiatric disorder	*N with diagnosis*	*Standard. HR (95% CI)*	*Cases / disc. siblings*	*Sibling. HR (95% CI)*
Any psychiatric diagnosis	27,138	1.60 (1.54-1.66)	978 / 1011	1.53 (1.16-2.01)
Any psychiatric diagnosis (other than substance use)	18,047	1.47 (1.41-1.53)	782 / 815	1.44 (1.05-1.96)
Schizophrenia spectrum	2,032	1.84 (1.67-2.02)	112 / 127	3.70 (1.37-10.0)
Bipolar	690	1.13 (0.91-1.40)	39 / 43	1.15 (0.22-6.04)
Depression	5,447	1.10 (1.02-1.19)	278 / 300	0.92 (0.55-1.54)
Anxiety	5,604	1.28 (1.20-1.38)	327 / 353	1.52 (0.91-2.54)
Personality disorder	2,671	2.18 (2.00-2.37)	173 / 190	2.03 (1.00-4.09)
Attention-deficit hyperactivity	3,370	1.57 (1.44-1.71)	202 / 212	1.26 (0.72-2.20)
Other developmental or childhood	3,246	1.68 (1.56-1.82)	192 / 200	1.20 (0.70-2.04)
Substance use	18,680	1.71 (1.64-1.78)	788 / 821	1.68 (1.24-2.27)
Alcohol use	11,569	1.72 (1.63-1.80)	554 / 580	1.61 (1.13-2.28)
Drug use	11,864	1.73 (1.65-1.81)	592 / 626	1.61 (1.13-2.30)
Women (N = 11,772)	Full cohort		Comparison between same-sex full siblings	
Previous psychiatric disorder	*N with diagnosis*	*Standard. HR (95% CI)*	*Cases / disc. siblings*	*Sibling. HR (95% CI)*
Any psychiatric diagnosis	7,062	2.19 (1.88-2.54)	48 / 48	0.23 (0.02-2.18)
Any psychiatric diagnosis (other than substance use)	5,486	1.92 (1.68-2.20)	65 / 66	1.35 (0.30-6.10)
Schizophrenia spectrum	563	2.44 (1.94-3.08)	7 / 7	-
Bipolar	340	1.30 (0.86-1.95)	6 / 7	-
Depression	2,037	1.09 (0.91-1.31)	35 / 37	0.02 (0.00-Inf)
Anxiety	1,869	1.52 (1.29-1.79)	44 / 46	0.25 (0.03-2.25)
Personality disorder	1,324	2.61 (2.21-3.07)	25 / 26	3.58 (0.36-35.7)
Attention-deficit hyperactivity	608	1.71 (1.28-2.28)	8 / 9	-
Other developmental or childhood	777	2.43 (1.99-2.97)	19 / 19	0.00 (0.00-Inf)
Substance use	4,825	2.09 (1.83-2.39)	53 / 56	0.54 (0.13-2.20)
Alcohol use	2,961	2.16 (1.88-2.49)	41 / 40	0.91 (0.17-4.87)
Drug use	3,345	1.81 (1.58-2.08)	46 / 49	0.22 (0.02-2.73)

Table 3-5. Violent reoffending in individuals given community sentences with psychiatric disorders with and without substance use disorder comorbidity

Hazard ratios were estimated by comparing individuals with psychiatric diagnoses to individuals without known psychiatric diagnoses. *Compared with individuals without any psychiatric disorder.

Outcome: **Violent reoffending**

Men (N = 70,643)				
	Incidence of reoffending Reoffended / N of individuals with disorder (%)		Hazard ratio (95% CI)* Adjusted for age	
Previous psychiatric disorder	*With substance use*	*Without substance use*	*With substance use*	*Without substance use*
Any psychiatric (other than substance use)	3,272 / 18,680 (18%)	1,264 / 8,458 (15%)	1.78 (1.70-1.86)	1.30 (1.22-1.38)
Schizophrenia spectrum	314 / 1,368 (23%)	111 / 664 (17%)	2.83 (2.52-3.17)	1.53 (1.27-1.85)
Bipolar	50 / 451 (11%)	32 / 239 (13%)	1.56 (1.18-2.06)	1.37 (0.96-1.93)
Depression	476 / 3,220 (15%)	244 / 2,227 (11%)	1.69 (1.54-1.86)	1.01 (0.89-1.15)
Anxiety	490 / 2,841 (17%)	385 / 2,763 (14%)	1.99 (1.81-2.19)	1.24 (1.12-1.37)
Women (N = 11,772)				
	Incidence of reoffending Reoffended / N of individuals with disorder (%)		Hazard ratio (95% CI)* Adjusted for age	
Previous psychiatric disorder	*With substance use*	*Without substance use*	*With substance use*	*Without substance use*
Any psychiatric (other than substance use)	479 / 4,825 (10%)	153 / 2,237 (7%)	2.48 (2.12-2.90)	1.65 (1.35-2.03)
Schizophrenia spectrum	61 / 369 (17%)	21 / 194 (11%)	5.01 (3.75-6.71)	3.26 (2.06-5.17)
Bipolar	18 / 234 (8%)	6 / 106 (6%)	2.94 (1.81-4.79)	1.53 (0.68-3.45)
Depression	97 / 1,255 (8%)	41 / 782 (5%)	2.27 (1.79-2.89)	1.38 (0.99-1.93)
Anxiety	117 / 1,050 (11%)	61 / 819 (7%)	2.97 (2.37-3.71)	1.82 (1.37-2.41)

Table 3-6. The association between new psychiatric diagnosis during follow-up and reoffending in men given community sentence without prior identified psychiatric diagnoses

The hazard ratios are estimated by fitting Cox regression models with a first new diagnosis as a time-dependent covariate. The Cox models were additionally adjusted by covariates measured at the time of a sentence: age, sex, sociodemographic variables (employment status, civil status, years of education, receipt of income support), criminal history variables (prior crime, prior violent crime, index violent crime), history of self-harm prior to an index sentence. In the final model, new substance use and other new psychiatric diagnosis were entered together.

Men (N = 43,505)			**Progressive adjustment. Adjusted hazard ratios (95% CI)**				
New diagnosis during follow-up	N of individuals	Median time until new diagnosis in months (interquartile range)	Age	+ sociodemographic	+ criminal history	+ prior self-harm	+ new diagnoses
General reoffending							
No new diagnosis (reference)	38,889	NA	1	1	1	1	1
New substance use disorder	2,589	18 (IQR: 6-39)	2.02 (1.90 - 2.16)	1.84 (1.72 - 1.97)	1.76 (1.60 - 1.88)	1.75 (1.64 - 1.87)	1.73 (1.61 - 1.86)
Other new psychiatric disorder	2,956	21 (IQR: 8-43)	1.37 (1.20 - 1.46)	1.18 (1.10 - 1.27)	1.22 (1.13 - 1.31)	1.21 (1.13 - 1.30)	1.03 (0.96 - 1.12)
Violent reoffending							
No new diagnosis (reference)	38,889	NA	1	1	1	1	1
New substance use disorder	2,589	24 (IQR: 9-49)	1.92 (1.76 - 2.10)	1.74 (1.58 - 1.90)	1.71 (1.56 - 1.88)	1.71 (1.56 - 1.88)	1.65 (1.49 - 1.82)
Other new psychiatric disorder	2,956	28 (IQR: 11-55)	1.48 (1.34 - 1.62)	1.32 (1.20 - 1.46)	1.34 (1.21 - 1.48)	1.34 (1.21 - 1.48)	1.13 (1.02 - 1.26)

Table 3-7. The association between new psychiatric diagnosis during follow-up and reoffending in women given community sentence without prior identified psychiatric diagnoses

The hazard ratios are estimated by fitting Cox regression models with a first new diagnosis as a time-dependent covariate. The Cox models were additionally adjusted by covariates measured at the time of a sentence: age, sex, sociodemographic variables (employment status, civil status, years of education, receipt of income support), criminal history variables (prior crime, prior violent crime, index violent crime), history of self-harm prior to an index sentence. In the final model, new substance use and other new psychiatric diagnosis were entered together.

Women (N = 4,710)			**Progressive adjustment. Adjusted hazard ratios (95% CI)**				
New diagnosis during follow-up	N of individuals	Median time until new diagnosis in months (interquartile range)	Age	+ sociodemographic	+ criminal history	+ prior self-harm	+ new diagnoses
General reoffending							
No new diagnosis (reference)	3,762	NA	1	1	1	1	1
New substance use disorder	396	18 (IQR: 6-46)	2.20 (1.81 - 2.67)	2.07 (1.68 - 2.54)	1.98 (1.59 - 2.46)	1.98 (1.59 - 2.46)	1.89 (1.50 - 2.38)
Other new psychiatric disorder	758	25 (IQR: 8-47)	1.57 (1.32 - 1.86)	1.29 (1.07 - 1.55)	1.35 (1.12 - 1.62)	1.33 (1.11 - 1.61)	1.17 (0.96 - 1.42)
Violent reoffending							
No new diagnosis (reference)	3,762	NA	1	1	1	1	1
New substance use disorder	396	26 (IQR: 8-51)	3.05 (2.16 - 4.31)	2.99 (2.09 - 4.26)	2.95 (2.06 - 4.22)	2.96 (2.07 - 4.23)	2.74 (1.87 - 4.02)
Other new psychiatric disorder	758	29 (IQR: 10-53)	1.75 (1.24 - 2.47)	1.67 (1.18 - 2.39)	1.66 (1.17 - 2.37)	1.66 (1.17 - 2.37)	1.29 (0.88 - 1.88)

In women, all individual psychiatric diagnoses, except for bipolar disorder and depression, were associated with an increased risk of violent reoffending. Hazard ratios for individual disorders in women ranged from 1.09 (95% CI 0.91-1.31) for depression to 2.61 (95% CI 2.21-3.07) for personality disorder. Overall, in the female cohort, 188 of 467 violent reoffending cases were potentially attributable to psychiatric disorders, which corresponds to the PAF of 40.3% (95% CI 33.7-46.9%).

The female cohort contained a relatively small number of same-sex full siblings, discordant by a given diagnosis. Consequently, all associations between individual psychiatric disorders and violent reoffending estimated using sibling models were non-significant with wide confidence intervals.

Receiving a first-time psychiatric diagnosis during the follow-up was associated with a higher risk of general reoffending in women (Table 3-7). For new substance use diagnosis, the hazard ratio was 3.05 (95% CI 2.16-4.31). For other new psychiatric diagnosis, the hazard ratio was 1.75 (95% CI 1.24-2.47).

3.5.4 Comorbidity and violent reoffending

Among individuals with psychiatric disorders, individuals with substance use comorbidity had a higher risk of violent reoffending compared to individuals without such comorbidity (Table 3-5).

This was the case for men and women. In men, hazard ratios for individual disorders accompanied by comorbid substance use ranged from 1.56 (95% CI 1.18-2.06) for bipolar disorder to 2.83 (95% CI 2.52-3.17) for schizophrenia spectrum disorder. For individual disorders without comorbid substance use, hazard ratios ranged from 1.01 (95% CI 0.89-1.15) for depression to 1.53 (95% CI 1.27-1.85) for bipolar disorder.

In women, hazard ratios for individual disorders accompanied by comorbid substance use ranged from 2.27 (95% CI 1.79-2.89) for depression to 5.01 (95% 3.75-6.71) for schizophrenia spectrum disorder. For individual disorders without comorbid substance use, hazard ratios ranged from 1.38 (95% CI 0.99-1.93) for depression to 3.26 (95% CI 2.06-5.17) for bipolar disorder.

In men and women, having multiple psychiatric diagnoses (other than drug or alcohol use disorders) was associated with an increased risk of violent reoffending (Appendix 8). The risk increased in a stepwise manner with each additional diagnosis. This pattern persisted even when individuals with and without comorbid substance use were analysed separately. Overall, in our cohort, the relationship between having multiple disorders and violent reoffending was only partially mediated by comorbid substance use diagnosis.

3.5.5 Sensitivity analyses

The findings did not differ when I separately analysed individuals sentenced before and after 2001 or used imputed data.

3.6 Discussion

I examined the association between psychiatric disorders and general and violent reoffending in a Swedish nationwide population-based study of 82,386 individuals given community sentences between 1991 and 2013. I followed individuals from the day their community sentence began until the date of committing any new violent or other offence. Our study has three principal findings.

First, psychiatric disorders were significantly associated with an increased risk of general and violent reoffending. The magnitude of this association was comparable to the pooled estimate from preliminary studies in community sentenced populations (Yukhnenko et al., 2020). The effects of psychiatric disorders were larger in the female cohort. The causal mechanisms underlying such observed relationships are likely to have both direct (for example, the impact of psychotic symptoms on violent reoffending) and indirect (for example, mediated by the impact of the psychiatric disorder on employment and social support) components. Moreover, the relationships between psychiatric disorders and other measured covariates, including criminological variables, are likely bidirectional. Life-course persistent delinquent lifestyle is likely to detrimentally affect the mental health of an individual (Reising, 2021). These bidirectional relationships complicate causal interpretations of the effects obtained from adjusted statistical models.

Another possible approach to obtain adjusted estimates suitable for causal interpretation was implementing sibling comparisons to adjust for unmeasured familial confounding. I noted minimal evidence of familial confounding on the association between most psychiatric disorders and risk of general or violent reoffending in men. This finding supports the causal interpretation of the effect of some individual psychiatric disorders on reoffending since sibling comparisons rule out the effect of any environmental differences that vary between families (Lahey & D'Onofrio, 2010). Substantial attenuation of the effect in the sibling model suggests either that certain psychiatric disorders have common familial causes with reoffending or that the effect of the psychiatric disorders is mediated through factors shared in the family (Sjölander & Zetterqvist, 2017). In women, adjustment for familial confounding rendered the findings largely non-significant (with the exception of the impact of drug use on the risk for general reoffending), which is likely because of the low statistical power inherent in the small numbers of reoffending women in the cohort with same-sex siblings.

Establishing the potential causal effect of psychiatric disorders on reoffending would also require establishing the clear temporal order between the manifestation of symptoms and committing an offence (Grzymala-Busse, 2011). In the study cohort, the first-time psychiatric diagnosis during the follow-up substantially increased the risk of subsequent general and violent reoffending. This finding provides additional support to the causal interpretation of the effect of psychiatric disorder on criminal recidivism. However, this finding should be interpreted with caution given the high underreporting of psychiatric symptoms in individuals given community sentences. It is possible that many first-time psychiatric diagnoses identified during the follow-up manifested prior to the index sentence but were not officially diagnosed.

The population attributable fraction estimates suggest that a substantial number of general and especially violent reoffending cases were potentially attributable to psychiatric disorders and substance use disorders. These estimations assume causality and, therefore, are likely to overestimate the true effect and should be interpreted with caution. The PAF can be interpreted as a maximum possible effect of full treatment of the psychiatric disorders in this population and provides one measure of population impact, which may help to inform public health measures and health resource allocation within criminal justice services.

Second, the magnitude of the association between psychiatric disorders and reoffending varied depending on individual diagnosis. Schizophrenia spectrum disorders, personality disorders, and substance use disorders were more strongly associated with violent offending than other disorders. Schizophrenia spectrum disorders were significantly associated with general and violent reoffending, and this association remained significant even after adjustment for unmeasured familial confounding in sibling analyses. There was no apparent association between mood disorders and the increased risk of general and violent reoffending in our cohort. This finding contrasts with prior research in released prisoners (Z. Chang, Larsson, et al., 2015) but corroborates prior research in individuals given community sentences (Grann et al., 2008). These findings might reflect different symptom pathways in lower versus higher-risk individuals with mood disorders. For example, in individuals diagnosed with bipolar disorder, having antisocial personality symptoms, higher impulsivity, and a predominantly manic course of illness have been associated with more criminal convictions (Swann et al., 2011). It is possible that diagnosed individuals with higher impulsivity and antisocial personality traits would be more likely to commit a severe violent offence and, subsequently, more likely to be imprisoned instead of given a community sentence.

Third, most of the increased risk for general reoffending in individuals with mental health disorders could be attributed to comorbid substance use problems. This is perhaps unsurprising given that many such crimes are directly related to drug possession or substance intoxication. Almost a quarter of all sentences in

Sweden were handed down for drug or alcohol-related crimes (Swedish National Council for Crime Prevention, 2021a) during the years of the study. However, the increased risk for violent reoffending in individuals with mental health disorders could not be simply attributed to comorbid substance use problems: comorbid substance use disorder only partially explained the association with reoffending as was previously demonstrated in a released prisoner cohort (Z. Chang, Larsson, et al., 2015). Having a schizophrenia spectrum disorder was associated with violent reoffending even in individuals without known substance use comorbidity. It is possible that some symptoms of schizophrenia, such as reality distortion symptomatology or a lack of empathy, may be specifically associated with violence but not general criminality, and these factors warrant further specification.

Fourth, controlling for comorbid substance use, individuals with more psychiatric diagnoses had a higher risk of violent reoffending. The stepwise increase in risk suggests the potential cumulative effect of psychiatric symptoms on violence. However, controlling for comorbid substance use, the was no such relationship between the number of diagnoses and the risk of general reoffending. One possible explanation could be that individuals with several different diagnoses have a higher chance of having symptoms specifically associated with violence. In this scenario, the number of psychiatric diagnoses indicates the probability of having these symptoms. Another possible explanation is that violent behaviour and psychopathology share a common genetic *p* factor (Selzam et al., 2018) that is not associated with general criminality. In this scenario, the number of psychiatric diagnoses indicates higher levels of genetic vulnerability. These findings warrant further investigation.

3.7 Strengths and limitations

To my knowledge, the current study is the first to examine risk factors for reoffending in community sentenced individuals using a sibling-comparison design to account for unmeasured familial confounding. I explored the mediating role of substance use in this association and demonstrated that substance use differentially contributed to different types of reoffending. The study was conducted on a large nationwide cohort with sufficient power to explore all associations of interest (with the potential exception of the sibling comparison analyses in the female reoffending group).

The study has several limitations. I relied on data from patient registers for ascertainment of psychiatric diagnoses. I did not have access to outpatient data until 2001. Hence, my approach was likely to lead to a conservative estimate for the prevalence of psychiatric disorders. Because of the nature of our data, I could not fully account for the unobserved confounding of the relationship between substance use and other psychiatric disorders. Therefore, a detailed formal mediation analysis of this relationship was not possible (Lapointe-Shaw et al., 2018). The research was done in one country with a freely accessible public health system. This might lead to conservative estimates of the effect of psychiatric disorders on repeated criminal behaviour, as sentenced individuals were more likely to receive appropriate community interventions than in other countries. The outcome definition and estimates were also specific to Sweden. Since the recidivism data are highly sensitive to definition, crime detection, and legal practices (Yukhnenko et al., 2019), this further limits the generalisability of these estimates.

3.8 Conclusion

Individuals with psychiatric disorders are overrepresented within the criminal justice system, and community sentencing is no exemption. My findings underscore the need for better assessment of mental health needs of sentenced individuals along with improved access to treatment. The facilitation of voluntary participation in mental health treatment, housing and vocational programmes may be one way forward. Substance use should be considered the most important target for intervention in the community sentenced population. This study also highlights the importance of continuous risk assessment throughout the supervision process. Such continuous risk monitoring could allow to proactively respond to an individual’s mental health needs and provide timely interventions.

Chapter 4. Psychiatric disorders and mortality: a national cohort study of individuals given community sentences in Sweden

4.1 Abstract

Individuals given community sentences have higher mortality rates compared to the general population. Since psychiatric disorders are highly prevalent in community sentenced populations and associated with premature death in the general population, they are potential targets for interventions to reduce mortality in sentenced individuals. I examined the association between psychiatric disorder and mortality in a large nationwide cohort of individuals given community sentences in Sweden, employing a sibling design to account for potential unmeasured familial confounding.

I did a longitudinal cohort study of 109,751 individuals given community sentences between 1991 and 2013 in Sweden. During the follow-up, most potentially preventable deaths occurred in individuals with a psychiatric diagnosis. The proportion of preventable deaths was higher among younger individuals with a psychiatric diagnosis. The leading cause of death was suicide. Having substance use or any other psychiatric disorder at the time of a sentence or receiving a new diagnosis during the follow-up was associated with an increased risk of all-cause and external-cause mortality in the study cohort. The association remained significant even after controlling for measured sociodemographic factors, prior criminal history, the history of self-harm, and unmeasured familial confounding (in sibling comparisons). Comorbid substance use partially mediated the association of psychiatric disorders with mortality.

4.2 Introduction

Chapter 3 identified the association between psychiatric disorders and criminal recidivism in individuals given community sentences. The identified association was mostly mediated by comorbid substance use. Chapter 3 also underlined the importance of timely detection of psychiatric disorders and subsequent interventions in sentenced individuals. However, the potential intervention programmes and policies do not have to revolve exclusively around lowering criminal recidivism. They also have the potential to improve general health outcomes and prevent excessive mortality. The current chapter examines the association between psychiatric disorders and mortality in individuals given community sentences to facilitate this.

Prior studies have identified high mortality rates in correctional populations (Graham et al., 2015; Wildeman et al., 2019). A large portion of correctional populations consists of individuals given community sentences, which include community service, probation and mandatory community treatment. Community sentences often aim to provide sentenced individuals with better access to healthcare and welfare services, thus reducing the risk of potential adverse health outcomes and premature death. However, similarly to prisoners, the mortality rates among individuals given community sentences are substantially higher compared to the general population (Sattar, 2001). Over the period from 2015 to 2020, in the UK and Wales, 2,415 deaths of individuals under community supervision occurred (Ministry of Justice, 2020b). 30% of those deaths were due to natural causes, and 34% were self-inflicted. In the USA, from 2001 to 2012, the mortality rates of individuals on probation were higher than both prisoners and the general population (Wildeman et al., 2019). High mortality rates in individuals given community sentences indicate the apparent need to examine the role of risk factors of premature death. The identified risk factors can serve as potential targets for interventions and inform policies that aim to reduce potentially preventable deaths. Psychiatric disorders constitute a group of such risk factors.

Psychiatric disorders have been overrepresented in individuals given community sentences across many countries (Brooker et al., 2012; Sirdifield, 2012). The most common diagnoses included substance use disorders, psychotic disorders, antisocial personality disorder, mood disorders, anxiety, and adjustment disorders. Prior research has demonstrated the association between psychiatric disorders and the risk of all-cause mortality and suicide in the general population (Chesney et al., 2014; Lawrence et al., 2010) as well as in released prisoners (Z. Chang, Lichtenstein, et al., 2015; Spittal et al., 2019). However, very few studies have examined the association between psychiatric disorders in individuals given community sentences (Gelsthorpe et al., 2012; Phillips et al., 2017). One such study conducted in Australia reported that involvement with mental health services had been associated with an increased risk of mortality in individuals on probation (Sodhi-Berry et al., 2015).

The need to examine risk factors for mortality in community sentences separately from custodial sentences is due to the substantial differences in the supervision environments. In community sentenced individuals with a higher risk of self-harm and drug overdose, these risks may be much harder to manage (Phillips et al., 2018). During their sentence, individuals under community supervision have greater access to means of suicide and illicit substances compared to prisoners. On the other hand, individuals given community sentences can maintain employment and receive additional social support. Overall, as other researchers have pointed out, deaths outside a custodial setting are less understood compared to deaths in prison (Phillips et al., 2017). This is worrying since failure to take into account the differences between different correctional settings can lead to incorrect conclusions with detrimental policy implications.

Research examining the effect of psychiatric disorders on mortality in community sentenced populations is lacking, despite a large number of individuals being given community sentences, many of whom have been diagnosed with psychiatric disorders, and community supervision being substantially different from custodial supervision. Published studies on this topic often have substantial limitations. These limitations include the lack of diagnostic specificity and the follow-up period limited to post-sentence supervision (Gelsthorpe et al., 2012; Sodhi-Berry et al., 2015). The lack of diagnostic specificity entails using broad categories such as ‘history of mental health service contact’ instead of codified medical diagnostic data. Restricting the follow-up period only to post-sentence supervision (1-5 years) limits the assessment of the potential long-term effects of risk factors. The current study aimed to address these limitations.

In this population-based longitudinal study of individuals given community sentences, I investigated the association between psychiatric disorders and post-sentence mortality. I aimed to address three main questions. First, whether psychiatric disorders are associated with all-cause and external-cause mortality after receipt of a community sentence. Second, whether the comorbid substance use mediates the association between other psychiatric disorders and mortality in individuals given community sentences. Third, what the population effects of identified risk factors on all-cause and external-cause mortality are in individuals given community sentences. To account for potential familial confounding, I additionally utilised sibling comparison.

4.3 Approach

The study is conducted using the same approach as described in Chapter 3. I additionally calculated mortality rates. Mortality rates in the context of survival analysis are estimated as the probability of dying per *n* person-years of observed survival (Fink & Brown Jr, 2006). To be consistent with prior research in released prisoners (Z. Chang, Lichtenstein, et al., 2015), I estimated mortality per 100,000 person-years.

4.4 Methods

I followed STROBE guidelines (von Elm et al., 2007) for the reporting of observational studies (see Appendix D1 for the checklist). The current study followed a general design of a previously published research in released prisoners (Z. Chang, Lichtenstein, et al., 2015).

4.4.1 Study setting

I linked several longitudinal, nationwide Swedish registers: the National Crime Register, which contains information about criminal offences and convictions since 1973; the National Patient Register, which provides information about diagnoses for individuals admitted to inpatient hospitals (since 1973) and outpatient care (since 2001); Migration Register, containing dates of migration to and from Sweden; Cause of Death Register, which includes information on dates and causes of dates since 1958; Multi-Generation Register, containing information about biological relationships for all individuals living in Sweden since 1933; Longitudinal Integration Database for Health Insurance and Labour Market studies that include yearly estimations of income benefit reception, marital and employment status, and education since 1990. In Sweden, all residents (including immigrants) have a unique personal identifier used in all national registers, thus enabling data linkage (Ludvigsson et al., 2009).

This study was approved by the Regional Ethics Committee at the Karolinska Institutet (Stockholm, Sweden).

4.4.2 Participants

I included all adult (18 years old or older) Swedish residents who received any community sentence at any point from November 1, 1991, to December 31, 2013. Community sentences included conditional sentences with community service, probation with community service, and probation with contract treatment. I selected individuals whose sentences came into legal force based on the information from the National Crime Register. This approach allowed us to exclude individuals whose cases were successfully appealed or dismissed. For each individual, I used the date when a community sentence came into force as the start of the follow-up period. If an individual received multiple community sentences, then the index sentence was selected at random.

I excluded individuals born before 1958 as they would not have a continuous criminal record on the National Crime Register. I additionally identified full siblings within the cohort using the Multi-Generation Register.

4.4.3 Measures

I extracted sociodemographic information, criminal and medical history at the start of the community sentence. The sociodemographic information included sex, age, years of education, marital status, employment information, and the receipt of income support.

Criminal history included dates of all prior convictions that were legally enforced and their corresponding crime codes. I separately recorded if the index sentence was an individual's first entry in the Crime Register or if they had been previously sentenced. I also recorded if an individual was sentenced for a violent offence in the past and whether their index offence was a violent offence. The violent offence was defined as homicide, assault, robbery, arson, any sexual offence, illegal threats, or intimidation.

Medical history included a history of any psychiatric diagnosis received before the index sentence. I utilised a hierarchical approach to coding main diagnostic categories in line with previous research using Swedish national registers (Z. Chang, Larsson, et al., 2015). The hierarchy was as follows: schizophrenia spectrum disorders, bipolar disorder, depression, and anxiety disorder. If an individual had a diagnosis of schizophrenia and any other diagnoses, I classified that individual as having schizophrenia. If an individual did not have schizophrenia but had bipolar disorder and depression or anxiety, I classified that individual as having bipolar disorder, and so on.

To explore the effects of comorbidity between psychiatric disorders, I also investigated alcohol use disorder, drug use disorder, personality disorder, attention-deficit hyperactivity disorder, and other developmental or childhood disorders. I did not use a hierarchical approach for these comorbidities but examined whether they were present or not. ICD codes for the psychiatric diagnoses are the same as were used in Chapter 3 (Appendix C3). Also, to assess the cumulative effect of multiple diagnoses, I recorded the number of distinct diagnostic categories that an individual belonged to.

4.4.4 Missing data

0.7% of individuals within the cohort did not have demographic information and 4.1% did not have education data at baseline. A sensitivity analysis demonstrated that the results did not differ significantly if the missing data were imputed. Thus, in the final analysis, I did not replace missing data by imputation or other methods.

4.4.5 Outcomes and censoring

The main outcome was death after receiving a community sentence. The Cause of Death Register includes all people who, at the time of death, were registered as residents in Sweden, regardless of whether the death occurred in Sweden or abroad. The underlying and contributing (secondary) causes of death are coded according to ICD-10 based on death certificates issued by physicians or forensic doctors. I extracted both all-cause mortality data and mortality data separated by ICD chapter in accordance with the underlying cause of death. Within external-cause mortality (deaths caused by environmental events and circumstances, ICD-10 Chapter XX), I further examined deaths by traffic and non-traffic accidents, suicide, and homicide. In keeping with previous work, I included undetermined deaths (ICD-10: Y10–Y34) as suicides.

All individuals were followed up until their death, permanent emigration from Sweden or the end of the available register (December 31, 2013).

4.4.6 Statistical analysis

I calculated mortality for all-cause death, 11 causes of death by ICD chapter, and sub-causes for external causes. Mortality rates were calculated as the number of deaths for a given cause divided by person-years at risk.

I used Kaplan-Meier survival curves to examine the timing of post-sentence mortality in individuals given community sentences. I plotted separate survival curves for individuals with and without substance use, and individuals with and without any other psychiatric disorders. I tested proportional hazards assumptions by visually examining the Kaplan-Meier curves and Schoenfeld residuals diagrams.

The explore the association between individual disorders and mortality, I fitted Cox proportional hazard models for each diagnosis investigated. To estimate the total effect of individual disorders on all-cause and external cause mortality, I fitted the models adjusted only for age and sex. To estimate the direct effect of individual disorders on all-cause and external cause mortality relative to measured sociodemographic and criminological factors, I fitted proportional hazard models with progressive adjustment. To examine whether unmeasured familial factors partially explained the association between psychiatric disorders and death, I fitted a fixed effect Cox regression model (Allison, 2009) to a cohort of full siblings given community sentences. The model was stratified by family, so each sibling within one family had the same baseline hazard. To test the effect of first-time diagnosis during the follow-up, I additionally fitted a Cox regression model with time-varying covariates in the subsample of individuals without a psychiatric diagnosis at sentence (Zhang et al., 2018).

To estimate the population effect of substance use and other psychiatric disorders on mortality, I calculated the population attributable fraction (PAF). The

PAF estimates the proportion of deaths that can be attributed to a given risk factor, assuming that a causal association exists. To calculate PAF and corresponding CIs, I used the model-based adjusted attributable fraction function for Cox proportional hazard models in *AF* package for R (Dahlqwist & Sjolander, 2019).

The analyses were done in R using *survival* package (Therneau, 2021). The visualizations were created using the *survminer* package (Kassambara et al., 2017), the *ggplot2* package (Wickham et al., 2016), and Tableau Desktop (Tableau Software, 2021).

4.5 Results

I identified 109,751 individuals (94,221 men and 15,530 women), who received at least one community sentence in Sweden from November 1, 1991, to December 31, 2013 (see Appendix D2 for the selection flowchart). These individuals were followed up for 685,453.2 person-years after their index sentence (see Appendix D3 for overall survival). I identified 9,439 full siblings from 4,479 families who had been given a community sentence (see Appendix D4 for sibling model estimation of individual diagnoses).

The baseline sociodemographic and criminological information, psychiatric diagnoses, and follow-up data are presented in Table 4-1. Out of 109,751, 34,918 (31.8%) individuals had prior substance use disorder diagnoses, and 31,748 (28.9%) individuals had other psychiatric diagnoses. A higher proportion of women in the cohort (61.4%) had been diagnosed with a substance use disorder or other psychiatric disorder compared to men (41.2%). The univariate associations between baseline characteristics and death are presented in Appendix D5. The baseline measures demonstrated some collinearity (Appendix D6). Many psychiatric diagnoses were associated with others, reflecting psychiatric comorbidity.

In total, 5,749 individuals died during the follow-up of the study with 2,709 deaths (47% of all deaths) caused by external factors (Table 4-2). 3,125 deaths (54%) occurred within a median follow-up time of 5.4 years (IQR: 2.7-8.8) after receiving a community sentence. 1,799 deaths (31%) occurred within 3 years after being sentenced. The overall all-cause mortality rate was 839 deaths per 100,000 person-years. The mortality rate for external cause mortality was 395 deaths per 100,000 person-years. The most common cause of death in the study cohort was suicide. Cardiovascular disorders, traffic accidents and cancer were other major causes of death. Most deaths from external causes occurred in individuals younger than 30 years old with suicide being the leading cause of death in this group. Most deaths from non-external causes occurred in individuals older than 50 years old with diseases of the circulatory system being a leading cause of death in this group.

Table 4-1. Baseline characteristics and follow-up data of adult individuals receiving community sentences from November 1, 1991, to December 31, 2013

573 men and 69 women have missing values for marital status, employment, and income support. 3,557 men and 497 women have missing values for education.

	Men	Women	Total
Follow-up data			
Number of individuals	94,221 (85.8%)	15,530 (14.2%)	109,751 (100.0%)
Person-years at risk	593,088.1	92,365.1	685,453.2
Follow-up time	5.5 (IQR: 2.7-8.8)	5.2 (IQR: 2.6-8.6)	5.4 (IQR: 2.7-8.8)
Time until death	4.9 (IQR: 2.4-8.5)	4.5 (IQR: 2.3-7.3)	4.8 (IQR: 2.4-8.4)
Median age at death	49.4 (IQR: 35.0-59.8)	49.9 (IQR: 39.2-58.2)	49.5 (IQR: 35.7-59.7)
Deaths during follow-up	5,096 (5.4%)	653 (4.2%)	5,749 (5.2%)
Deaths from external causes during follow-up	2,396 (2.5%)	313 (2.0%)	2,709 (2.5%)
Emigrated during follow-up	1,901 (2.0%)	232 (1.5%)	2,133 (1.9%)

Baseline characteristics				
Any prior conviction		75,264 (79.9%)	10,735 (69.1%)	85,999 (78.4%)
Prior conviction for a violent crime		37,227 (39.5%)	3,095 (19.9%)	40,322 (36.7%)
Prior prison sentence		25,937 (27.5%)	2,186 (14.1%)	28,123 (25.6%)
Violent index sentence		40,292 (42.8%)	4,652 (30.0%)	44,944 (41.0%)
Median age at sentence		30 (IQR: 22-43)	35 (IQR: 24-45)	31 (IQR: 22-44)
Age groups				
	18-24 years	33,048 (35.1%)	4,065 (26.2%)	37,113 (33.8%)
	25-39 years	30,774 (32.7%)	5,417 (34.9%)	36,191 (33.0%)
	≥ 40 years	30,399 (32.3%)	6,048 (38.9%)	36,447 (33.2%)
Married or in a registered partnership		12,249 (13.0%)	2,424 (15.6%)	14,673 (13.4%)
Employed		39,142 (41.5%)	5,042 (32.5%)	44,184 (40.3%)
Highest level of education				
	< 9 yr	4,937 (5.2%)	844 (5.4%)	5,781 (5.3%)
	9-11 yr	78,289 (83.1%)	12,403 (79.9%)	90,692 (82.6%)
	≥ 12 yr	7,145 (7.6%)	1,739 (11.2%)	8,884 (8.1%)
Recipient of income support		32,583 (34.6%)	7,172 (46.2%)	39,755 (36.2%)
Any psychiatric disorder		38,807 (41.2%)	9,539 (61.4%)	48,346 (44.1%)
Any psychiatric disorder (other than substance use)		24,570 (26.1%)	7,178 (46.2%)	31,748 (28.9%)
	Schizophrenia spectrum disorder	2,930 (3.1%)	739 (4.8%)	3,669 (3.3%)
	Bipolar disorder	1,019 (1.1%)	452 (2.9%)	1,471 (1.3%)
	Depression	7,603 (8.1%)	2,745 (17.7%)	10,348 (9.4%)
	Anxiety disorder	7,462 (7.9%)	2,421 (15.6%)	9,883 (9.0%)
	Alcohol use disorder	18,690 (19.8%)	4,278 (27.5%)	22,968 (20.9%)
	Drug use disorder	16,714 (17.7%)	4,550 (29.3%)	21,264 (19.4%)
	Substance (drug or alcohol) use disorder	28,154 (29.9%)	6,764 (43.6%)	34,918 (31.8%)
	Personality disorder	3,885 (4.1%)	1,667 (10.7%)	5,552 (5.1%)
	Attention-deficit hyperactivity disorder	4,076 (4.3%)	702 (4.5%)	4,778 (4.4%)
	Other developmental or childhood disorder	3,974 (4.2%)	906 (5.8%)	4,880 (4.4%)
History of self-harm or prior suicide attempts		7,901 (8.4%)	2,975 (19.2%)	10,876 (9.9%)

Table 4-2. Mortality rates in individuals given community sentences in Sweden from 1991 until 2013

Data are n (%) or mortality per 100,000 person-years (95% CI). Causes are classified by ICD-10 chapters. *Out of 1,893 deaths caused by neoplasms, 556 (29%) were malignant neoplasms of digestive organs including 228 cases of malignant neoplasms of the liver. †Out of 401 deaths caused by diseases of the digestive system, 356 (89%) were caused by alcoholic liver disease.

	Men		Women		Overall	
Cause	Deaths	Mortality	Deaths	Mortality	Deaths	Mortality
All causes	5,096 (100%)	859 (836-883)	653 (100%)	707 (653-761)	5,749 (100%)	839 (817-860)
Certain infectious and parasitic diseases (chapter I)	129 (3%)	22 (18-26)	18 (3%)	19 (10-28)	147 (3%)	21 (18-25)
Neoplasms (chapter II)*	575 (11%)	97 (89-105)	81 (12%)	88 (69-107)	656 (11%)	96 (88-103)
Endocrine, nutritional and metabolic diseases (chapter IV)	82 (2%)	14 (11-17)	14 (2%)	15 (7-23)	96 (2%)	14 (11-17)
Mental and behavioural disorders (chapter V)	221 (4%)	37 (32-42)	23 (4%)	25 (15-35)	244 (4%)	36 (31-40)
Diseases of the nervous system (chapter VI)	60 (1%)	10 (8-13)	3 (0%)	3 (0-7)	63 (1%)	9 (7-11)
Diseases of the circulatory system (chapter IX)	892 (18%)	150 (141-160)	103 (16%)	112 (90-133)	995 (17%)	145 (136-154)
Diseases of the respiratory system (chapter X)	163 (3%)	27 (23-32)	23 (4%)	25 (15-35)	186 (3%)	27 (23-31)
Diseases of the digestive system (chapter XI)†	354 (7%)	60 (53-66)	47 (7%)	51 (36-65)	401 (7%)	59 (53-64)
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified (chapter XVIII)	190 (4%)	32 (27-37)	21 (3%)	23 (13-32)	211 (4%)	31 (27-35)
Other non-external causes (chapters III, VII, VIII, XII-XVII)	34 (1%)	6 (4-8)	7 (1%)	8 (2-13)	41 (1%)	6 (4-8)
External causes of morbidity and mortality (chapter XX)	2,396 (47%)	404 (388-420)	313 (48%)	339 (301-376)	2,709 (47%)	395 (380-410)
Traffic accidents	210 (4%)	35 (31-40)	12 (2%)	13 (6-20)	222 (4%)	32 (28-37)
Non-traffic accidents	130 (3%)	22 (18-26)	11 (2%)	12 (5-19)	141 (2%)	21 (17-24)
Suicide	1,004 (20%)	169 (159-180)	166 (25%)	180 (152-207)	1,170 (20%)	171 (161-180)
Homicide	118 (2%)	20 (16-23)	9 (1%)	10 (3-16)	127 (2%)	19 (15-22)

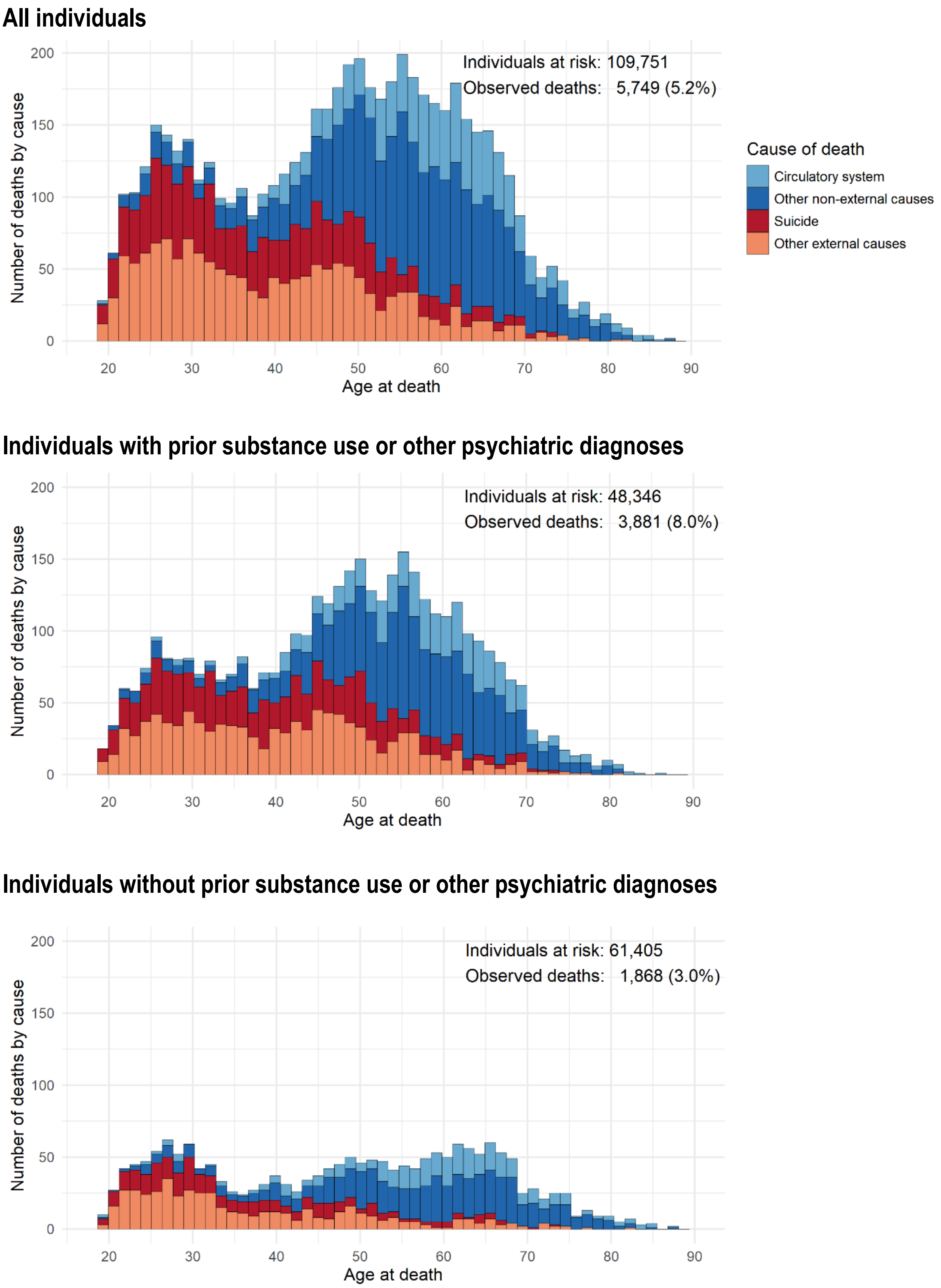
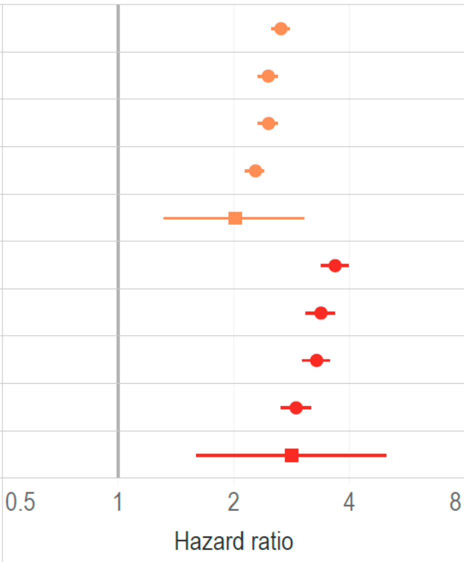


Figure 4-1. Deaths of individuals given community sentences during the follow-up period by primary cause of death, age at death, and prior psychiatric diagnosis

Suicide was a leading cause of external deaths; circulatory system diseases were a leading non-external cause of death.

A. Substance use disorder

Mortality	Model	HR (95% CI)
All-cause	Baseline	2.64 (2.51-2.79)
	+ demographic covariates	2.45 (2.32-2.59)
	+ criminal history	2.46 (2.32-2.60)
	+ prior self-harm	2.27 (2.14-2.41)
	Sibling model	2.01 (1.32-3.06)
External-cause	Baseline	3.66 (3.38-3.96)
	+ demographic covariates	3.36 (3.09-3.66)
	+ criminal history	3.27 (3.00-3.57)
	+ prior self-harm	2.89 (2.65-3.16)
	Sibling model	2.82 (1.60-4.99)



B. Other psychiatric disorder

Mortality	Model	HR (95% CI)
All-cause	Baseline	1.72 (1.63-1.82)
	+ demographic covariates	1.58 (1.50-1.68)
	+ criminal history	1.60 (1.51-1.69)
	+ prior self-harm	1.43 (1.35-1.52)
	Sibling model	1.27 (0.84-1.91)
External-cause	Baseline	2.40 (2.22-2.59)
	+ demographic covariates	2.22 (2.05-2.41)
	+ criminal history	2.24 (2.07-2.43)
	+ prior self-harm	1.93 (1.78-2.10)
	Sibling model	1.43 (0.83-2.49)

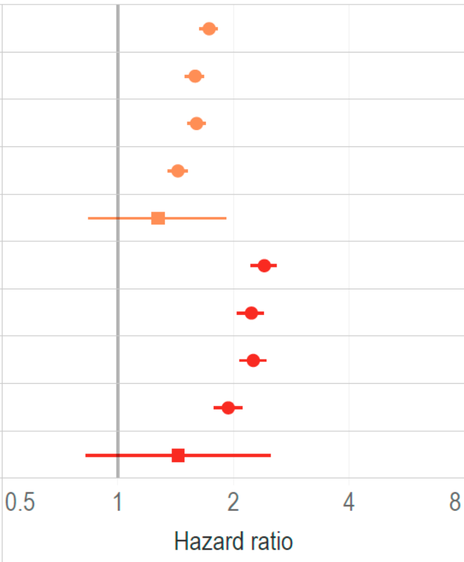


Figure 4-2. The association between mortality and prior psychiatric diagnosis.

The hazard ratios were derived by fitting standard Cox regression models, progressively adjusting for measured covariates, and a fixed-effect Cox regression sibling model. The baseline model was only adjusted for age and sex. The second model was additionally adjusted for sociodemographic covariates. In addition to the covariates from the second model, the third model was also adjusted for criminal history covariates, and so on. The sibling model was adjusted for age and sex.

Table 4-3. The association between new psychiatric diagnosis during the follow-up period and mortality in individuals given community sentence without a prior known psychiatric diagnosis

The hazard ratios are estimated by fitting Cox regression models with a first new diagnosis as a time-dependent covariate. The Cox models were progressively adjusted by covariates measured at the time of a sentence: age, sex, sociodemographic variables (employment status, civil status, years of education, receipt of income support), criminal history variables (prior crime, prior violent crime, index violent crime, having a pseudoreconviction), history of self-harm prior to an index sentence. The median time until a new diagnosis in months: 43 (IQR: 18-82). The median time until a new substance use diagnosis in months: 39 (IQR: 15-76).

		Adjusted Cox regression models, HR (95% CI)			
New diagnosis during follow-up	N	*Model 1.* *Age + sex*	*Model 2.* *Age + sex +* *sociodemographic factors*	*Model 3.* *Age + sex +* *sociodemographic factors +* *criminal history*	*Model 4.* *Age + sex +* *sociodemographic factors +* *criminal history +* *history of self-harm*
All-cause mortality					
No new diagnosis	44,933	1	1	1	1
New substance use diagnosis	11,282	4.00 (3.61 - 4.43)	3.86 (1.47 - 4.29)	3.86 (3.47 - 4.29)	3.83 (3.45 - 4.27)
Other new psychiatric diagnosis	10,817	2.82 (2.04 – 2.55)	2.21 (1.97 - 2.48)	2.24 (1.99 - 2.52)	2.23 (1.98 - 2.50)
All individuals	61,405				
External-cause mortality					
No new diagnosis	44,933	1	1	1	1
New substance use diagnosis	11,282	3.88 (3.34 - 4.51)	3.47 (2-96 - 4.07)	3.31 (2.83 - 3.89)	3.21 (2.83 - 3.89)
Other new psychiatric diagnosis	10,817	1.54 (1.29 - 1.83)	1.50 (1.26 - 1.80)	1.55 (1.30 - 1.86)	1.55 (1.30 - 1.86)
All individuals	61,405				

4.5.1 All-cause mortality and psychiatric disorders

Figure 4-3 shows the overall Kaplan-Meier curves for all-cause mortality by prior psychiatric diagnosis in individuals given a community sentence. Individuals with prior diagnoses of substance use disorders were more likely to die during the follow-up period than were individuals without substance use disorders (Figure 4-2). The corresponding hazard ratio adjusted for age and sex was 2.64 (95% CI 2.51-2.79). This association remained significant after adjustment for other measured sociodemographic covariates, criminal history and prior self-harm (Table 4-4, Figure 4-2). In the sibling comparison model, additionally adjusted for age and sex, the hazard ratio for the association between substance use and all-cause mortality was 2.01 (95% CI 1.32-3.06). Overall, 1,531 of 5,749 all deaths were potentially attributable to substance use, which corresponds to the PAF of 26.6% (95% CI 24.5-28.8%) (Table 4-5).

Having any psychiatric disorder other than substance use was also associated with an increased risk of death (Figure 4-2). The corresponding hazard ratio adjusted for age and sex was 1.72 (95% CI 1.63-1.82). This association remained significant after adjustment for other measured sociodemographic covariates, criminal history and prior self-harm (Table 4-4, Figure 4-2). In the sibling model, adjusted for unmeasured familial confounding, the hazard ratio for the association between other psychiatric disorders and death was 1.27 (95% CI 0.84-1.91). Overall, 674 of all 3,279 deaths were potentially attributable to psychiatric disorders other than substance use, which corresponds to the PAF of 12.4% (95% CI 11.0-13.8%).

Hazard ratios for individual diagnoses other than substance use ranged from 1.47 (1.36-1.60) for anxiety disorder to 2.47 (2.33-2.61) for ADHD (Table 4-4). The associations between individual psychiatric diagnosis and mortality were partially mediated by comorbid substance use (Table 4-5).

Receiving a first-time psychiatric diagnosis during the follow-up was associated with a higher risk of death from any cause (Table 4-3). For new substance use diagnosis, the hazard ratio was 4.00 (95% CI 3.61-4.43). For other new psychiatric diagnosis, the hazard ratio was 2.82 (95% CI 2.04-2.55).

Table 4-4. Estimation of the direct effect of individual psychiatric disorders on all-cause mortality relative to measured covariates

Sociodemographic factors included being employed, receiving income support, education level, marital status. Criminal history included any prior conviction, a prior conviction for a violent crime, having a violent index offence. History of self-harm included hospitalisations and outpatient medical visits with ICD-10 codes X60-X84, Y10-Y34.

Outcome: **All-cause mortality**	Adjusted Cox regression models, HR (95% CI)			
Previous psychiatric disorder	*Model 1. Age + sex*	*Model 2. Age + sex + sociodemographic factors*	*Model 3. Age + sex + sociodemographic factors + criminal history*	*Model 4. Age + sex + sociodemographic factors + criminal history + history of self-harm*
Any psychiatric diagnosis	2.47 (2.33-2.61)	2.28 (2.15-2.41)	2.28 (2.15-2.42)	2.10 (1.98-2.23)
Any psychiatric diagnosis (other than substance use)	1.72 (1.63-1.82)	1.58 (1.50-1.68)	1.60 (1.51-1.69)	1.43 (1.35-1.52)
Schizophrenia spectrum	1.56 (1.40-1.74)	1.35 (1.20-1.51)	1.36 (1.21-1.53)	1.26 (1.12-1.41)
Bipolar	1.56 (1.29-1.90)	1.42 (1.17-1.74)	1.44 (1.18-1.75)	1.24 (1.02-1.52)
Depression	1.48 (1.36-1.60)	1.39 (1.28-1.51)	1.40 (1.29-1.52)	1.19 (1.09-1.29)
Anxiety	1.47 (1.36-1.60)	1.43 (1.31-1.55)	1.42 (1.31-1.54)	1.32 (1.22-1.44)
Personality disorder	1.75 (1.61-1.91)	1.51 (1.39-1.65)	1.50 (1.37-1.64)	1.32 (1.20-1.44)
Attention-deficit hyperactivity	2.38 (2.02-2.80)	2.11 (1.78-2.49)	2.09 (1.77-2.48)	1.80 (1.52-2.14)
Other developmental or childhood	1.82 (1.60-2.07)	1.63 (1.42-1.86)	1.63 (1.42-1.87)	1.43 (1.25-1.64)
Substance use	2.64 (2.51-2.79)	2.45 (2.32-2.59)	2.46 (2.32-2.60)	2.27 (2.14-2.41)
Alcohol use	2.20 (2.08-2.32)	2.03 (1.92-2.15)	2.03 (1.92-2.15)	1.87 (1.76-1.98)
Drug use	2.46 (2.33-2.60)	2.22 (2.09-2.35)	2.23 (2.10-2.36)	2.01 (1.89-2.14)

Table 4-5. Association between all-cause mortality and substance use comorbidity in individuals given community sentences with prior psychiatric diagnosis.

Hazard ratios were estimated by comparing individuals with psychiatric diagnoses to individuals without known psychiatric diagnoses.

All-cause mortality				
	Incidence of death Died / N of individuals with disorder (%)		Hazard ratio (95% CI)s Adjusted for age	
Previous psychiatric disorder	*With substance use*	*Without substance use*	*With substance use*	*Without substance use*
Any psychiatric (other than substance use)	3,432 / 34,918 (10%)	449 / 13,428 (3%)	2.72 (2.57-2.88)	1.31 (1.18-1.45)
Schizophrenia spectrum	259 / 2,528 (10%)	73 / 1,141 (6%)	2.78 (2.44-3.17)	1.73 (1.37-2.18)
Bipolar	80 / 985 (8%)	26 / 486 (5%)	2.93 (2.34-3.67)	1.69 (1.15-2.49)
Depression	570 / 6,495 (9%)	120 / 3,853 (3%)	2.84 (2.58-3.13)	1.20 (0.99-1.44)
Anxiety	510 / 5,419 (9%)	136 / 4,464 (3%)	3.09 (2.81-3.41)	1.15 (0.97-1.37)

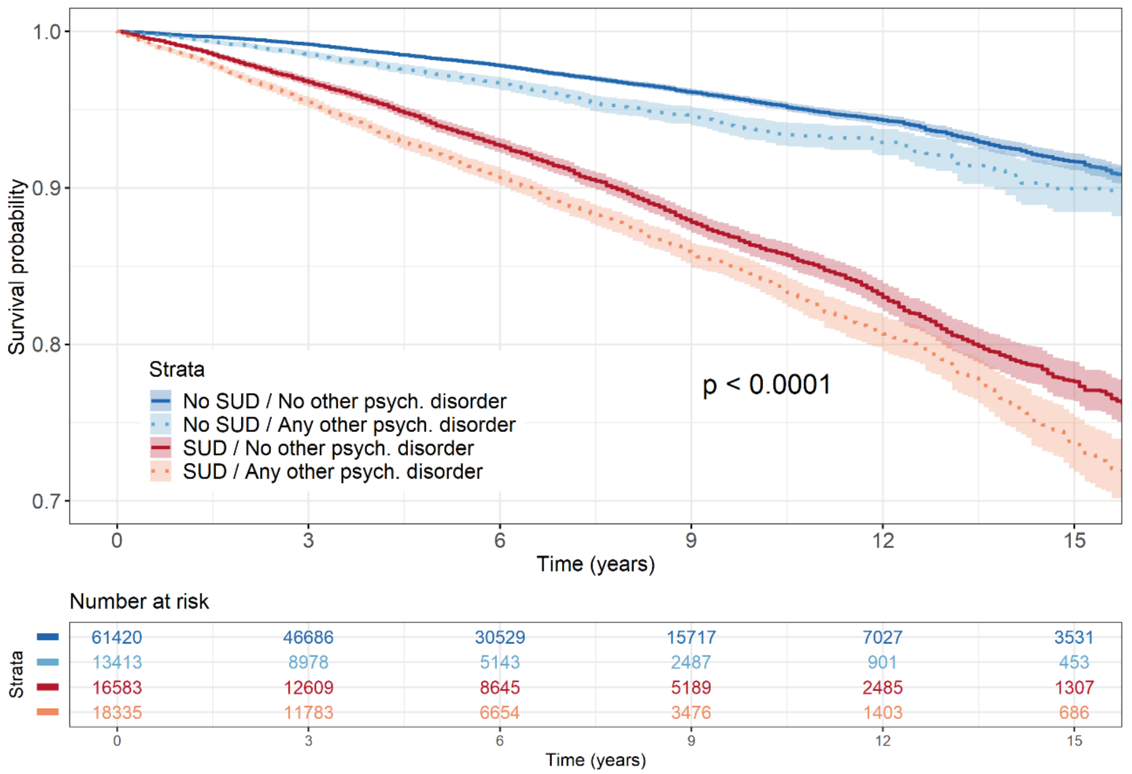


Figure 4-3. Kaplan-Meier curves for all-cause mortality in individuals given community sentences by prior diagnosis.

SUD = substance use disorder.

Table 4-6. Population attributable fraction (PAF) for all-cause mortality by substance use and other psychiatric disorder diagnoses.

The analyses were done for the maximum duration of the follow-up period (22 years) and adjusted for age and sex.

	N of deaths	N died with diagnosis	Adjusted hazard ratio (HR [95% CI])	N of deaths attributable to diagnosis	PAF (% [95% CI])
Substance use disorder					
Overall	5,749	3,432	2.64 (2.51-2.79)	1,531	26.6 (24.5-28.8)
Men	5,096	2,960	2.62 (2.47-2.77)	1,280	25.1 (22.9-27.3)
Women	653	455	2.91 (2.46-3.44)	257	39.4 (33.2-45.6)
Other psychiatric disorder					
Overall	5,749	2,164	1.72 (1.63-1.82)	712	12.4 (11.0-13.8)
Men	5,096	1,788	1.71 (1.62-1.81)	566	11.1 (9.7-12.5)
Women	653	364	1.80 (1.54-2.11)	148	22.7 (16.7-28.7)

4.5.2 External-cause mortality and psychiatric disorders

Figure 4-4 shows the overall Kaplan-Meier curves for external-cause mortality in individuals given a community sentence. Individuals with previous diagnoses of substance use disorders were more likely to die from an external cause during the follow-up than were individuals without substance use disorders (Figure 4-2). The corresponding hazard ratio adjusted for age and sex was 3.66 (95% CI 3.38-3.96). This association remained significant after adjustment for other measured sociodemographic covariates, criminal history and prior self-harm (Table 4-7, Figure 4-2). In the sibling model, adjusted for unmeasured familial confounding, the hazard ratio for the association between substance use and external-cause mortality was 2.82 (95% CI 1.60-4.99). Overall, 1,136 of 2,709 deaths from external causes were potentially attributable to substance use, which corresponds to the PAF of 42.0% (95% CI 39.2-44.7%).

Having any psychiatric disorder other than substance use was also associated with an increased risk of death from an external cause (Figure 4-2). The corresponding hazard ratio adjusted for age and sex was 2.40 (95% CI 2.22-2.59). This association remained significant after adjustment for other measured sociographic covariates, criminal history and prior self-harm (Table 4-7, Figure 4-2). In the sibling model, adjusted for unmeasured familial confounding, the hazard ratio for the association between other psychiatric disorders and death was 1.43 (95% CI 0.83-2.49). Overall, 704 of 2,709 deaths from external causes were potentially attributable to psychiatric disorders other than substance use, which corresponds to the PAF of 26.0% (95% CI 23.5-28.5%) (Table 4-9).

Table 4-7. Estimation of the direct effect of individual psychiatric disorders on external-cause mortality relative to measured covariates.

Sociodemographic factors included being employed, receiving income support, education level, marital status. Criminal history included any prior conviction, a prior conviction for a violent crime, having a violent index offence. History of self-harm included hospitalisations and outpatient medical visits with ICD-10 codes X60-X84, Y10-Y34.

Outcome: **External-cause mortality**	**Adjusted Cox regression models, HR (95% CI)**			
Previous psychiatric disorder	*Model 1. Age + sex*	*Model 2. Age + sex + sociodemographic factors*	*Model 3. Age + sex + sociodemographic factors + criminal history*	*Model 4. Age + sex + sociodemographic factors + criminal history + history of self-harm*
Any psychiatric diagnosis	3.42 (3.15-3.72)	3.17 (2.90-3.46)	3.11 (2.85-3.40)	2.76 (2.52-3.03)
Any psychiatric diagnosis (other than substance use)	2.40 (2.22-2.59)	2.22 (2.05-2.41)	2.24 (2.07-2.43)	1.93 (1.78-2.10)
Schizophrenia spectrum	1.96 (1.67-2.30)	1.65 (1.40-1.95)	1.68 (1.43-1.99)	1.47 (1.25-1.74)
Bipolar	2.34 (1.80-3.05)	2.15 (1.64-2.82)	2.21 (1.68-2.89)	1.75 (1.33-2.30)
Depression	2.09 (1.87-2.34)	1.97 (1.76-2.21)	2.01 (1.79-2.25)	1.60 (1.42-1.80)
Anxiety	1.81 (1.62-2.03)	1.73 (1.54-1.94)	1.72 (1.53-1.93)	1.52 (1.35-1.71)
Personality disorder	2.34 (2.08-2.64)	2.00 (1.77-2.26)	1.97 (1.74-2.24)	1.61 (1.42-1.83)
Attention-deficit hyperactivity	2.32 (1.93-2.80)	2.12 (1.75-2.57)	2.10 (1.73-2.55)	1.70 (1.40-2.07)
Other developmental or childhood	1.83 (1.56-2.14)	1.64 (1.39-1.94)	1.64 (1.38-1.94)	1.40 (1.18-1.66)
Substance use	3.66 (3.38-3.96)	3.36 (3.09-3.66)	3.27 (3.00-3.57)	2.89 (2.65-3.16)
Alcohol use	2.52 (2.32-2.73)	2.28 (2.09-2.48)	2.26 (2.08-2.47)	1.96 (1.79-2.14)
Drug use	3.65 (3.38-3.93)	3.33 (3.06-3.62)	3.24 (2.97-3.52)	2.82 (2.59-3.08)

Table 4-8. Association between external-cause mortality and substance use comorbidity in individuals given community sentences with prior psychiatric diagnosis.

Hazard ratios were estimated by comparing individuals with psychiatric diagnoses to individuals without known psychiatric diagnoses.

External-cause mortality				
	Incidence of death Died / N of individuals with disorder (%)		Hazard ratio (95% CI)s Adjusted for age	
Previous psychiatric disorder	*With substance use*	*Without substance use*	*With substance use*	*Without substance use*
Any psychiatric (other than substance use)	1,656 / 34,918 (5%)	239 / 13,428 (2%)	3.96 (3.63-4.32)	1.69 (1.46-1.96)
Schizophrenia spectrum	135 / 2,528 (5%)	29 / 1,141 (3%)	4.68 (3.88-5.64)	2.20 (1.52-3.19)
Bipolar	39 / 985 (4%)	18 / 486 (4%)	5.62 (4.04-7.80)	4.16 (2.60-6.65)
Depression	301 / 6,495 (5%)	68 / 3,853 (2%)	5.27 (4.58-6.06)	1.98 (1.54-2.54)
Anxiety	287 / 5,419 (5%)	67 / 4,464 (2%)	5.12 (4.46-5.88)	1.40 (1.09-1.80)

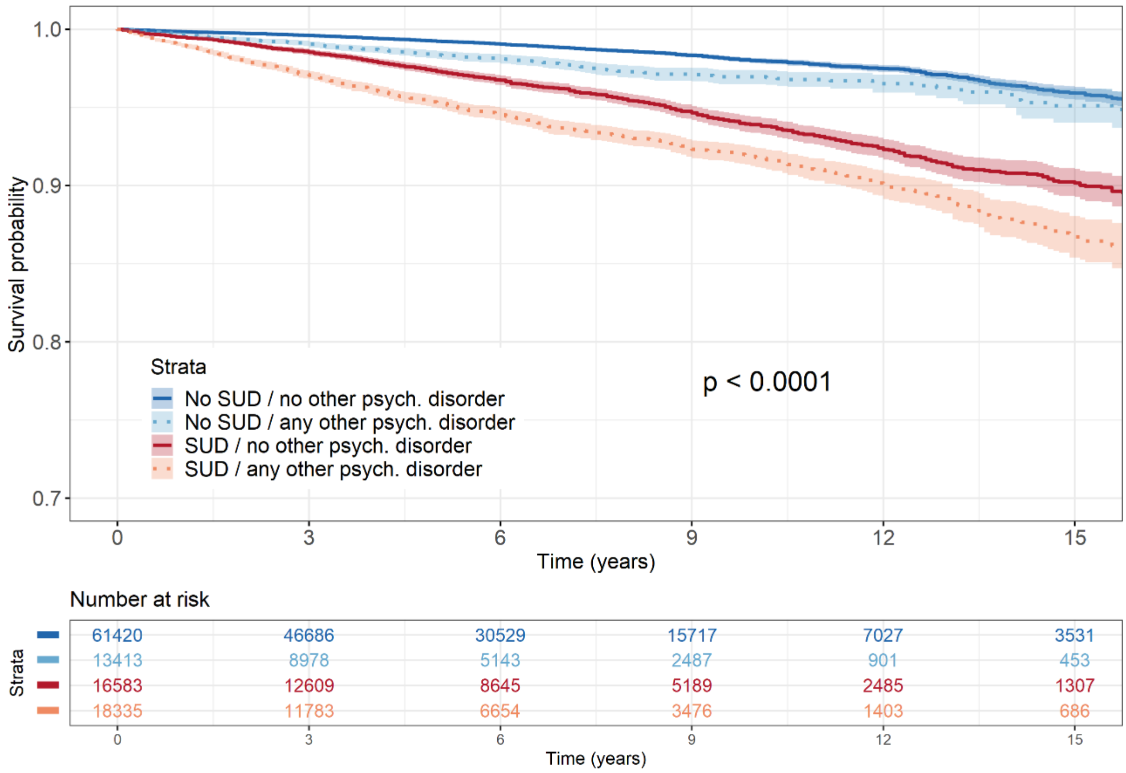


Figure 4-4. Kaplan-Meier curves for external-cause mortality in individuals given community sentences by prior diagnosis.

SUD = substance use disorder.

Table 4-9. Population attributable fraction (PAF) for external-cause mortality by substance use and other psychiatric disorder diagnoses.

The analyses were done for the maximum duration of the follow-up period (22 years) and adjusted for age and sex.

	N of deaths	N died with diagnosis	Adjusted hazard ratio (HR [95% CI])	N of deaths attributable to diagnosis	PAF (% [95% CI])
Substance use disorder					
Overall	2,709	1,656	3.66 (3.38-3.96)	1,136	42.0 (39.2-44.7)
Men	2,396	1,410	3.61 (3.32-3.93)	961	40.1 (37.2-43.0)
Women	313	239	4.33 (3.33-5.63)	181	57.8 (49.6-66.1)
Other psychiatric disorder					
Overall	2,709	1,161	2.40 (2.22-2.59)	704	26.0 (23.5-28.5)
Men	2,396	943	2.32 (2.14-2.52)	559	23.4 (20.8-25.9)
Women	313	212	3.13 (2.46-3.98)	151	48.1 (39.6-56.6)

Hazard ratios for individual diagnoses other than substance use ranged from 1.81 (95% CI 1.62-2.03) for anxiety disorder to 2.34 (95% CI 1.80-3.05) for bipolar disorder and 2.34 (95% CI 2.08-2.64) for personality disorder (Table 4-7). Comorbid substance use partially mediated the associations between individual psychiatric diagnosis and death (Table 4-8).

Receiving a first-time psychiatric diagnosis during the follow-up was associated with a higher risk of death from an external cause (Table 4-3). For new substance use diagnosis, the hazard ratio was 3.88 (95% CI 3.34-4.51). For other new psychiatric diagnoses, the hazard ratio was 1.54 (95% CI 1.29-1.83).

4.6 Discussion

I examined the association between psychiatric disorders and mortality in a Swedish nationwide population-based study of 109,751 individuals given community sentences from 1991 to 2013. I followed individuals from the day their community sentence began until their death, emigration, or the end of the follow-up period. The study has three principal findings.

First, substance use and other psychiatric disorders were significantly associated with increased all-cause and external-cause mortality in individuals given community sentences. The association remained significant even after controlling for measured sociodemographic factors, prior criminal history, and the history of self-harm. The association attenuated but remained significant in sibling comparison, which adjusts for genetic and environmental factors shared between siblings discordant by a given diagnosis. Substantial attenuation of the effect in the sibling model suggests either that certain psychiatric disorders have common familial causes with mortality or that the effect of the psychiatric disorders is mediated through factors shared in the family (Sjölander & Zetterqvist, 2017). The comorbid substance use only partially mediated the association between other psychiatric disorders and mortality. The magnitude of the association between substance use, other psychiatric disorders and mortality were comparable to estimates from other studies in community sentenced populations (Sodhi-Berry et al., 2015) and in released prisoners (Z. Chang, Lichtenstein, et al., 2015).

As with the aetiology of criminal behaviour, there is likely no single aetiological pathway that connects psychiatric disorders and mortality. Prior studies have shown that the association between psychiatric disorders and physical health can be realised through different combinations of psychosocial and biological mechanisms. Identified psychosocial mediators of general psychiatric psychopathology can include antisocial lifestyle and attitudes, deficits of executive functions, and poor emotional regulation (Brennan et al., 2017; Simpson et al., 2015). Biological factors can include acute or cumulative chronic damage to bodily organs and systems or interference with their normal functions. Substance use was shown to cause serious neurological, gastrointestinal (including liver), and dermatological problems (Keaney et al., 2011). The association between other psychiatric disorders and mortality can be mediated through non-specific stress response and increased inflammatory burden (Rasmussen et al., 2020).

Second, substance use and other psychiatric disorders had a stronger association with external-cause mortality than with all-cause mortality. The association between psychiatric diagnoses and potentially preventable deaths suggest the importance of mental health as an intervention target in individuals given community sentences. Chapter 3 highlighted the importance of mental health risk factors as potential targets for interventions that aim to lower criminal recidivism. Given the shared interventional targets between recidivism and mortality outcomes, it may be practical to provide interventions for community sentenced individuals within an integrated interventional framework (Liu et al., 2017). Identifying barriers that prevent sentenced individuals with mental health disorders from access to interventions should also be addressed. The barriers can include lack of medical insurance, low literacy or difficulty understanding a local language, low help-seeking behaviour, low health literacy (inability to recognise one's symptoms as non-normative), mistrust of the healthcare system, fear of stigmatisation, and superficial provision of services by healthcare providers (Howerton et al., 2007; Owens et al., 2011; Sirdifield & Brooker, 2020).

Third, suicide was the leading cause of death in our cohort. Most suicides occurred around the age of 25, which contrasts with the trend in the general population. During the years of study, the highest average suicide rates per age group were 38.0 per 100,000 for men aged 65+ and 16.4 per 100,000 for women aged 45-64 (National Centre for Suicide Research and Prevention, 2021). Compared to these benchmark rates, the suicide rates in community sentenced individuals were at least 4.4 times higher for men and 11 times higher for women. In our cohort, the number of suicides peaked around the age of 25-30. These findings underscore the need to pay additional attention to suicide risk screening and the employment of preventive measures during post-sentence supervision. In addition to higher access to means of suicide, community supervision itself can serve as a trigger for self-harm. Community supervision can be a significant source of stress associated with the loss of control over one's mental state and situation

(Mackenzie et al., 2017). While being less harsh than prison, community supervision lacks the mitigating effect of the controlled environment with the constant observation by others. Moreover, the deaths from external causes, including suicide, were not limited to the immediate post-sentence supervision period. This finding emphasizes the need for continuity of care after the sentence.

4.7 Strength and limitations

To my knowledge, this is the first study that examined the association between psychiatric disorder and post-sentence mortality in individuals given community sentences using sibling designs and additionally estimating the effect of new diagnoses during the follow-up. The estimates of the population effect of substance use and other psychiatric disorders on post-release mortality are also novel.

This study did not examine the effects of future sentences on the trajectory of an individual. It is possible that individuals with substance use and other psychiatric diagnoses, having a higher risk of criminal recidivism, are more likely to go to prison at some point after being given community sentences. Controlling for future effects of imprisonment would allow for a stronger causal interpretation of the specific effects of community sentences on an individual's mortality risk. It was demonstrated that multiple prison sentences are associated with higher mortality risk (Graham et al., 2015). Clear separation between the effects of custodial and non-custodial measures on mortality is a potential direction for future research.

4.8 Conclusion

Individuals with psychiatric disorders are at a higher risk of premature death compared to other sentenced individuals. The increased risk of premature death extended beyond the post-released supervision period. Most preventable deaths occurred in sentenced individuals with psychiatric disorders in the study cohort, especially young individuals. The leading cause of death was suicide. This study emphasises the need to focus on screening and prevention of suicide and external cause mortality in individuals with psychiatric disorders under post-release supervision. Community sentences should not have the reduction of criminal recidivism risk as their only goal. They also represent an opportunity to provide sentenced individuals with access to evidence-based integrated interventions to improve general health outcomes and prevent premature death.

Chapter 5. Dynamic prediction of reoffending in individuals given community sentences

5.1 Abstract

I used pre-specified criminal, sociodemographic, and clinical risk factors to develop a dynamic prediction model for criminal recidivism in individuals under community supervision. I used the dynamic prediction by landmarking approach that adjusts the predictive model for survival data during the follow-up.

The resulting prediction model takes into account adverse events that might occur during the community supervision (triggers for violence), changes in a supervised individual's circumstances, and desistance from crime. The model estimates the probability of general and violent reoffending within two years and can be used for ongoing risk monitoring. The model was deployed as an online dynamic risk assessment tool OxMore with good calibration and discrimination performance (c-index = 0.74 for violent reoffending, c-index = 0.69 for general reoffending).

As an important secondary outcome, the study demonstrated that actuarial recidivism risk assessment tools, which have not been developed as risk monitoring instruments but are used as such, will systematically overestimate the risk of recidivism over time.

5.2 Introduction

Chapters 2 and 4 demonstrated that psychiatric disorders are important risk factors for criminal recidivism and mortality in individuals given community sentences. Improved access to substance use treatment and general mental health treatment could substantially decrease the number of new offences and deaths in this population. However, access to such programmes is often limited, and their coverage depends on extraneous variables, such as the availability of public funding and trained personnel.

Risk assessment tools are often used to prioritize the limited resources of supervising agencies and healthcare providers. There are two main classes of risk assessment tools: structured clinical judgement instruments and actuarial instruments. Structured clinical judgement instruments incorporate empirically based risk factors that guide the assessor's subjective prediction of an individual's risk. The risk factors include both static (unchangeable) and dynamic (potentially changeable) characteristics of an individual or their environment. Actuarial risk assessment tools 'estimate the likelihood of misconduct by assigning numerical values to risk factors associated with offending' (Singh, Grann, & Fazel, 2010). Some actuarial recidivism prediction instruments rely exclusively on static risk factors (e.g., VRAG, Risk Matrix 2000, OGRS, Static-99), others also incorporate dynamic risk factors (e.g., LSI-R, VRS) (Singh, Desmarais, et al., 2014; Yang, Wong, & Coid, 2010). While the predictive validity of structured clinical judgement instruments and actuarial instruments is comparable, actuarial instruments are often easier to administer and require less training to use. The focus of the current study is the development of the actuarial risk assessment instrument that incorporates dynamic risk factors and is easy to administer.

More resources are allocated to individuals who present as high-risk, whereas low-risk individuals get less supervision and less intensive interventions. This principle is referred to as a ‘risk principle’ (Bonta & Andrews, 2007). However, every individual undergoing supervision has a unique set of static and dynamic risk factors. The interventions should aim to change the modifiable risk factors and mitigate the effect of static risk factors. This principle is referred to as a ‘need principle.’ Also, the interventions should be evidence-based and tailored to an individual's learning style to be effective. This principle is referred to as a ‘responsivity principle.’ The risk, need, and responsivity principles comprise the Risk-Need-Responsivity model (RNR), a standard approach to risk management of sentenced individuals (Bonta & Andrews, 2007; Maruna & Mann, 2019).

However, some researchers and practitioners argue that RNR overemphasizes the risk factors and does not focus enough on factors that prevent criminal behaviour (Maruna & Mann, 2019; Wong & Horan, 2021). They argue for the shift towards facilitating desistance from crime in the risk management of supervised individuals. The shift towards desistance may also be potentially useful for improving the accuracy and validity of recidivism risk assessment. It may also help to decrease the detrimental effect of unnecessary harsh sanctions on individuals deemed high risk (Fazel et al., 2012; Viglione & Taxman, 2018).

One way of incorporating desistance into risk assessment is a more detailed evaluation of protective factors. The instruments assessing protective factors for criminal recidivism have been developed (de Vogel et al., 2011), although their predictive validity remains moderate. Another possible way of taking desistance into account is the development of dynamic risk monitoring tools that incorporate offence-free time as a predictor of recidivism. As prior studies indicate, after five-ten years without offending, an individual with a criminal record has the same risk of committing a crime as an average person from the general population (Hanson et al., 2018). This effect of criminal recidivism risk decay over time could potentially provide useful information about recidivism risk that simple changes in dynamic risk factors cannot provide. In the present study, I developed the dynamic risk assessment tool for criminal recidivism risk monitoring that accounts for offence-free time.

Another line of research that can potentially improve the predictive validity of risk assessment instruments is the analysis of the acute risk factors. These factors can increase the likelihood of committing a new offence, especially a violent offence, although this effect can be short-lived. Prior studies investigated the effect of adverse events in individuals with psychotic disorders and the general population (Sariaslan et al., 2016). They demonstrated that exposure to violence, substance intoxication, unintentional injury, traumatic brain injury (TBI), and self-harm were associated with an increased risk of violence in the weeks following an adverse event in all individuals. Another study in the general population demonstrated that psychiatric hospitalization was associated with an increased risk of violent offending and other adverse outcomes, including suicide and non-fatal self-harm (Walter et al., 2019). The acute increase in risk was noted during the first months after discharge from a hospital, and the residual effect was constant throughout the 10-year follow-up period. The research into the association between adverse events and criminal recidivism in the community sentenced population is limited. However, a recent study identified an increased risk of rearrest in individuals given community sentences with prior traumatic brain injury (Norman et al., 2021). The growing body of research also highlights the association between unstable housing situation and reoffending in custodial and non-custodial populations (Clark, 2014; Jacobs & Gottlieb, 2020).

These adverse events with transient effects have been referred to as ‘triggers for violence’ (Sariaslan et al., 2016). In this thesis, I use the term ‘triggers’ to emphasize that their effect may not be limited to violent offences. Triggers represent understudied but potentially promising covariates to include in the risk assessment instruments. However, the triggers’ effect on general criminality and their predictive validity for criminal recidivism have not been investigated. Thus, their inclusion in predictive models has to depend on their association with an outcome of interest, and a pre-specified inclusion criterion has to be applied to filter out the spurious associations (Heinze et al., 2018).

Widely used actuarial instruments were not developed to capture the effect of triggers and have a low temporal resolution. The term ‘temporal resolution’ is not commonly used to describe recidivism risk assessment instruments. In general, temporal resolution refers to the density of time points over which the data are collected (Stępniak et al., 2019; Théau, 2008). A high temporal resolution allows the collection of high-density data that are more likely to contain rare but potentially critical events. Thus, the term can be applied to dynamic recidivism risk assessment with the same meaning. A risk assessment instrument with high temporal resolution captures quick and transient changes in the levels of risk factors and re-estimates the risk score accordingly. A risk assessment instrument with a low temporal resolution is not sensitive to these transient changes and cannot account for them in their estimates. In the present study, I developed the dynamic risk assessment tool for criminal recidivism risk monitoring with a high temporal resolution that accounts for the transient effect of triggers. Building on past research (Sariaslan et al., 2016; Walter et al., 2019), I investigated such triggers as being a victim of a violent assault (exposure to violence), substance intoxication, unintentional injury, traumatic brain injury, self-harm episode, and psychiatric hospitalisation.

This study aims to develop and validate a dynamic risk assessment model for criminal recidivism in individuals given community sentences. I used data from a large population-based cohort of individuals given community sentences in Sweden. I examined the set of empirically supported and theoretically based risk factors for general and violent recidivism (Fazel et al., 2016; Fazel, Wolf, Vazquez-Montes, et al., 2019; Wolf et al., 2018). The developed dynamic risk prediction model also accounts for offence-free time, the effect of triggers, and has a high temporal resolution.

5.3 Approach

5.3.1 Dynamic prediction with landmarking

The 'standard' Cox model that I used in Chapter 4 and 5 have been extensively employed for the development of the risk assessment instruments (Fazel et al., 2016; Wolf et al., 2018). However, it is not suitable for dynamic risk assessment as it cannot account for the change in covariate values. Simple time-dependent Cox regression is also not suitable for prediction modelling (Putter & van Houwelingen, 2017). Some machine learning approaches, including neural networks and survival forests, have been used to predict survival outcomes and can be used for the development of dynamic prediction models (Hazewinkel, 2018). However, they do not provide easily interpretable coefficients, which is crucial for risk assessment instruments in high-risk situations (Rudin, 2019). To obtain easily interpretable estimates, I used the dynamic model with sliding window landmarking as described in van Houwelingen & Putter (2011).

Landmarking approach allows the pooling of estimates from cross-sectional datasets over time. The density of landmark points represents the temporal resolution of the developed model. Ideally, one wants to maximize the number of landmark points to capture all possible influences of dynamic risk factors on the outcome of interest. However, it is often not computationally feasible as adding only one landmark point increases the size of a dataset, sometimes doubling it. In the current study, the dataset with 43,192 observations was transformed into a landmark superset of 1,126,328 observations after being divided by 37 landmark points. Therefore, the choice of the number and frequency of landmarks is a trade-off between precision and feasibility. Ultimately this choice should be informed by the characteristics of the process one aims to predict and the intended use of the resulting model.

The standard time-event data structure is represented in Table 5-1. Figure 5-1 depicts the same data graphically. These data were transformed into a landmark superset in a wide format (Table 5-2, Figure 5-2).

Each row in a landmark superset is a separate observation of an individual within a given period of LM + w, where w is an observational horizon of 24 months. If an individual does not reoffend and is not censored prior to the time LM + w, they are considered outcome-free at the end of the observation. In Table 5-2, column *LM* contains the landmark times. Each new observation starts at the landmark time *LM* and lasts until *time.stop*. The length of the observational period is represented by *time.stop.rel* (for relative), which is the difference between *time.stop* and *LM* times.

The covariates were also transformed in this example dataset. *Employed* was changed to reflect employment status at time 0, and thus cannot change throughout the follow-up period. *Substance use* was changed to reflect a substance use disorder diagnosis received before a given landmark, i.e., prior to the time of a risk assessment. *Current substance use* variable can therefore change at any point of the follow-up period. However, only in one direction, from zero to one.

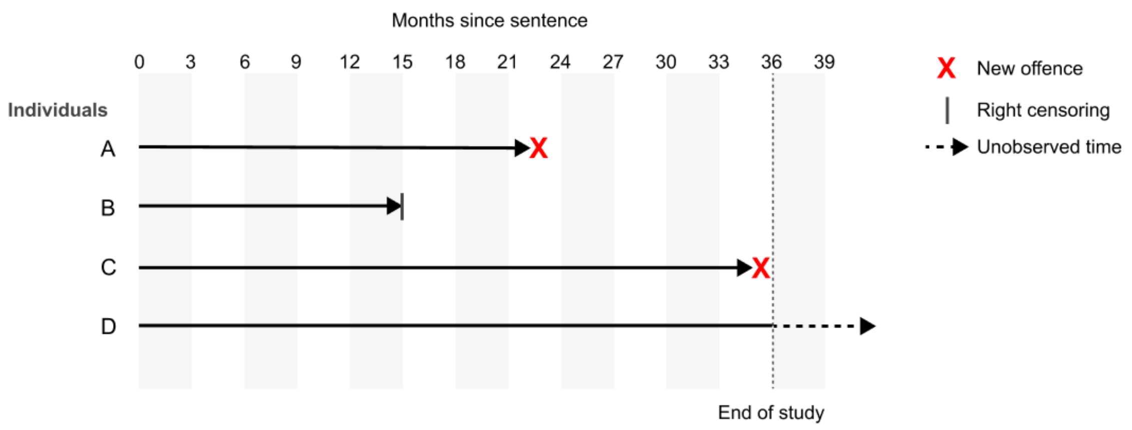


Figure 5-1. Graphical representation of the standard time-to-event data used for recidivism prediction. Letters A, B, C, D represent separate individuals. The observations correspond to the table below

Table 5-1. Hypothetical reoffending dataset in a regular time-to-event format

id	time.stop	Reoffended	Sex	Age	Employed	Substance use
A	23	1	1	25	1	0
B	15	0	0	35	0	0
C	35	1	0	45	1	0
D	36	0	0	23	1	1

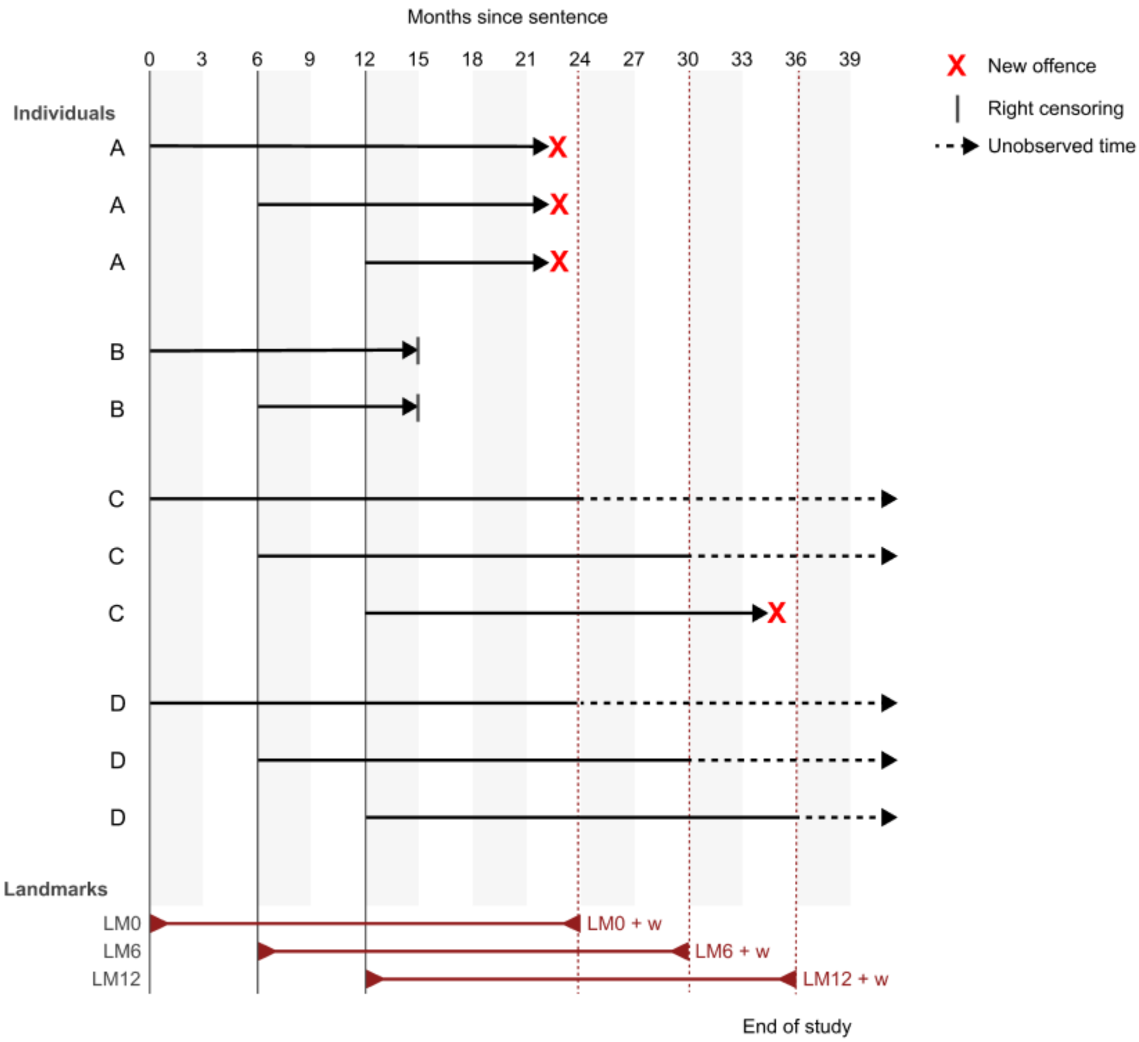


Figure 5-2. Graphical representation of time-to-event data organised for recidivism prediction using landmark approach. Letters A, B, C, D represent separate individuals. The observations correspond to the table below. The observational period starts at a given landmark and continues for the duration of the sliding time window of w = 24 months

Table 5-2. Hypothetical landmark superset with recidivism data

id	LM	time.stop.rel	time.stop	Reoffended	Sex	Current age	Employed at baseline	Current substance use
A	0	23	23	1	1	25	1	0
A	6	17	23	1	1	26	1	1
A	12	11	23	1	1	26	1	1
B	0	15	15	0	0	35	0	0
B	6	9	15	0	0	35	0	0
C	0	24	24	0	1	41	1	0
C	6	24	30	0	1	41	1	0
C	12	23	35	1	1	42	1	0
D	0	24	36	0	0	23	1	1
D	6	24	36	0	0	23	1	1
D	12	24	36	0	0	23	1	1

5.3.2 Choosing landmark point and the follow-up time

On the one hand, the landmark times have to be chosen so that they capture the process of interest. In the case of triggers, which are episodic events with a transitory effect (Sariaslan et al., 2016), the period between landmarks has to be as small as possible. On the other hand, the computation with many landmark datasets stacked into one large superset quickly becomes unfeasible. The solution is to set the number of landmarks that would be meaningful in the practical application of the predictive landmark model. The intended application of the developed model is to monitor the post-sentence reoffending risk of an individual given a community sentence. My understanding of the Swedish Prison and Probation Service processes informed the model development, as I used Swedish data provided by them. However, the way community sentences are carried on in Sweden is not unique, and many other countries employ similar methods.

A typical community sentence (probation) in Sweden lasts three years with supervision during the first year (Swedish Prison and Probation Service, 2021). The frequency of the meetings between probation officers and their clients varies depending on a client's estimated reoffending risk. The average period between sessions for one client is around 2-4 weeks (information obtained during a personal discussion with Swedish Prison and Probation Service employees). Using these numbers as reference points, I chose the set of 37 equally spaced landmarks. The first landmark was set at the time of receiving a sentence (LM = 0), with the other 36 covering the follow-up period of 3 years. The period between each landmark is 30.25 days. The 2-year reconviction was chosen as a primary outcome, as it is the most common measure used for criminal recidivism globally (as demonstrated in Chapter 1). It was operationalised as 2-year violent and general reoffending.

5.3.3 Presentation of findings and risk threshold

Some authors argue that using risk thresholds to stratify the level of risk is problematic, and the probabilistic estimates have to be used as outputs of risk assessment instruments (Wynants et al., 2019). Using the threshold value that optimises the number of correctly classified individuals may seem like a good idea at first glance. However, as it assumes the same cost for false-negative and false-positive results, it may not be appropriate for the practice setting. Practitioners, especially in the forensic psychiatric setting, often want to lower the false-negative rate more than the false-positive rate, as the latter would lead to less supervision allocated to potentially high-risk individuals.

In a forensic psychiatric setting, the use of risk categories can lead to asymmetric allocation of resources, where low-risk individuals get fewer resources than high-risk individuals (Ryan et al., 2010). It might be problematic as low-risk individuals are often more likely to benefit from additional treatment resources than high-risk individuals. Assigning the pre-determined probabilities to future risk of violence without providing local base rates for violent acts could be problematic as well, especially for the individuals classified as high risk (Singh et al., 2014). The base rates estimated on the population on which the tool is used can serve as reference points for decision-making. I added the base rate estimates for a given landmark estimated from Kaplan-Meier estimates for the derivation sample.

I argue that the predictive models have to be used as an instrument of risk communication and as a part of an integrated risk management approach. The dynamic tools are more convenient for this purpose because they are sensitive to changing conditions and can be used as an ongoing measure on regular check-ups or visits.

5.3.4 Generasibility

The predictive model is derived from the data and thus depends on how the data is obtained, coded, and stored. Data collection can vary drastically from country to country and sometimes from province to province within one country (as demonstrated in Chapter 1). This high data source variability can limit the utility of the predictive models (Sáez, Romero, Conejero, & García-Gómez, 2021). When data routinely obtained by state agencies are used to develop a predictive model, the resulting model may not be easily generalizable to another state with different procedures in place. Subsequently, to use the predictive model in a new setting, the model often has to be re-trained and re-calibrated (Fazel, Wolf, Vazquez-Montes, et al., 2019). The model might contain variables that are highly specific to its country of origin and not useful in a new setting. I aimed to avoid the inclusion of the highly specific variables to improve the model's potential generalisability. I did not include such factors as immigrant status and neighbourhood deprivation despite their use in prior instruments developed by our research group. Their operationalization is very specific to Sweden and may not translate well to other countries.

5.4 Methods

The current study is a model development and validation study that uses the retrospective cohort design. The study follows the TRIPOD checklist guidelines for model development and validation (Appendix E1).

This study was approved by the Regional Ethics Committee at the Karolinska Institutet (Stockholm, Sweden).

5.4.1 Data extraction

I used data from the same population registers as in Chapters 3 and 4. I extracted data for individuals given community sentences from 2007 to 2013. This period allows the 2-year outcome estimates over a 3-year supervision period while only using the most recent available data. This was done to ensure practically meaningful model development as the trial period of 3 years is a standard for individuals given community sentences in Sweden, and 2-year reoffending is the most commonly reported criminal recidivism outcome. Consequently, five years was the maximum length of the follow-up in the cohort.

The selection procedure was identical to the one described in Chapter 3. Community sentences included conditional sentences with community service, probation with community service, and probation with contracted treatment. I only selected individuals whose sentences came into legal force and were not appealed or dismissed. For each individual, I used the date when a community sentence came into force as the start of the follow-up period. If an individual received multiple community sentences, the index sentence was selected at random. For each individual, the outcome and covariate data were extracted for the period of five years since their community sentence came into legal force. I excluded individuals who were born before 1958 because these individuals would not have a continuous criminal record on the National Crime Register. I also excluded individuals who committed a crime before the start of the follow-up period but were not sentenced for it by that time – this is referred to as a pseudo-reconviction. The inclusion of such pseudo-reconvictions could have resulted in an overestimation of recidivism risk.

These data comprised the total cohort that was split into the derivation sample, which was used for derivation and internal validation of the prediction model, and the external validation sample (the holdout dataset).

5.4.2 Data splitting

The data was divided into a derivation set (~80% of data) and an external validation set (~20% of data). The derivation set was selected from the total cohort based on the residential geographical location of the individual at the time of a sentence. Regions were primarily based on the counties of Sweden, derived from the first two digits of the Swedish Small Area Market Statistics (SAMS) code. Exceptions were the municipalities of Gothenburg and Malmö, which are separated from their respective counties, and Stockholm municipality, which is separated from its county and sub-divided into northern and southern parts by identifying each SAMS area with the historical province in which it is located. Regions were allocated to four groups, which served as proxy measures of urban/rural status: the four urban areas (Group I); the three counties in which the urban areas are located (Group II); four counties with a low population (Group III); and all other counties (Group IV). The external validation set was selected randomly, with equal probability, choosing one region from each of the first three groups and selecting sequentially from the fourth group. This approach was implemented before during the development of

OxMIV to maximise the number of outcomes event in the external validation sample (Fazel, Wolf, Larsson, et al., 2019). It will also increase the representativeness of the external validation sample.

5.4.3 Modelling process

The prediction modelling was based on Cox proportional hazards regression with sliding window landmarks (van Houwelingen & Putter, 2011), adjusting for covariates. The separate models were developed for two outcomes:

i. probability of committing a violent offence within 2 years from a given time point,
ii. probability of committing any offence within 2 years from a given time point.

I used 37 landmarks corresponding to the start of a sentence (landmark 0) and each month of the 3-year follow-up period (landmarks 1-36). Each landmark thus represented the point of risk re-evaluation. The separate datasets were created for each landmark. Each dataset contained only individuals at risk at a given landmark's time and corresponding values of covariates. All 37 landmark datasets were combined into a single landmark superset.

On this superset, I fitted a Cox regression model stratified by landmark. I used a simple landmark model, which assumes that the effect of the covariates is the same across landmarks. However, the baseline hazard is allowed to vary across landmarks, reflecting the effect of event-free survival (i.e., offence-free time).

The individuals were followed up until the occurrence of an outcome event, the end of the available follow-up period, or a censoring event. The censoring events were death, permanent emigration out of Sweden, and, in the case of violent reoffending, imprisonment for a non-violent crime.

5.4.4 Assumptions of the models

I checked the proportionality assumption over the whole observational period and over different time windows by assessing Schoenfeld residuals' plots in univariate analyses. The assumption was satisfied for all included covariates. In general, the landmark model can mediate the effect of the non-proportionality of covariates' effects by accounting for time effects (van Houwelingen & Putter, 2011). Martingale residual plots with LOESS lines were used to check the linearity assumption for 'Current age' and the covariates representing time since an event (Karlsson, 2016). 'Current age' variable satisfied linear assumption, whereas time-since-event covariates did not. I changed the functional form of time-since-event covariates to resolve this issue and improve the interpretability of the model (see 'Variable specification' below for details).

I assumed complete information about clinical covariates in our dataset. The database records are likely underreporting the true prevalence of the clinical covariates. This is one of the limitations of the study design.

5.4.5 Variable specification

The variable selection process for the current model followed the general approach implemented during the development of the OxRec (Fazel et al., 2016), FoVoX (Wolf et al., 2018), and OxMIV (Fazel, Wolf, Larsson, Mallett, & Fanshawe, 2019) prediction tools.

All variables were considered in four groups. The variable operationalisation and grouping were pre-specified in the model derivation protocol (Appendix E2). Group 1 consisted of covariates measured at the time of a sentence (at baseline). These included sex and prior criminal history. Group 2 consisted of covariates that were measured at baseline but could change value during the follow-up, reflecting the current status of an individual at landmark time. They included current age, sociodemographic variables assessed within a year prior to a given landmark, and psychiatric diagnoses prior to a given landmark. For demographic factors, the values for the first three months were forward propagated from the baseline. See 'Missing data' for rationale. Group 1 and Group 2 covariates were strongly suspected on the basis of previous research to be associated with the two outcome measures. These covariates were included in the model by default.

Group 3 consisted of covariates measured at baseline that were likely to show an association with the outcomes. These included several sociodemographic measurements and a history of self-harm. Group 4 consisted of covariates measured only during the follow-up period that were likely to show an association with the outcomes. These covariates represented adverse events that can occur during the follow-up period. Group 4 included triggers for violence (Sariaslan et al., 2016) and psychiatric hospitalisation, as a proximal measure of psychiatric symptomatology. The inclusion of Group 3 and 4 covariates was not required to achieve face validity. These covariates were included if they were significantly associated with the outcomes. To ensure a straightforward interpretation of the model's coefficients, interactions between covariates were not considered.

The time-varying effect of each trigger was entered in the model as three separate binary variables, representing three hypothetical components of a trigger’s effect. ‘Acute effect’ (risk surge or incapacitation) was coded as 1 if a trigger event occurred within a week before a given landmark, otherwise coded as 0. ‘Short-term effect’ was coded as 1 if a trigger event occurred within a month before a given landmark, otherwise coded as 0. Residual effect was coded as 1 if a trigger event occurred at any time from the start of the follow-up until a given landmark, otherwise coded as 0. The covariates for acute and short-term effects could be interpreted as modifiers of the residual trigger effect during their corresponding time windows.

5.4.6 Variable selection

All variables from Groups 1 and 2 were included in the final model by default. To further select predictors from Groups 3 and 4, I implemented automatic backwards elimination on combined estimates from all imputed datasets (Wood, White, & Royston, 2008). A p-value of 0.157 was used as a threshold for the exclusion of a variable as suggested by Heinze and Dunkler (Heinze & Dunkler, 2017).

This strategy recognizes that the final model must demonstrate face validity whilst simultaneously allowing the inclusion of additional risk factors if they show an association with outcome variables. The variables are considered in four groups in this way to recognise that a parsimonious model is preferable (i.e., easier to use in practice), provided that it has an acceptable predictive ability.

5.4.7 Missing data

The proportion of missing data for included covariates ranged from 0.1% to 3.2% (Appendix E3). I imputed the values of the sociodemographic records at the time of a sentence (at baseline) and during the follow-up. In case of missing records for sociodemographic factors during the follow-up, I used values at baseline for the first three months of the follow-up period. Three months was the median time from the start of the follow-up to the subsequent measurement. For education level, I forward propagated the last measurement until the next new available measurement without time restrictions. I imputed all other missing sociodemographic records using an Expected-Maximisation algorithm implemented in *Amelia* package for R (Honaker et al., 2011). The landmark time was used as a cross-sectional time-series indicator. All measured covariates and the outcome variables were used as predictors for missing data points (Sterne et al., 2009).

Clinical covariates, including triggers, were measured in an ongoing manner. The complete information about clinical covariates in the dataset was assumed; thus, no missing value imputation was required for them.

Table 5-3. The rule for splitting the full dataset into derivation and external validation samples

Regions were selected at random, with equal probability, as follows: one region from Group I, one region from Group II (under the constraint that no more than one region in Groups I and II from the same county can be selected), one region from Group III, sequentially select from Group IV until the number of individuals in the external validation sample is equal to or exceeds 20% of the entire dataset.

Group I Major urban centres	Group II Counties with major urban centres removed	Group III Counties with small population	Group IV Counties with medium population
1 Stockholm City North	1 Stockholm County Other	7 Kronoberg	3 Uppsala
1 Stockholm City South	12 Skåne Other	9 Gotland	4 Södermanland
12 Malmö	14 Västra Götaland Other	10 Blekinge	5 Östergötland
14 Gothenburg		23 Jämtland	6 Jönköping
			8 Kalmar
			13 Halland
			17 Värmland
			18 Örebro
			19 Västmanland
			20 Dalarna
			21 Gävleborg
			22 Västernorrland
			24 Västerbotten
			25 Norrbotten

5.4.8 Internal validation

The model's predictive accuracy on the derivation dataset was assessed with optimism corrected c-index and ROC curve analysis. To obtain c-index corrected for optimism, I implemented Harrell's bias correction method (Harrell, Lee, & Mark, 1996; Iba, Shinozaki, Maruo, & Noma, 2021). The final estimates and the estimates of the model's optimism were obtained by pooling the results from all iterations. The 95% confidence intervals were derived empirically from the corresponding quantiles.

The proportions of predicted and observed events at different levels of predicted probability and different time windows will be compared using calibration plots.

5.4.9 External validation

Estimates of coefficients in the final prediction rule were combined across imputations using Rubin's rule (Barnard & Rubin, 1999). The final model's discrimination was assessed on the external validation dataset using Harrel's c-index. The prediction error was estimated by calculating the Brier score. In addition, the Brier Skill Score was estimated relative to a performance of a naïve predictor, which assigns zero reoffending probability to all observations.

To test the final model, I used the external validation (holdout) sample derived based on the residential geographical location of the individual at the time of the sentence. The rationale for the selection of the external validation sample is described in 'Splitting'. The predictive accuracy of the model was summarised by Harrell's concordance index (c-index), the Brier score, and time-dependent ROC curve plots. The proportions of predicted and observed events at different levels of predicted probability will be compared using a calibration plot.

5.4.10 Model comparison

The performance of the final dynamic landmark model (DLM) was compared with two other models. First, the baseline model with fixed covariates (FBL) was trained using only the information available at a time of a sentence. For this model, the values of covariates were fixed at their baseline level and were not updated during the follow-up. Also, this model only incorporated the baseline hazard estimated at time 0 of the follow-up. Thus, FBL corresponds to the basic approach to developing a risk prediction model.

Second, the landmark model with fixed covariates (FLM) was fitted using the information available at all landmark times but without a dynamic update of the covariate values. For this model, the values of covariates were fixed at their baseline values and did not change during the follow-up. The occurrence of the triggers was not considered either. However, this model incorporated separate baseline hazard estimations for each landmark time. FLM corresponds to the basic approach of developing a simple dynamic risk prediction model.

No variable selection was performed for FBL and FLM. They consisted only of covariates included by default.

5.4.11 Software and coding

I extracted and pre-processed data using SAS, Version 9 for Windows (SAS Institute Inc., 2013). The data transformation and the model development were done in RStudio (R Core Team, 2013; RStudio Team, 2020). For the main analyses and modelling, I used the following R packages: *dynpred* (Putter & Putter, 2011), *survival* (Therneau, 2021), *Amelia* (Honaker et al., 2011), and *pROC* (Robin et al., 2011). The plots were created with the *survminer* (Kassambara et al., 2017) and *ggplot2* packages (Wickham et al., 2016). Additional visualisations were created in Figma (Figma Design, 2017).

Some code was adapted from van Houwelingen and Putter (2011) and Hazewinkel (2018). These works were instrumental in guiding my prediction model development process.

I also wrote custom code from scratch for variable selection, optimism corrected internal validation, out-of-sample prediction for the trained models, and calibration assessment. This was a necessity as many built-in functions of the available packages either do not work with landmark data or work very inefficiently on large datasets. The custom code is available in Appendix E8 (numbered out of order for the reader's convenience).

5.5 Results

I identified a cohort of 59,676 individuals given community sentences in Sweden between 2007 and 2013 (Table 5-4). The median follow-up time was 23 months. During the follow-up, 18,307 (31%) were reconvicted, and 4,416 (7%) committed a violent offence. The external validation sample contained 28% of all individuals. The derivation and external validation samples had similar baseline characteristics.

The number of individuals who have experienced triggers is presented in Figure 5-3 for both the derivation and external validation samples. The observed reoffending rates for selected landmarks (0, 12, 24, 36 months since the start of a community sentence) are presented in Figure 5-4 for the derivation sample and in Figure 5-5 for the external validation sample.

Table 5-4. Baseline characteristics and follow-up data of adult individuals receiving community sentences in Sweden from January 1, 2007 to December 31, 2013

The follow-up and censoring data are presented for violent reoffending outcome.

	Derivation sample	Validation sample
Number of individuals	43,192 (100%)	16,484 (100%)
Follow-up data		
Number of persons-year at risk	2,240.4	852.7
Incidents of general reoffending	13,300 (30.8%)	5,007 (30.4%)
Incidents of violent reoffending	3,220 (7.5%)	1,196 (7.3%)
Imprisoned during follow-up	2,489 (5.8%)	968 (5.9%)
Emigrated during follow-up	549 (1.3%)	237 (1.4%)
Died during follow-up	1,337 (3.1%)	542 (3.3%)
Covariates at baseline		
Sex (male)	36,794 (85.2%)	14,021 (85.1%)
Median age	32 (IQR: 23-45)	31 (IQR: 22-45)
Prior criminal history	30,050 (69.6%)	11,376 (69.0%)
Prior conviction for a violent crime	13,629 (31.6%)	5,124 (31.1%)
Prior prison sentence	7,781 (18.0%)	2,975 (18.0%)
Violent index offence	20,513 (47.5%)	7,732 (46.9%)
Married or in a registered partnership	6,505 (15.1%)	2,284 (13.9%)
Employed	20,354 (47.1%)	7,133 (43.3%)
Recipient of income support	12,037 (27.9%)	5,275 (32.0%)
Unstable housing situation	544 (1.3%)	254 (1.5%)
Highest level of education		
<9 years	2,187 (5.1%)	819 (5.0%)
9-11 year	35,381 (81.9%)	13,645 (82.8%)
>= 12 years	4,223 (9.8%)	1,527 (9.3%)
History of self-harm	3,934 (9.1%)	1,577 (9.6%)
Previous psychiatric disorder		
Any psychiatric disorder	19,454 (45.0%)	7,425 (45.0%)
Any psychiatric disorder (other than substance use)	13,731 (31.8%)	5,582 (33.9%)
Any severe psychiatric disorder	2,250 (5.2%)	784 (4.8%)
Schizophrenia spectrum disorder	1,463 (3.4%)	546 (3.3%)
Bipolar disorder	787 (1.8%)	238 (1.4%)
Substance (drug or alcohol) use disorder	13,591 (31.5%)	4,955 (30.1%)
Alcohol use disorder	9,507 (22.0%)	3,200 (19.4%)
Drug use disorder	7,702 (17.8%)	3,047 (18.5%)

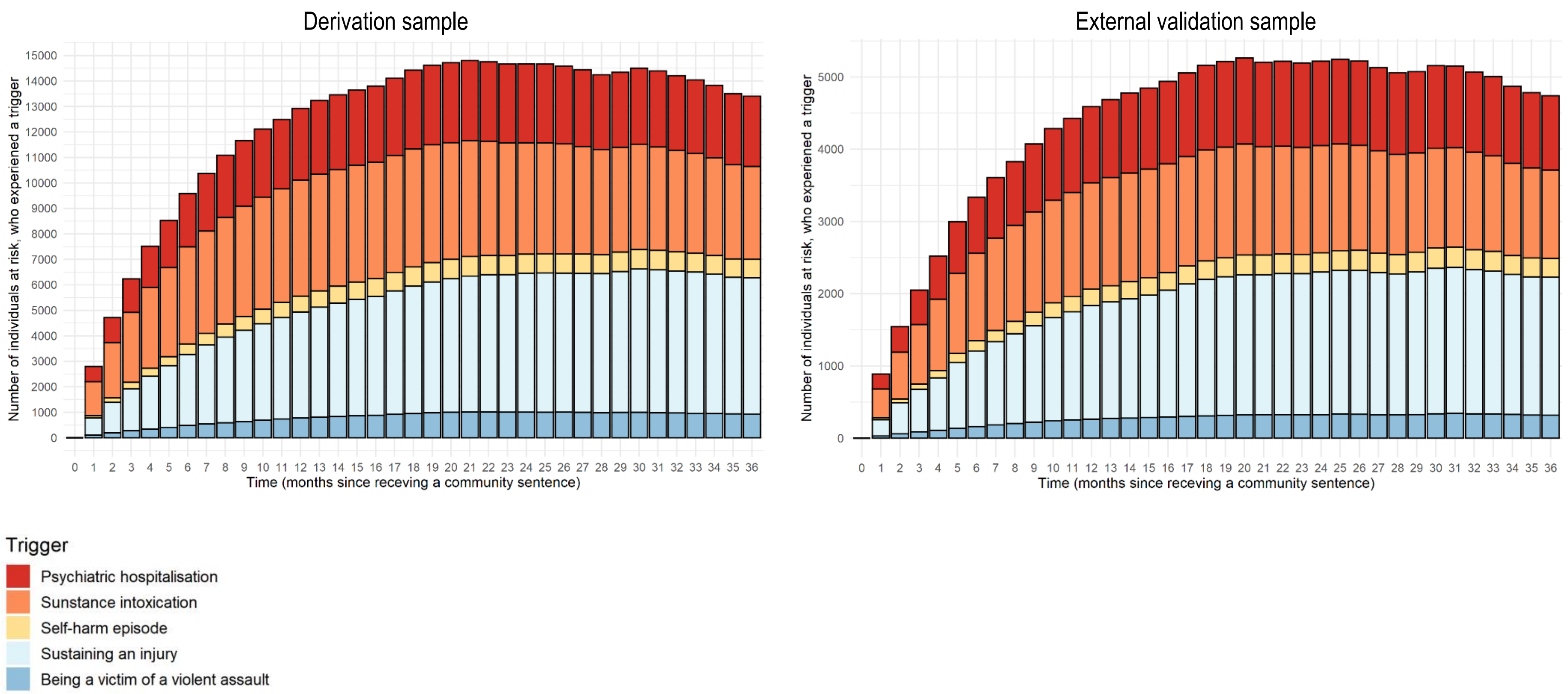


Figure 5-3. Cumulative prevalence of individuals who have experienced a trigger in the derivation and external validation samples over time

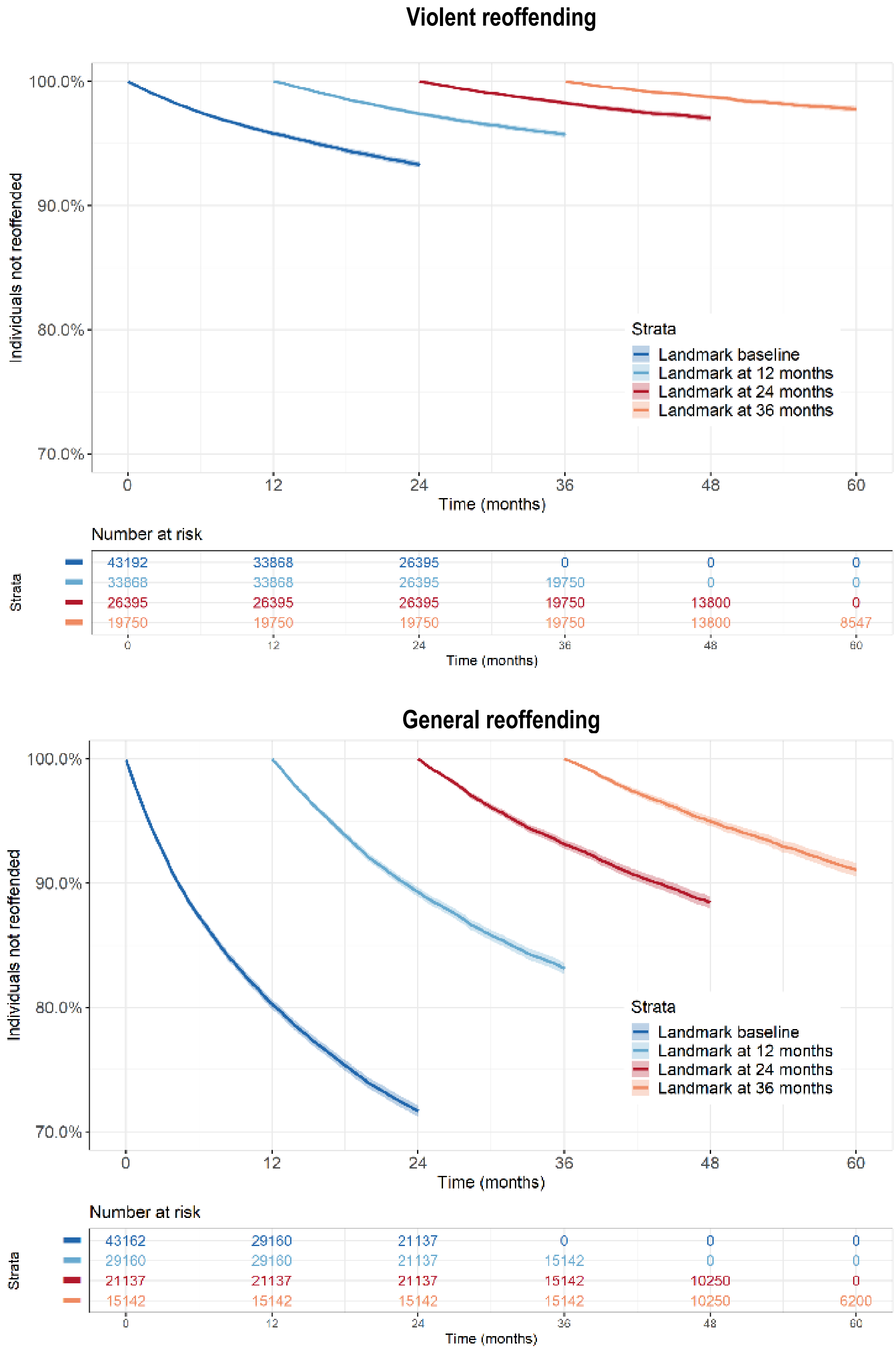


Figure 5-4. Kaplan-Meier curves for selected landmarks in the derivation dataset

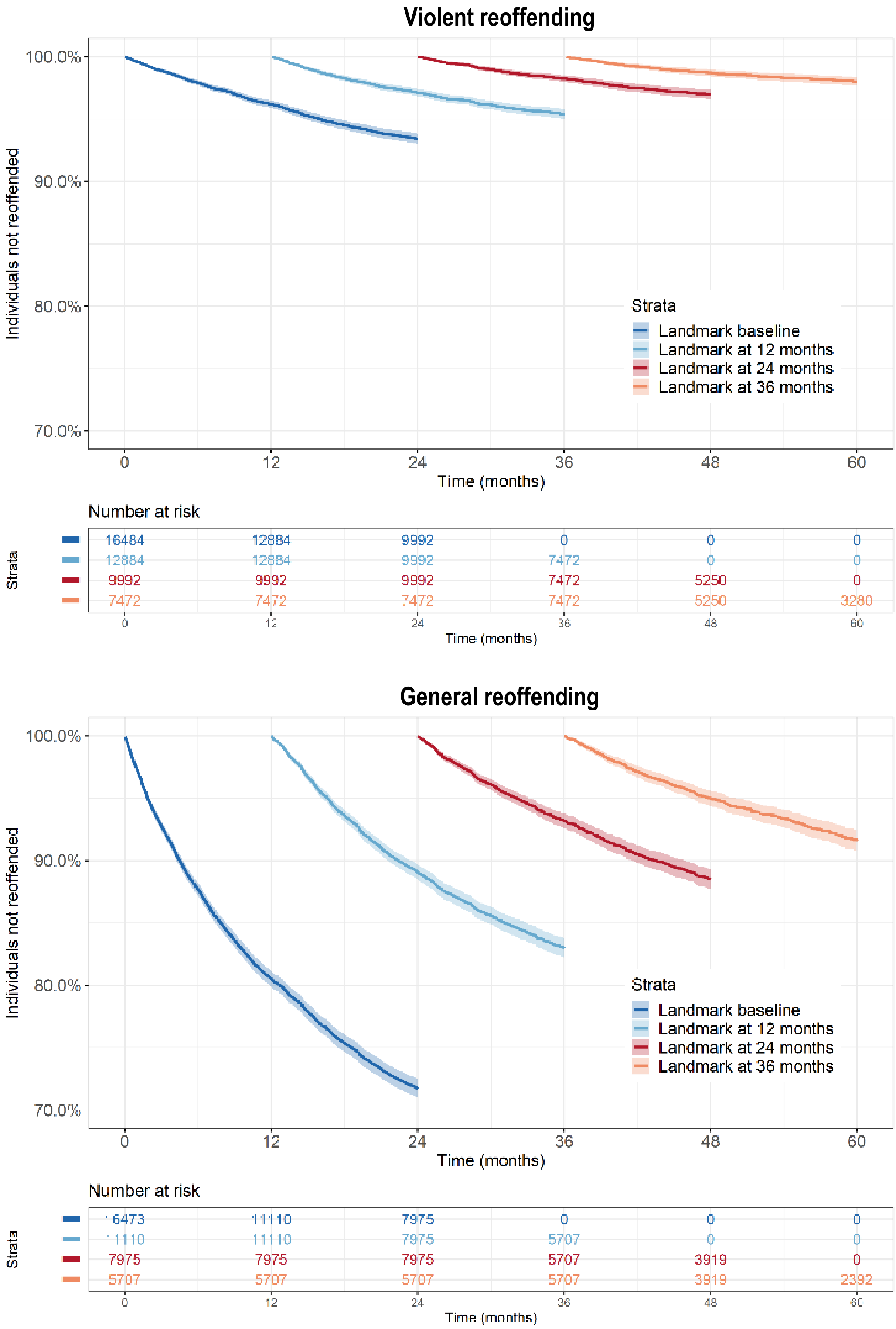


Figure 5-5. Kaplan-Meier curves for selected landmarks in the external validation dataset

5.5.1 Final model specification

Several variables were re-operationalised after the first variable selection (see results in Appendix E4). The dynamic *Self-harm episode* covariate was excluded as a predictor for general offending as a result of variable selection. However, *History of self-harm or suicide attempts* measured at the time of the sentence remained in the model. This resulting model could lead to a tool that ignores a suicide attempt during the post-sentence supervision but accounts for a self-harm episode several years ago. The two self-harm variables were subsequently transformed into *Any self-harm episode or suicide attempt prior to the time of the assessment*. Following the same rationale, the variables *Injuries from other causes* and *TBI* were also combined into one variable *TBI or injuries from other causes*. I also eliminated baseline *Civil status* from the model. It did not add anything to the models' performance above what *Current civil status* adds. Baseline *Civil status* is also irrelevant for risk management. Having two separate variables for civil status could lead to a practitioner's confusion without a good reason. So, I repeated the variable selection with the re-operationalised covariates and removed baseline *Civil status* (Appendix E5).

Besides the covariates included by default, the final dynamic landmark model (DLM) for violent reoffending included: *Being a victim of a violent assault*, *TBI or injuries from other causes*, *Any psychiatric hospitalisation*, *Substance intoxication*, *Any prior self-harm or suicide attempt*, *Employment at baseline.* All these variables were also included in the final model for general reoffending, although included triggers had different time components. DLM for general reoffending also included the receipt of income support at baseline as a predictor. The baseline hazards for general and violent reoffending within two years were also extracted for each landmark. For DLM formulae and coefficients, refer to Appendix E6.

5.5.2 Internal validation

The optimism corrected c-index for DLM was 74.4% for 2-year violent reoffending and 69.4% for 2-year general reoffending, indicating good overall discrimination (Table 5-5). In the ROC analysis, DLM also demonstrated good discrimination over time. For violent reoffending, the AUC at the start of a sentence was 76.2 and decreased to 74.9 in individuals with an offence-free time of 36 months (Figure 5-6). For general reoffending, the AUC at the start of a sentence was 74.6 and decreased to 66.9 in individuals with an offence-free time of 36 months (Figure 5-7). DLM also demonstrated moderate prediction error for 2-year violent and general reoffending probabilities as assessed by Brier scores (Figure 5-9). The performance gain relative to a naïve estimator, which assigned zero probability of outcomes to every individual, was more apparent for general reoffending. The prediction error also increased over time, approaching the performance level of a naïve estimator.

Table 5-5. The results of internal validation using Harrell's bias correction algorithm

Ten imputed datasets, 200 bootstrap iterations for each imputed datasets (2,000 bootstrap iteration in total). The 95% confidence intervals were estimated from quantiles of the bootstrapped values.

	Violent	**General**
Apparent c-index (%)	74.60	69.49
Bootstrapped c-index (%)		
Fit on a bootstrapped sample	74.68 (73.83-75.50)	69.51 (69.06-69.97)
Fit on the original sample	74.51 (74.39-74.60)	69.45 (69.41-69.49)
Optimism-corrected c-index (%)	***74.44***	***69.43***

5.5.3 External validation

DLM demonstrated good overall discrimination on the external validation dataset. For violent reoffending, the AUC at the start of a sentence was 76.0 and decreased to 71.4 in individuals with an offence-free time of 36 months (Figure 5-6). For general reoffending, the AUC at the start of a sentence was 75.5 and decreased to 67.9 in individuals with an offence-free time of 36 months (Figure 5-7).

DLM demonstrated marginally better discrimination performance over time for general and violent reoffending compared to the landmark model with fixed covariates (FLM) and the baseline model with fixed covariates (FBL) (Figure 5-8). However, despite being consistent, this difference in performance was not statistically significant. The point estimates for the discrimination performance of FLM and FBL were almost identical. Both models used the covariate values estimated at baseline.

Similarly to the internal validation, DLM demonstrated moderate prediction error for 2-year violent and general reoffending probabilities as assessed by Brier scores (Figure 5-9). The performance of DLM and FLM was comparable. Both landmark models outperformed FBL, which accumulated prediction error at a much faster rate.

DLM demonstrated good overall calibration in the external validation dataset for violent reoffending (Figure 5-11, Figure 5-13) and general reoffending (Figure 5-12, Figure 5-14). The calibration for low- and medium-risk individuals was better than for high-risk individuals. For the individuals in the highest risk decile at the start of their sentence, DLM overestimated the probability of violent reoffending within two years by 6% and the probability of general reoffending within two years by 8%. For the individuals in the highest risk decile at month 24 of their sentence, DLM overestimated the probability of violent reoffending within two years by 4% and the probability of general reoffending within two years by 8%. FLM also demonstrated good calibration. The calibration curves for DLM and FLM were similar.

The calibration of FBL was poor. Although the model demonstrated good calibration at baseline, over time, FLM substantially overestimated the probabilities of violent and general reoffending for all risk deciles. For the individuals in the highest risk decile at the start of their sentence, FBL overestimated the probability of violent reoffending within two years by 6% and the probability of general reoffending within two years by 9%. For the individuals in the highest risk decile at month 24 of their sentence, FBL overestimated the probability of violent reoffending within two years by 13% and the probability of general reoffending within two years by 37%. The calibration gradient shows the continuous accumulation of bias towards overestimation in FBL over time (Figure 5-13, Figure 5-14).

Overall, DLM demonstrated virtually the same discrimination performance as FLM and FBL. DLM and FLM also had similar calibration and prediction error estimates. Both landmark models (DLM and FLM) that accounted for the offence-free time had substantially better calibration over time compared to the baseline model with fixed covariates (FLM) that did not account for offence-free time.

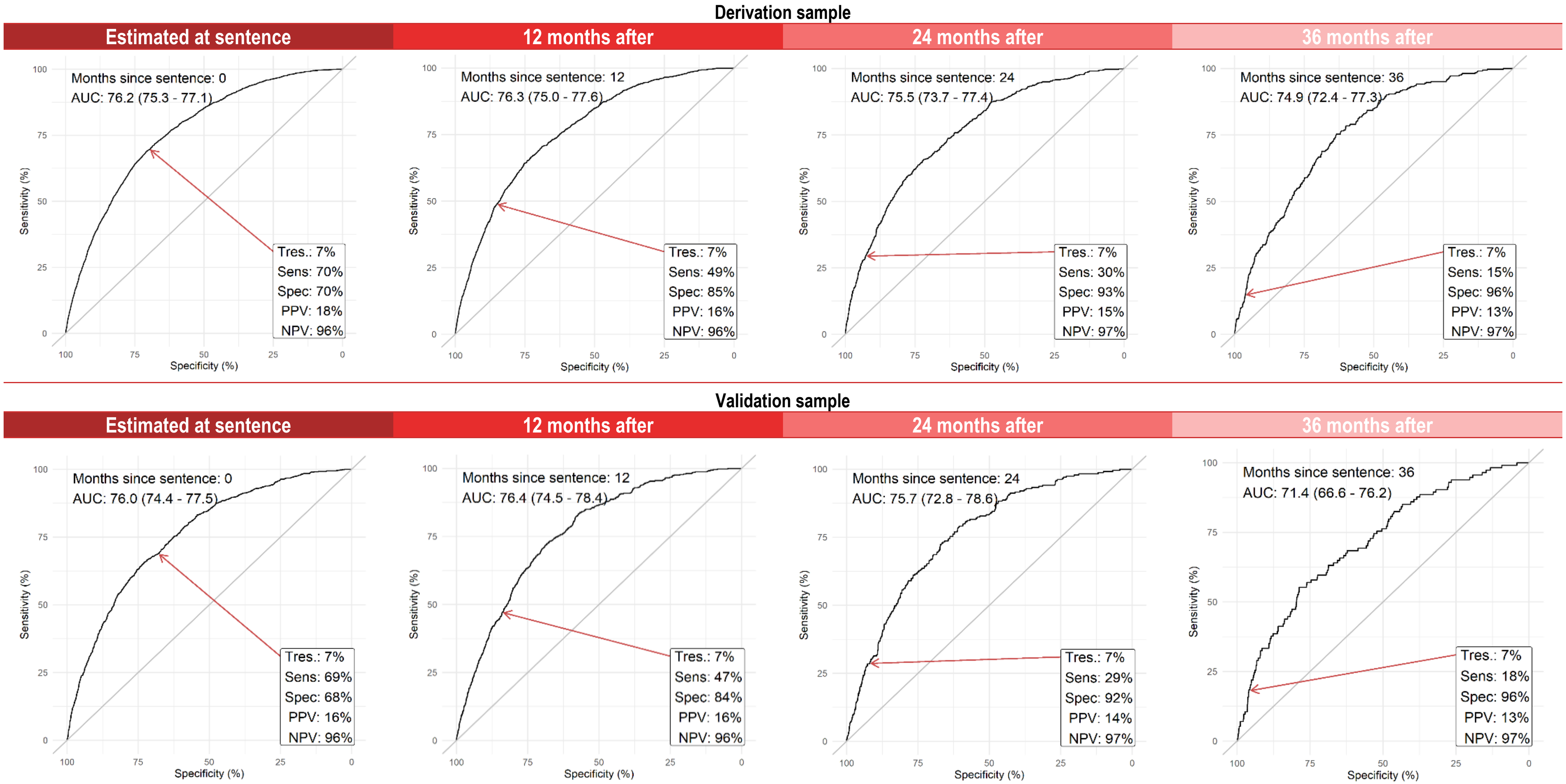


Figure 5-6. Dynamic Landmark Model (DLM) discrimination shown by receiver operating characteristics curves for 2-year violent reoffending

Estimates for a given threshold. Tres. = threshold. Sens = sensitivity. Spec = specificity. PPV = positive predictive value. NPV = negative predictive value. AUC = area under curve.

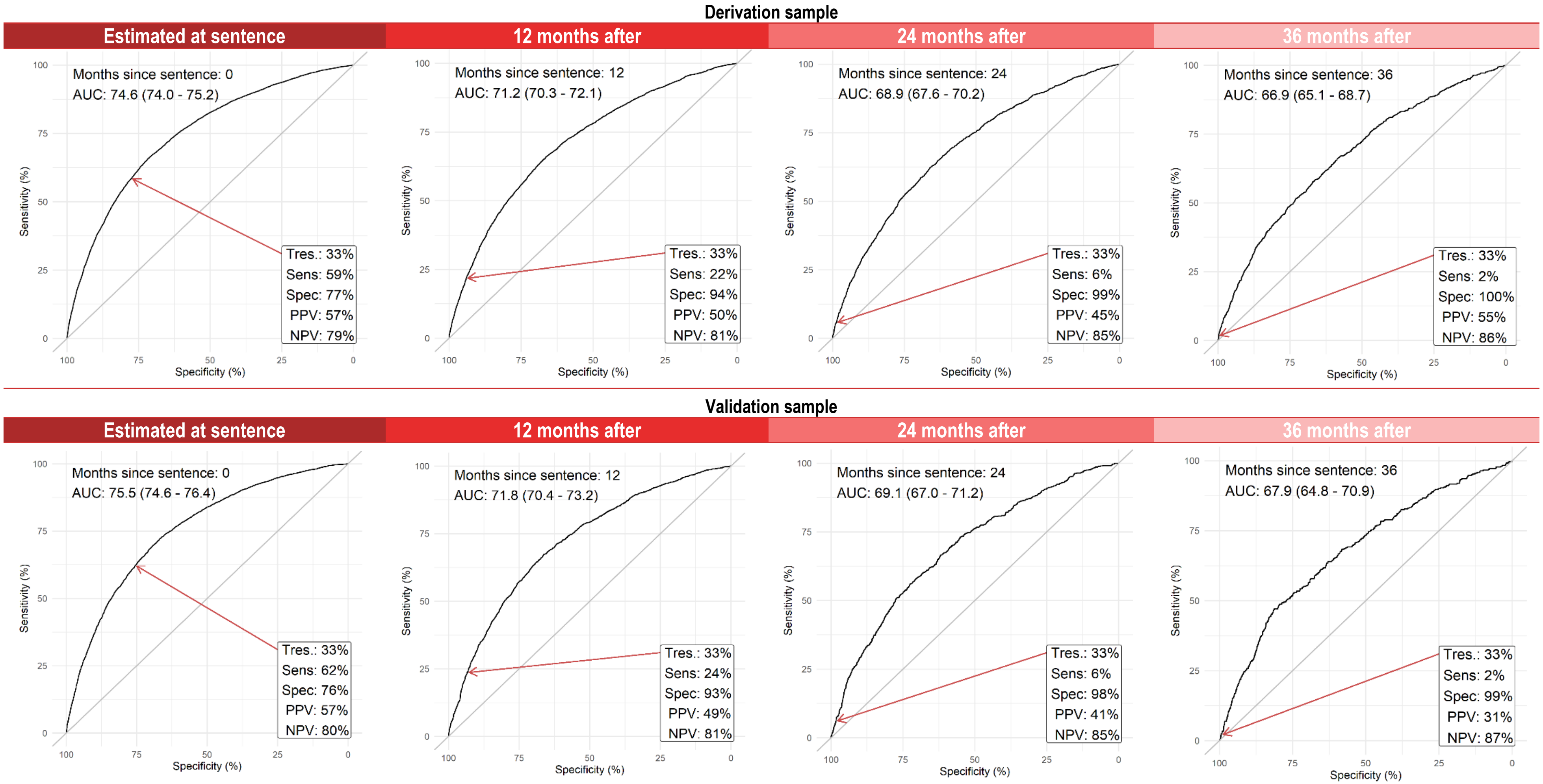


Figure 5-7. Dynamic Landmark Model (DLM) discrimination shown by receiver operating characteristics curves for 2-year general reoffending

Estimates for a given threshold. Tres. = threshold. Sens = sensitivity. Spec = specificity. PPV = positive predictive value. NPV = negative predictive value. AUC = area under curve.

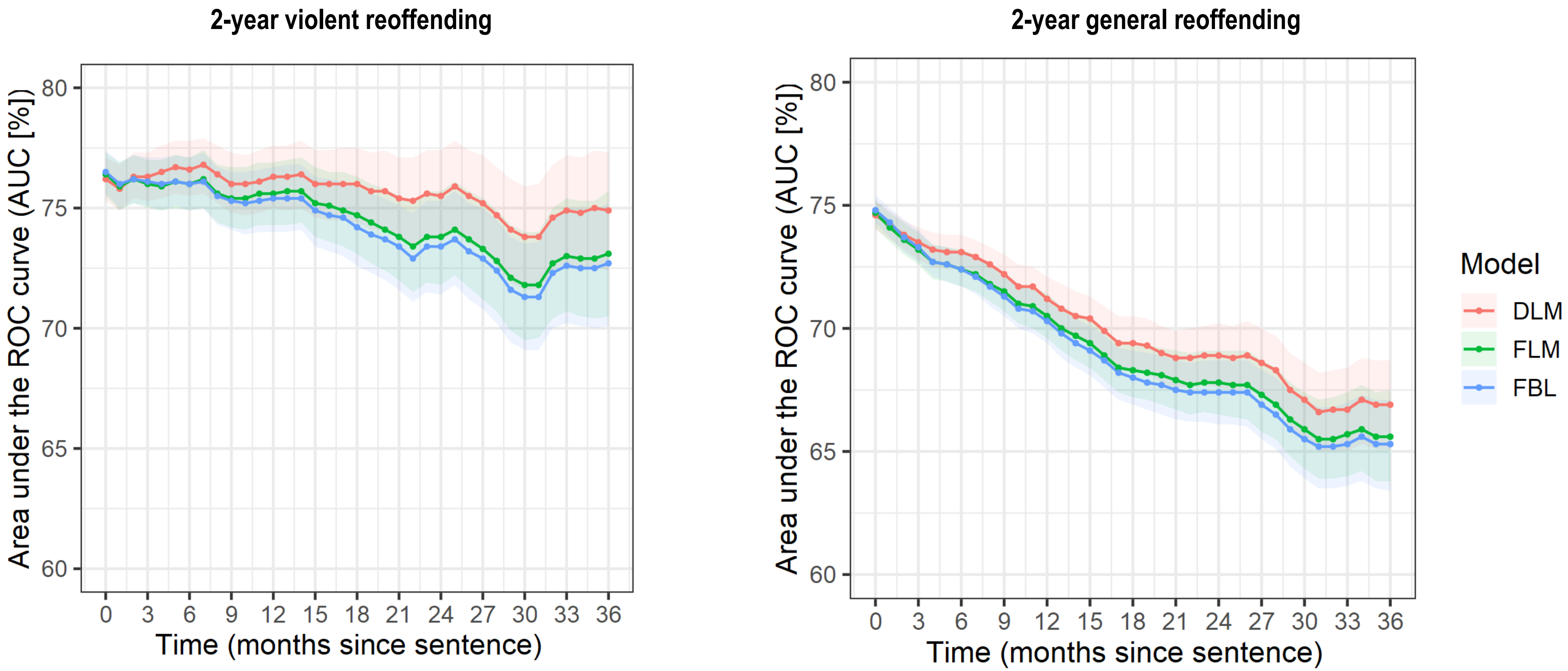


Figure 5-8. Discrimination performance of the prediction models over time in the external validation dataset

DLM = dynamic landmark model. FLM = landmark model with fixed covariates. FBL = baseline model with fixed covariates. DLM and FBL account for offence-free time. FBL only uses covariate values at baseline and does not account for offence-free time.

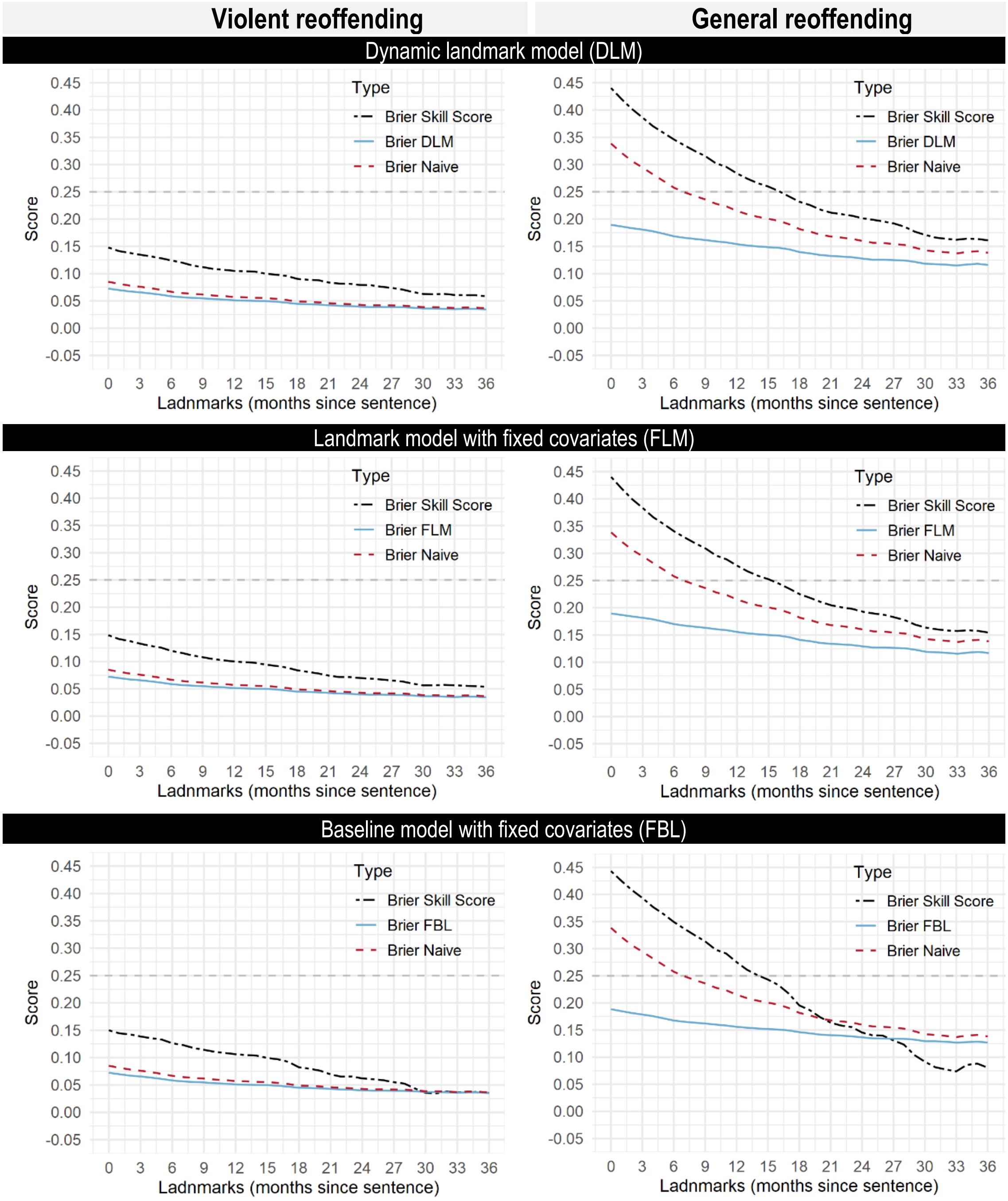


Figure 5-9. Prediction error estimates in the derivation dataset

Baseline hazards in models DLM and FLM vary over time. DLM is the main model for risk monitoring developed in the current study. FLM is the same model but with all covariate values fixed at time 0 and no trigger variables included. FBL is the same as FLM, but its baseline hazard value was fixed at time 0. The naïve predictor assigned zero probability of reoffending to all individuals.

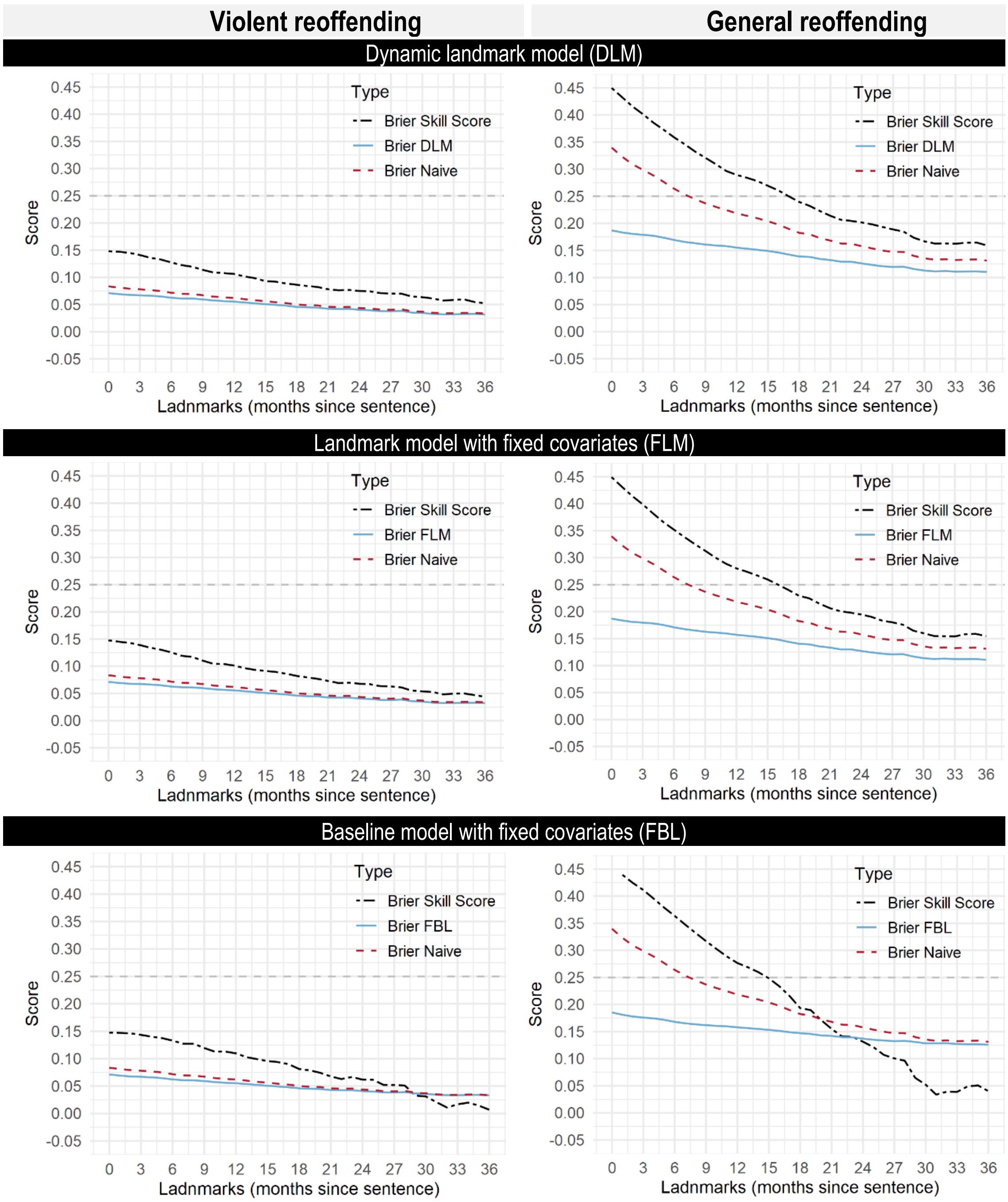


Figure 5-10. Prediction error estimates in the external validation dataset

Baseline hazards in models DLM and FLM vary over time. DLM is the main model for risk monitoring developed in the current study. FLM is the same model but with all covariate values fixed at time 0 and no trigger variables included. FBL is the same as FLM, but its baseline hazard value was fixed at time 0. The naïve predictor assigned zero probability of reoffending to all individuals.

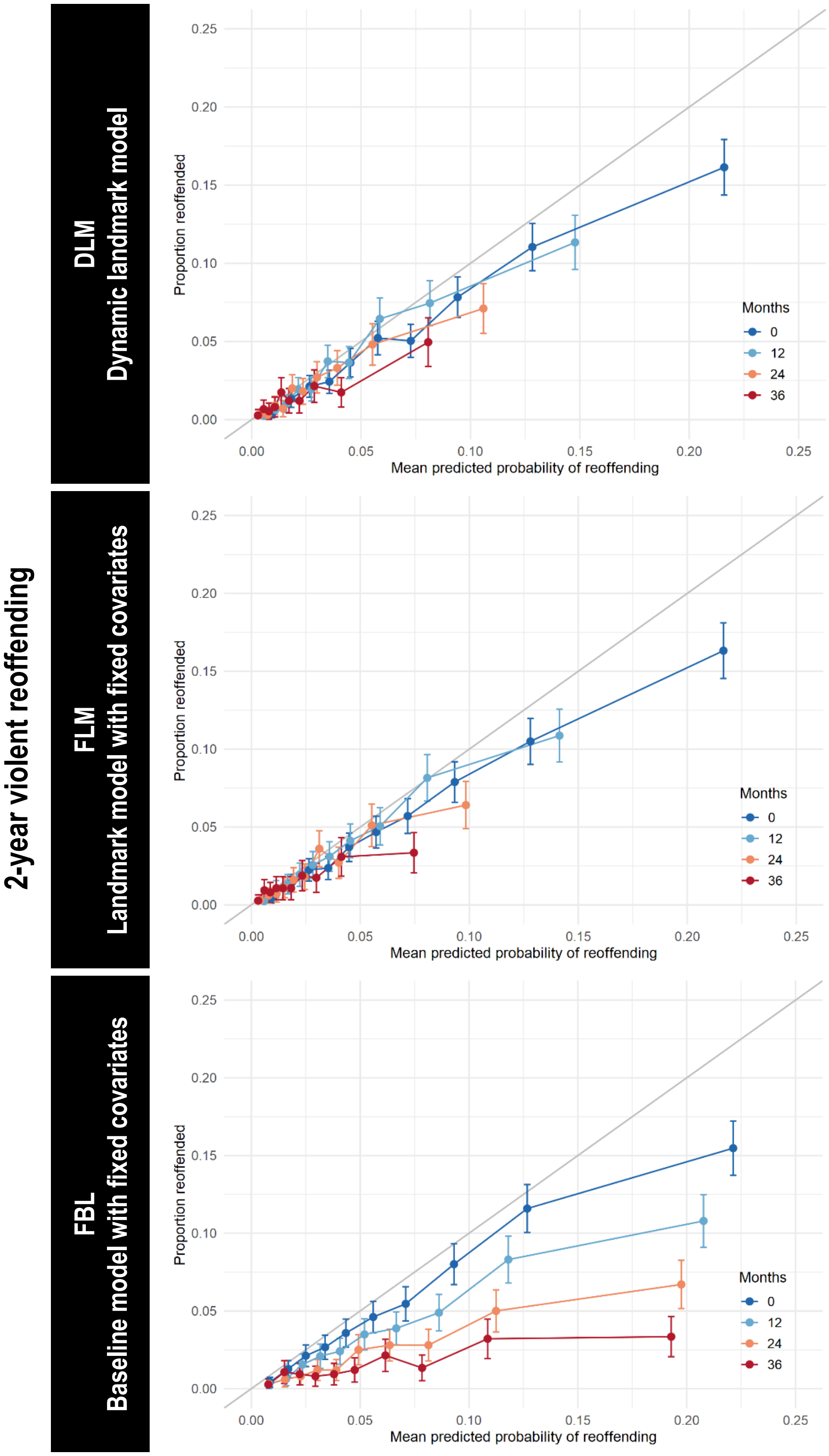


Figure 5-11. Calibration curves for the prediction of violent reoffending

Estimates for selected landmarks in the external validation dataset. The individuals grouped in deciles by predicted probability of general reoffending within 2 years.

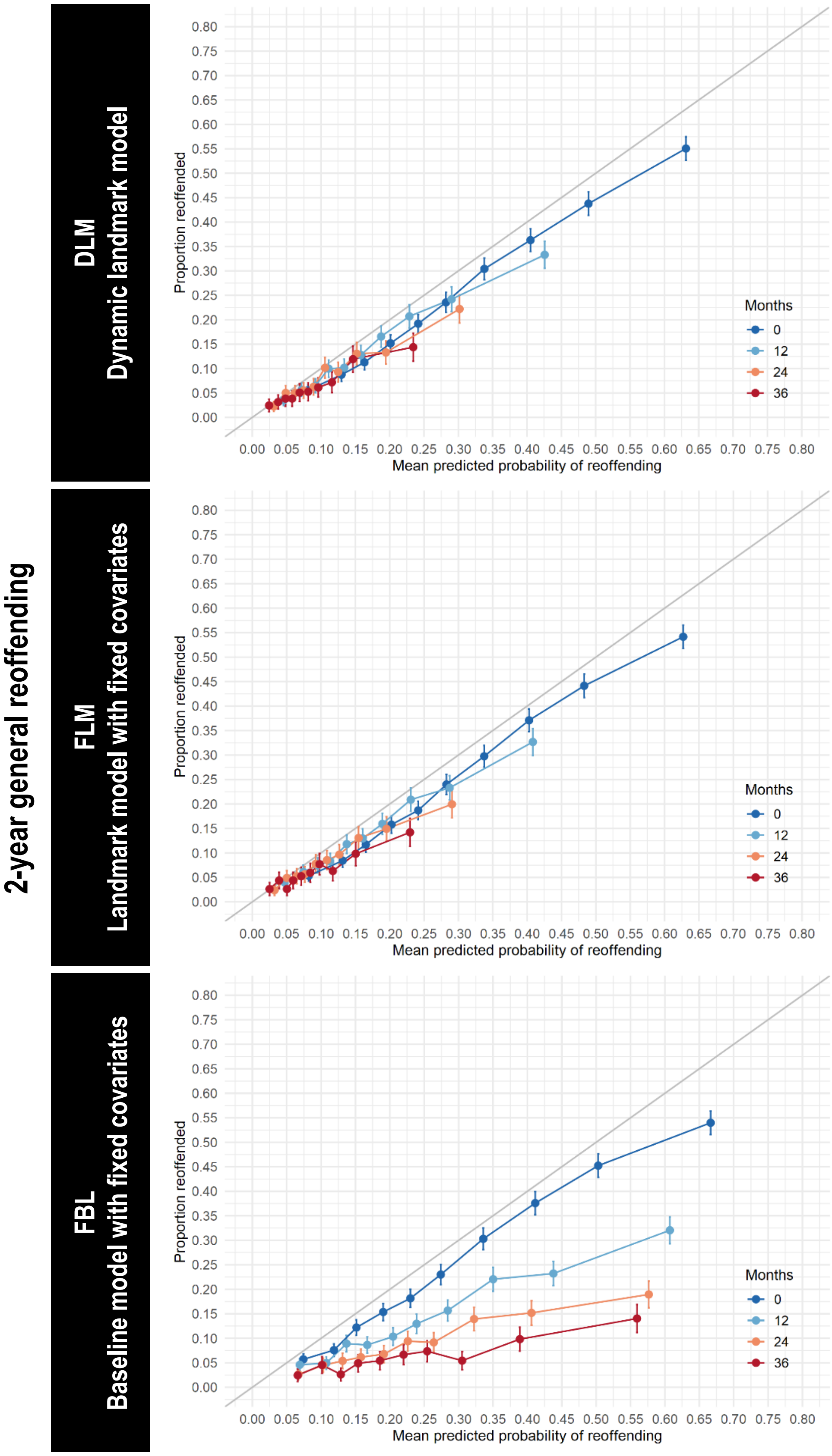


Figure 5-12. Calibration curves for the prediction of general reoffending

Estimates for selected landmarks in the external validation dataset. The individuals grouped in deciles by predicted probability of general reoffending within 2 years.

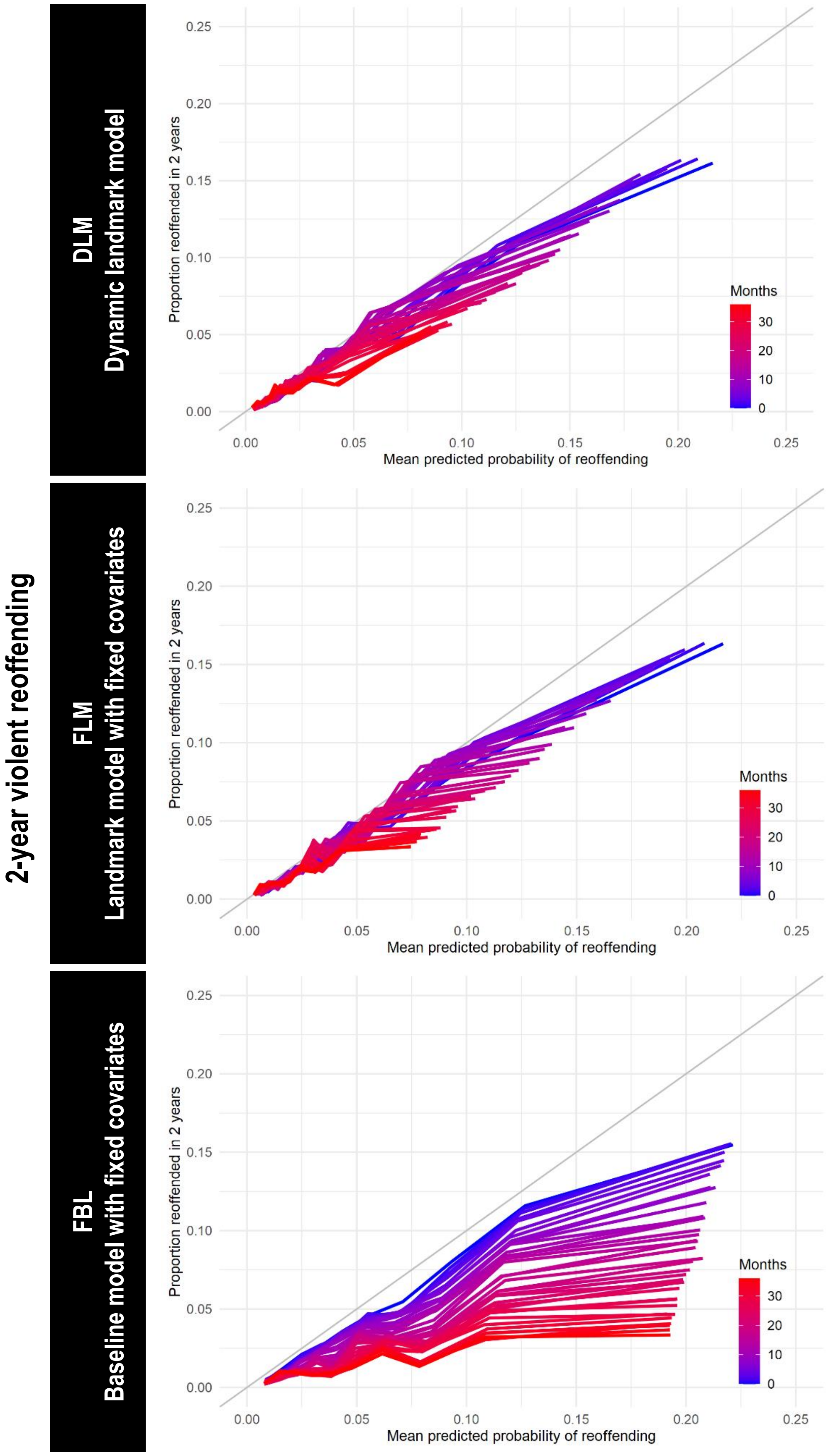


Figure 5-13. Calibration gradient for violent reoffending (calibration over time)

Each point of the curve is the mean probability of reoffending within a decile.

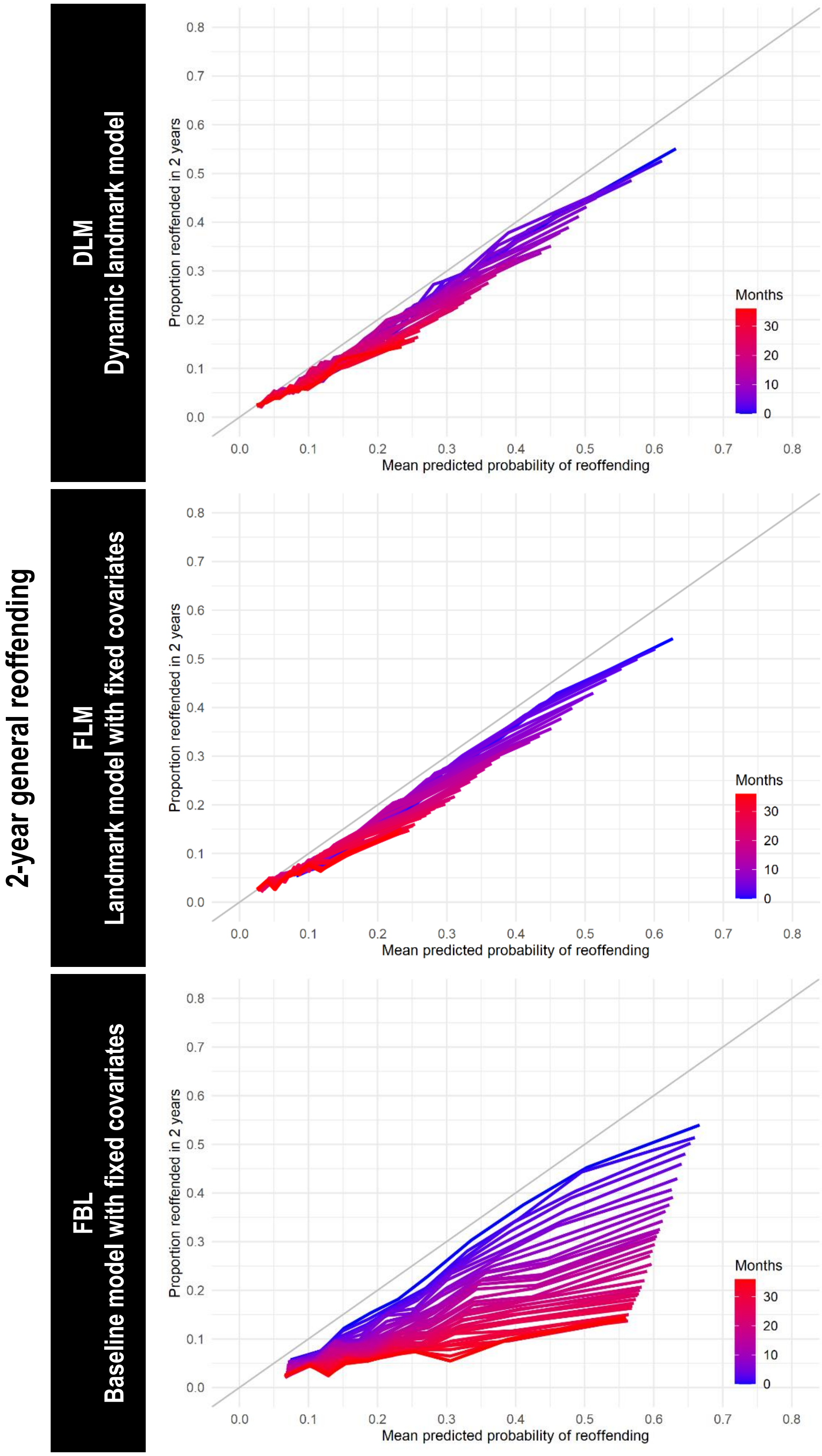


Figure 5-14. Calibration gradient for general reoffending (calibration over time)

Each point of the curve is the mean probability of reoffending within a decile.

5.5.4 Hypothetical vignettes

I applied DLM to simulated datasets containing information about hypothetical individuals given community sentences. The individuals were supervised in the community for three years since the start of their sentence. Their probation officer conducted regular monthly assessments. The resulting vignettes illustrate a practical application of the developed criminal recidivism risk monitoring tool.

High risk

Vlad is 24 y.o. (at the start of the community supervision) unemployed man with a prior criminal history. He had been imprisoned for a violent offence in the past. His index offence was non-violent. He was diagnosed with bipolar disorder, and he also experienced several self-harm episodes before the current sentence.

During month six of his sentence, Vlad got involved in a fight at the local bar. He was assaulted and injured. During month 13, he was hospitalized with substance intoxication and received an alcohol use diagnosis. During month 25, he was hospitalized with a manic episode with psychotic features for two weeks.

Despite his struggles, Vlad managed to remain offence-free and found a stable job by the end of his community sentence. For Vlad's risk trajectory, refer to Figure 5-15.

Low risk

Anastasia is 35 y.o. (at the start of the community supervision) an employed woman without a prior criminal history. She is married. Her index offence was violent. She was diagnosed with paranoid schizophrenia several years before the index sentence.

During month four of her sentence, she was hospitalized with a psychotic episode and spent one week in a hospital. During month 24, she lost her job and was not able to find a new one. She was granted income support seven months later. For Anastasia's risk trajectory, refer to Figure 5-16

Medium risk

Dmitriy is 25 y.o. (at the start of the community supervision) employed man with a prior criminal history. He received a community sentence in the past for a violent offence. His index offence was not a violent offence. In the past, he was diagnosed with schizoaffective disorder, drug use disorder, and alcohol use disorder.

Throughout his sentence, Dmitriy experienced the same events as Anastasia did. During month four of the sentence, he was hospitalized with a psychotic episode for a week. During month 24, he lost his job and was not able to find a new one. He was granted income support seven months later. For Dmitriy's risk trajectory, refer to Figure 5-17.

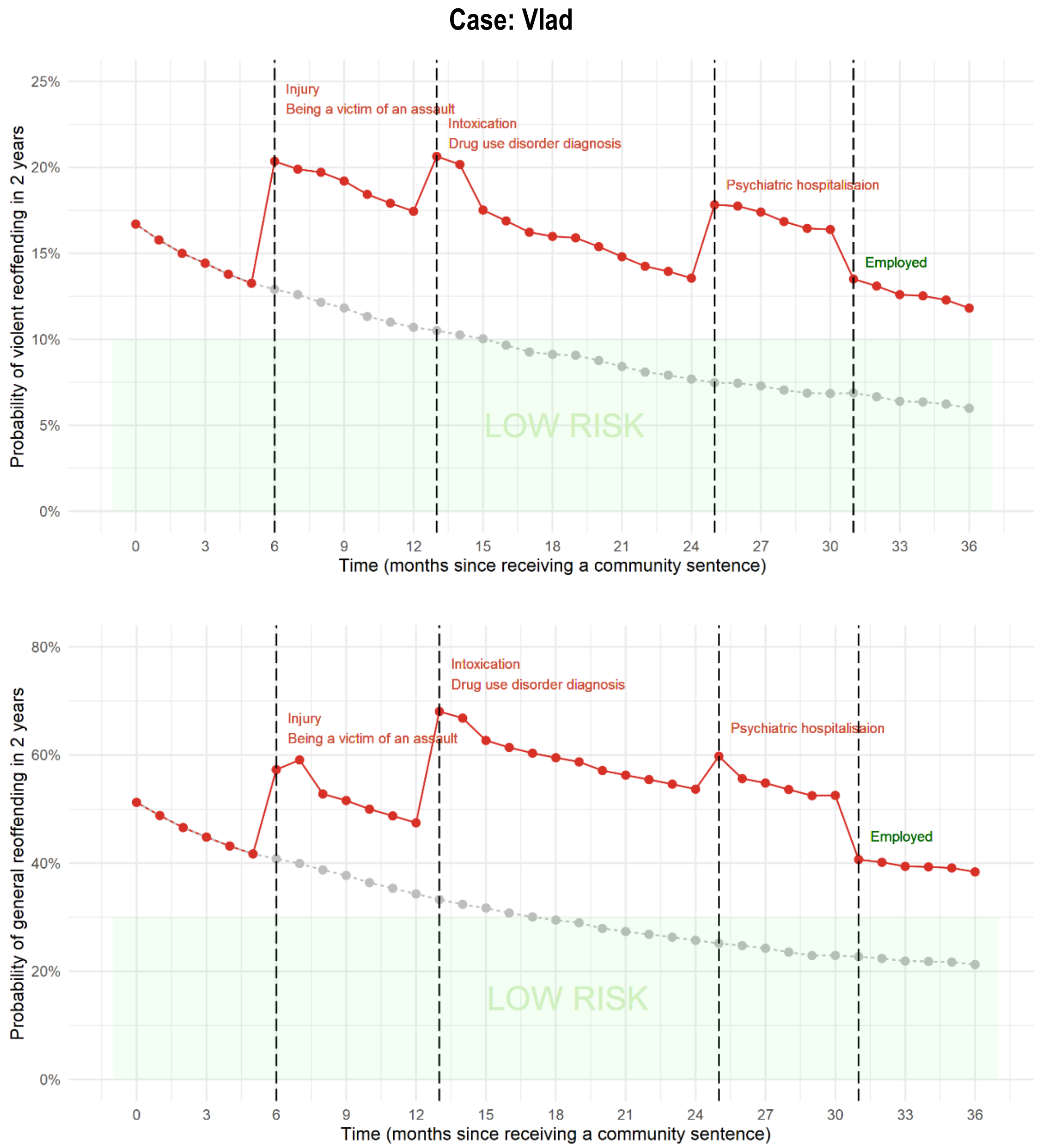


Figure 5-15. Predicted risk trajectories for a hypothetical high-risk individual

The grey line represents the reoffending risk trajectory given no covariates change their values from baseline. 'Low risk' area is depicted for comparison using a threshold of 30% for general reoffending and 10% for violent reoffending.

Case: Anastasia

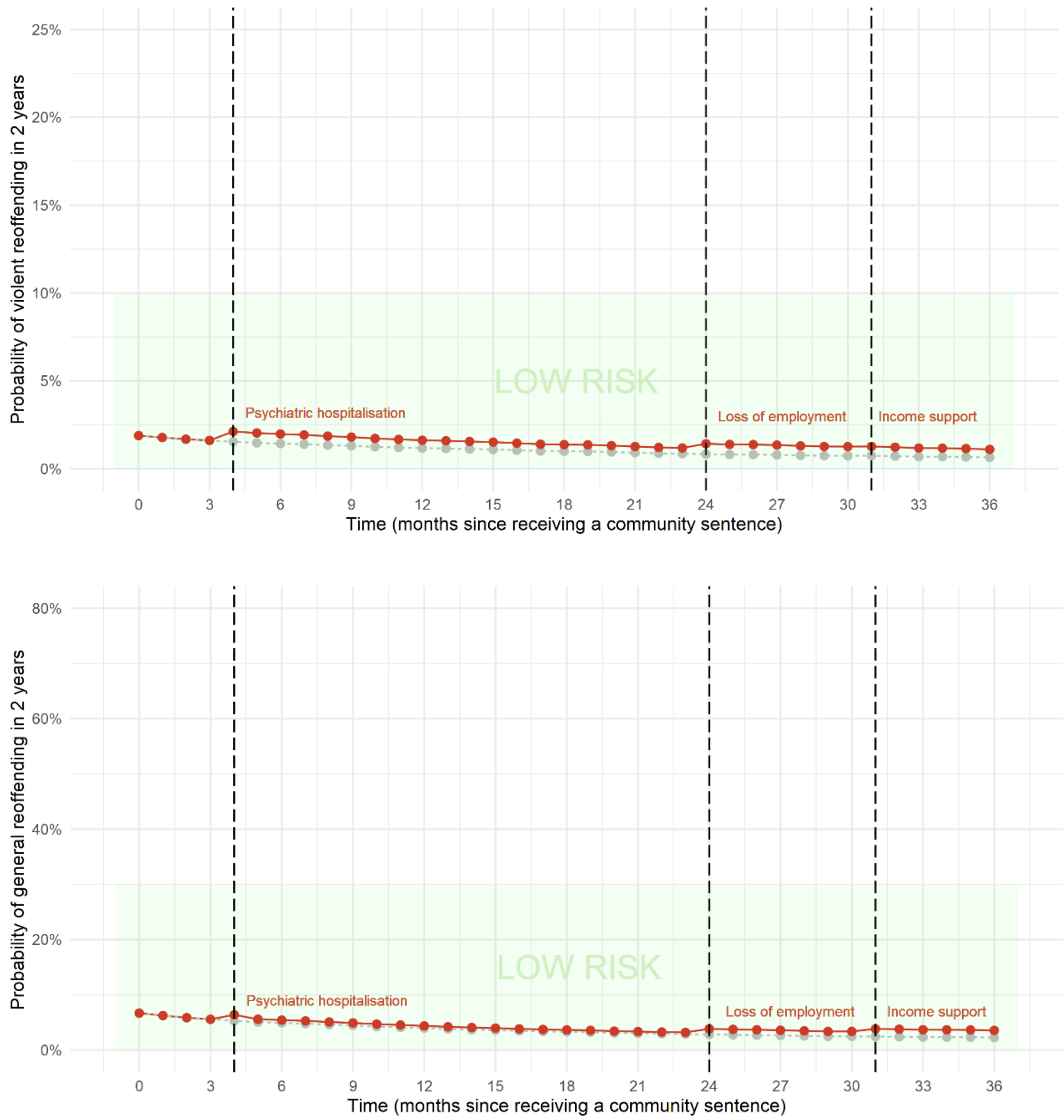


Figure 5-16. Predicted risk trajectories for a hypothetical low-risk individual

The grey line represents the reoffending risk trajectory given no covariates change their values from baseline. 'Low risk' area is depicted for comparison using a threshold of 30% for general reoffending and 10% for violent reoffending.

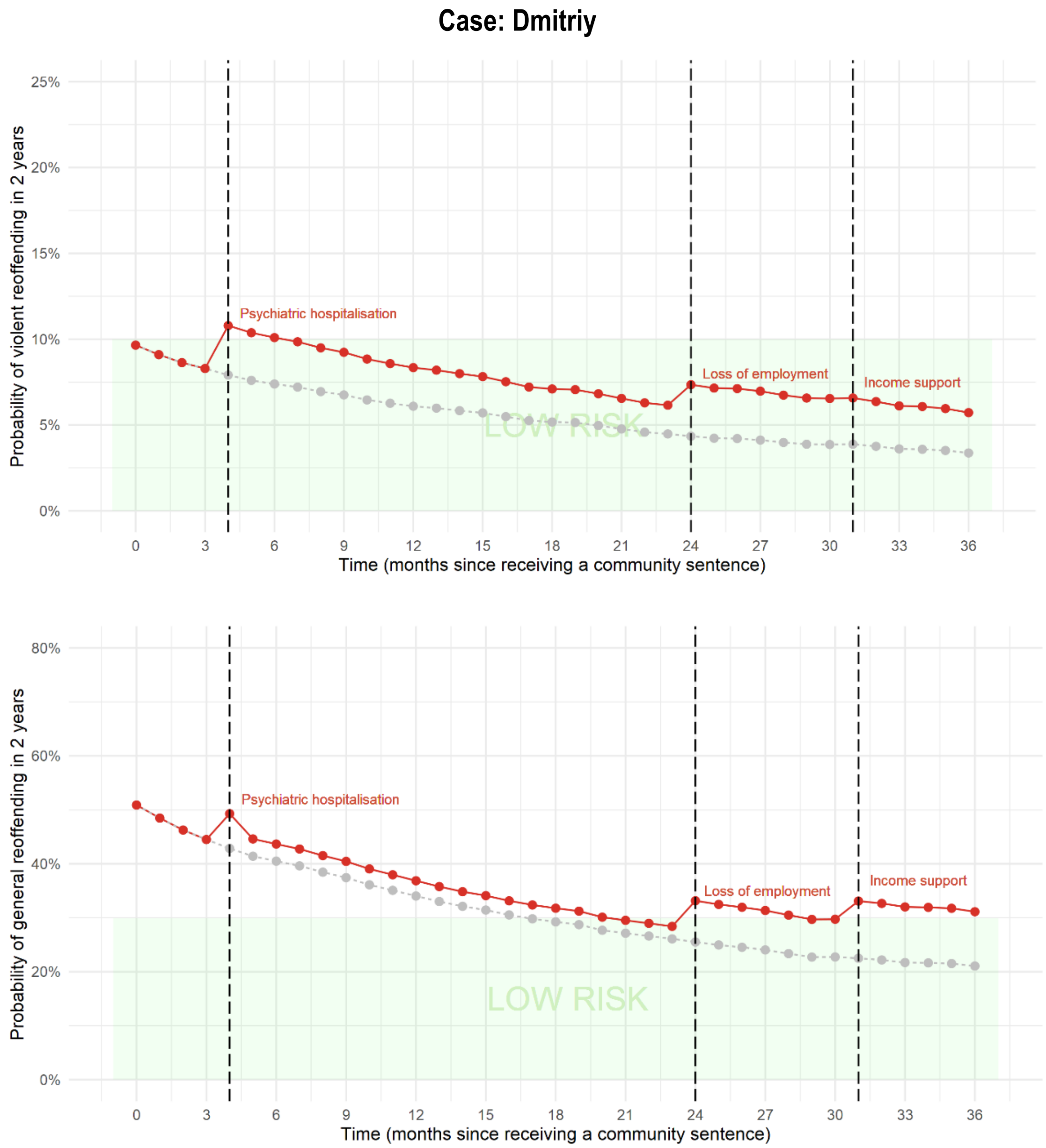


Figure 5-17. Predicted risk trajectories for a hypothetical medium-risk individual

The grey line represents the reoffending risk trajectory given no covariates change their values from baseline. ‘Low risk’ area is depicted for comparison using a threshold of 30% for general reoffending and 10% for violent reoffending.

5.6 Discussion

I derived a dynamic prediction model for criminal recidivism risk monitoring on a sample of 43,192 individuals given community sentences in Sweden from 2007 to 2013. The model was additionally validated using the holdout sample of 16,484 individuals, which was similar in composition to the derivation sample. The resulting prediction model takes into account adverse events that might occur during the community supervision (triggers), changes in a supervised individual’s circumstances, and offence-free time. The model has good calibration and discrimination performance (c-index = 0.74 for violent reoffending, c-index = 0.69 for general reoffending). The model was deployed as an online tool for criminal recidivism risk assessment. The tool included 23 predictors for violent reoffending and 28 predictors for violent reoffending. As an important secondary outcome, the study demonstrated that actuarial recidivism risk assessment tools, which have not been developed as risk monitoring instruments but are used as such, will systematically overestimate the risk of recidivism over time.

Many risk assessments instrument work under the assumption that an individual’s risk of reoffending remains constant over time, given that the individual's and their environment's measured characteristics remain the same. However, the studies on desistance from crime have demonstrated that the more time an individual remains offence-free, the less likely they reoffend (Bushway et al., 2011; Hanson et al., 2018). In this study, incorporating offence-free time significantly improved the developed model discrimination and calibration. However, the inclusion of the dynamic covariates leads to only non-significant, although consistent, improvement. The improvement was more pronounced at the later stages of the follow-up period, which might indicate higher predictive validity of temporally proximal covariates in the latter stages of the supervision. These findings underscore the need for dynamic predictive models of reoffending to incorporate the distance effect. Otherwise, they are likely to produce biased scores for individuals refraining from offending. Many standardized risk assessment tools will not pick up on this risk reduction as the model's desistance time was not built-in.

An example of the actuarial tool that incorporates offence-free time for the dynamic violence risk assessment in individuals given community sentences is the Offender Assessment System (OASys) used in the UK (Howard & Dixon, 2012). However, the OASys is a comprehensive tool battery that takes a considerable amount of time to complete. Because of its complexity, it cannot serve as a screening instrument, and it may be challenging to adapt the OASys for other countries. The risk assessment tool OxMore developed in this can be used for screening and is potentially more generalisable.

As part of this discussion, I would also like to point out that the actuarial risk assessment tools are not merely calculators. They are used in the context of meaningful interpersonal activity as instruments for one person to understand another person. Thus, they are likely to be interiorised by users and integrated into their mental model of the risk assessment process (Vygotsky, 2012). In other words, a frequently used risk assessment instrument will tend to become a part of a user’s cognitive schema. The developers of empirically based actuarial risk assessment instruments, like mine, should consider what ‘mental model of risk’ they implicitly create in a user’s mind. This mental model could depend on many factors: included covariates, their coefficients, their order, their wording, instructions to the user, and even the tool's interface.

I would argue that similar principles should be applied to developing risk assessment tools as to developing websites or mobile applications. The developed tools should undergo some usability testing not only to test whether users understand the interface but also to ensure that the intended mental model of risk is clearly communicated to them. The usability test is something that the tool developed in the current study can also benefit from.

5.7 Strengths and limitations

The study has several strengths. I developed a risk assessment tool using data from a large cohort of individuals given community sentences. Using a large population cohort ensured the stability of the derived models and the absence of selection bias. The model also included predictors that are easily generalisable across different countries and legal settings. This generalisability enables easy recalibration and redeployment of the developed dynamic risk assessment tool outside of Sweden. I also covered the calibration of the developed model in detail. Unfortunately, reporting the calibration can be considered a strength, given that most publications describing the development of prediction models omit the calibration entirely (Collins et al., 2014). The developed risk assessment tool has good discrimination and low optimism, ensuring good performance outside the derivation sample. Another strength is that the developed tool accounts for offence-free time, leading to more accurate reoffending risk estimates over time compared to other commonly used models. The risk assessment tool outputs the probability of reoffending and baseline risk for a given offence-free time. This approach does not use risk categories and provides considerable flexibility to an end-user.

The study also has several limitations. First, population registers do not contain data on potentially significant predictors of criminal recidivism that have been identified in published research, such as lack of social support and association with antisocial peers. I also did not have information on any intervention programmes that sentenced individuals might have undergone. The successful completion of the rehabilitation program could be a potential strong desistance factor. In addition, the sociodemographic covariates are only recorded in population registers once a calendar year for all individuals in Sweden. This cross-sectional approach limits the validity of sociodemographic variables as dynamic predictors since they are unlikely to capture sudden changes in an individual's life circumstances. Second, I relied on data from patient registers for ascertainment of psychiatric diagnoses and triggers. This approach could somewhat biased estimates of their effect because they are likely to be underreported. This bias is more likely to affect the negative predictive value of the developed models. Third, I developed the risk assessment tool using data from one country, which limits its generalisability in the current form. The developed risk assessment tool should be recalibrated before being deployed in any population other than individuals given community sentences in Sweden.

5.8 Conclusion

Using data from the large population cohort of individuals given community sentences, I developed a simple, scalable tool for dynamic risk assessment of general and violent reoffending. The tool has good discrimination and moderate calibration performance for both outcomes. For high-risk individuals, the calibration performance is moderate. The developed tool can be used in the criminal justice system for risk monitoring during post-sentence supervision. In addition, I explored the effects of using risk assessment tools developed without accounting for offence-free time for dynamic risk assessment. Such tools accumulate substantial bias over time and should not be used without proper recalibration.

GENERAL DISCUSSION

Key findings

The thesis aimed to examine the role of psychiatric disorders as factors for criminal recidivism and mortality in individuals given community sentences and to create a simple, scalable tool for criminal recidivism risk monitoring in this population.

Chapter 1 identified the 2-year reconviction as the most reported criminal recidivism outcome. Subsequently, the 2-year reconviction, operationalized as 2-year general and violent reoffending, was used as a primary outcome in the cohort studies in Chapters 3 and 5. Chapter 1 additionally provided a review of criminal recidivism rates worldwide. The 2-year reconviction rates ranged from 14% to 43% in men and 9% to 35% in women given community sentences. The recidivism rates substantially varied across countries. The main reasons for their variation were the differences in the follow-up period, outcome operationalisation, and the inclusion of technical violations as recidivism outcomes.

Chapter 2 was a meta-analysis that summarised the existing research on risk factors for criminal recidivism in individuals given community sentences. The meta-analysis results highlighted the lack of published research into psychiatric disorders and criminal recidivism in community sentenced populations. Out of 15 identified studies, which reported data for 246,608 individuals, 8 examined substance use as a potential risk factor for recidivism (ORs ranged from 1.1 to 4.2), and only four examined other mental health factors (ORs ranged from 1.2 to 1.7). The strength of reported associations between mental health risk factors and criminal recidivism was comparable to that of non-modifiable risk factors, such as age, gender, and criminal history. However, the studies reporting these associations had major limitations. Most of them lacked diagnostic specificity, did not provide information about individual psychiatric diagnoses or were underpowered. Chapter 3 addressed these limitations.

Chapter 3 examined the association between psychiatric disorders and criminal recidivism in a Swedish population cohort of adult individuals given community sentences (N = 82,415), using register data. Substance use disorder was associated with an increased risk of general (HR = 1.90 [95% CI: 1.86-1.95]) and violent reoffending (HR = 1.71 [95% CI: 1.64-1.78]). Other psychiatric disorders were also associated with increased risk of general (HR = 1.27 [95% CI: 1.24-1.31]) and violent reoffending (HR = 1.47 [95% CI: 1.41-1.53]). Schizophrenia spectrum disorders, personality disorders and substance use disorders had stronger effects on violent reoffending than other psychiatric disorders. The associations between psychiatric diagnoses and reoffending persisted in individuals matched with their sentenced siblings without a known psychiatric diagnosis. Comorbid substance use disorder fully mediated the association between psychiatric disorders and general reoffending. However, comorbid substance use only partially mediated the association between psychiatric disorders and violent reoffending. The time-dependent analysis further corroborated this effect of comorbid substance use on general and violent reoffending. These results highlight the contribution of psychiatric disorders to violent reoffending beyond the effect of comorbid substance use. The relationship between psychiatric disorders and violence is likely polyetiological with each individual disorder representing an increased sensitivity to a separate subset of environmental triggers (Arseneault et al., 2000). Overall, alcohol and drug use disorders remain the primary mental health risk factors for criminal recidivism.

Chapter 4 additionally examined the association between psychiatric disorders and mortality in a Swedish total population cohort of adult individuals given community sentences (N = 109,751), using register data. During the follow-up, most potentially preventable deaths occurred in individuals with a psychiatric diagnosis. The proportion of preventable deaths was higher among younger individuals with a psychiatric diagnosis. The leading cause of death was suicide. Having substance use diagnosis at the time of a sentence was associated with the increased risk of all-cause (HR = 2.64 [95% CI: 2.51-2.79]) and external-cause mortality (HR = 3.66 [95% CI: 3.38-3.96]). Having any other psychiatric diagnosis at the time of a sentence was also associated with the increased risk of all-cause (HR = 1.72 [95% CI: 1.63-1.82]) and external-cause mortality (HR = 2.40 [95% CI: 2.22-2.59]). Receiving new substance use or another psychiatric diagnosis during the follow-up was associated with mortality in individuals without psychiatric diagnoses known at the start of a community sentence. The association between psychiatric disorders and mortality remained substantial even when controlling for comorbid substance use. Overall, substance use and other psychiatric disorders represent major risk factors for external-cause mortality, especially suicide.

Chapter 5 used pre-specified criminal, sociodemographic, and clinical risk factors to develop a dynamic prediction model for criminal recidivism in individuals under community supervision. The model was developed using dynamic prediction by landmarking (van Houwelingen, 2007). The prediction model accounts for adverse events during community supervision (triggers), changes in a supervised individual's circumstances, and offence-free time. The triggers include unintentional injury, being a victim of a violent assault, self-harm, psychiatric hospitalisation and substance intoxication. The model estimates the probability of general and violent reoffending within two years and can be used for ongoing risk monitoring. The model was deployed as an online dynamic risk assessment tool OxMore with good calibration and discrimination performance (c-index = 0.74 for violent reoffending, c-index = 0.69 for general reoffending). The developed prediction tool is available online (https://oxrisk.com/oxmore). As an important secondary outcome, the study demonstrated that actuarial recidivism risk assessment tools, which have not been developed as risk monitoring instruments but are used as such, will systematically overestimate the risk of recidivism over time.

The thesis emphasises the role of substance use and other psychiatric disorders as important targets for intervention in individuals given community sentences. Addressing psychiatric disorders in this population could potentially prevent many cases of reoffending and improve the long-term health outcomes decreasing the burden on correctional and healthcare systems.

Strength and limitations

The primary strength of the thesis is the focus on the application of robust and transparent statistical methods to high-quality data sources. Swedish population registers have positive predictive values of 85-95% for officially recorded medical and sociodemographic information that covers the total country population. The extensive coverage and high accuracy of Swedish register data ensure high validity and reliability of the thesis' findings. Chapters 3 and 4 report the currently largest studies of psychiatric disorders in individuals given community sentences that use official medical records. These studies also benefited from using time-dependent analysis and the sibling comparison design. The sibling comparison allowed accounting for the effect of the unmeasured familial confounding.

Chapter 5 introduces a novel application of dynamic prediction modelling by landmarking to criminal recidivism risk monitoring. The dynamic prediction model was derived and validated according to the best practices of prediction model development. The process was designed to prevent overfitting, mitigate possible effects of missing data, and produce generalisable results. The performance estimates of the resulting model were corrected for optimism, and the model's performance was compared to the performance of several competing models.

To ensure the high quality of the studies' design and presentation, all studies in the thesis were reported using the commonly accepted reporting guidelines. The systematic reviews followed PRISMA guidelines, the cohort studies followed STROBE guidelines, and the prediction model development study followed TRIPOD guidelines. The protocols for the systematic review (Chapter 1) and the meta-analysis (Chapter 3) were pre-registered in the PROSPERO database to minimise bias and ensure the transparency of the review process.

The thesis also has several limitations. Although sibling comparison and time-dependent analysis provide the basis for a cautious causal interpretation of the cohort studies' results, causal inferences in a strict sense would be significantly limited. The cohort studies relied on routinely collected observational data, not experimental data; thus, the relationships between exposures and outcomes were not directly controlled. In addition, when a complex relationship exists between exposure and measured covariates, including the interrelated covariates in the model can lead to over-adjustment and ultimately biased estimates (Schisterman et al., 2009). The simultaneous adjustment for multiple measured covariates could lead to biased estimates of the direct effects of psychiatric disorders in the multivariate analyses. Therefore, the results of multivariate analyses should be interpreted with caution.

Another limitation is that population registers do not contain data on potentially significant predictors of criminal recidivism that have been identified in published research, such as lack of social support and association with antisocial peers. In addition, the sociodemographic covariates are only recorded in population registers once a calendar year for all individuals in Sweden. This cross-sectional approach limits the validity of sociodemographic variables as dynamic predictors since they are unlikely to capture sudden changes in an individual's life circumstances.

The systematic review and the meta-analysis in Chapters 1 and 2 also had limitations. The identified studies had high heterogeneity of outcome operationalizations, which complicated the general interpretation of the reviews' results. It was often not possible to reliably estimate potential sources of heterogeneity. In Chapter 2, only a small number of identified studies reported the association between psychiatric disorders and criminal recidivism in individuals given community sentences. This lack of published research limited the generalisability of the meta-analysis' results.

Implications for policy and practice

Community sentences were shown to be effective in reducing recidivism (Tabar, Miravalle, Ronco, & Torrente, 2016) and in preventing further criminalization of offenders (Nieuwbeerta et al., 2009; Robinson & McNeill, 2015; Wermink et al., 2013). However, their blanket or insufficiently informed use can be counterproductive for achieving general goals of decreasing crime rates and prison population. In many countries, including the UK, the ‘toughened enforcement’ of community sentences created a backdoor path from community to prison for sentenced individuals (Heard, 2015). Frequently, community sentences are revoked because of technical violations, and more low-risk individuals end up in prison. This backdoor pathway contributes to the increase in the prison population in many countries and diminishes the potential rehabilitative effect of community sentences. The effect of such sentencing policies can be especially detrimental for individuals with psychiatric disorders given community sentences. Imprisonment can lead to the recurrence and exacerbation of symptoms (Hoke, 2015), which could be problematic given that access to mental health care in prisons is often limited. When released, individuals with psychiatric symptoms are at higher risk of committing a new offence and getting a new sentence (Z. Chang, Larsson, et al., 2015). Such vicious circles are likely to lead to the chronification of substance use and other psychiatric disorders in correctional populations. Overall, the failure to adequately address the mental health needs of sentenced individuals substantially limits the effectiveness of community sentences in reducing recidivism.

A community sentence is often the first contact point that an individual has with the correctional system. It follows that community supervision also represents an opportunity for early interventions. As I highlighted in this thesis, interventions that target psychiatric disorders, especially substance use, can significantly reduce repeat offending in individuals given community sentences. Policy-makers should focus on improving the accessibility of mental health interventions for individuals given community sentences. Besides improving accessibility and reach, it is also important to identify and address the barriers to participation in mental health programmes. The individual-level barriers can include lack of medical insurance, low literacy or difficulty understanding a local language, low health literacy (inability to recognise one's symptoms as non-normative), mistrust of healthcare providers, and fear of stigmatisation (Owens et al., 2011; Sirdifield & Brooker, 2020). The systemic barriers can include poor access to general medical care, the complexity of the healthcare landscape, and the lack of trained personnel in organisations providing services. Under-use of the community treatment requirements was also highlighted as a major barrier to mental health care access in sentenced individuals in the UK (Royal College of Psychiatrists, 2021; Sirdifield & Brooker, 2020). A one potentially promising approach that could mitigate many listed barriers is liaison and diversion services (Ryland et al., 2021; Scott et al., 2013). The trained mental health professionals work with the police and courts to identify individuals with specific needs at the early stages of legal system involvement. Liaison and diversion teams ensure the continuity of care, provide referrals and assist in the management of high-risk cases (Heard, 2016).

Improving mental health outcomes in individuals given community sentences will likely lead to lower mortality rates in this population. Substance use disorder should be the primary target for intervention as it accounts for the majority of potentially preventable deaths. The mental health programmes should also address suicidal behaviour as a separate intervention target since suicide is a leading cause of death in individuals given community sentences. Given many directions for potential interventions to reduce mortality, the use of a unified integrative approach could be beneficial. The specialised framework to reduce excessive mortality in individuals with severe psychiatric disorders is one such approach (Liu et al., 2017) It could be transformed and adapted for use in community sentences populations.

Overall, for all mental health interventions, the continuity of care has to be implemented beyond the post-sentence supervision period to minimise relapses (de Andrade et al., 2019; Sirdifield et al., 2019).

Another set of recommendations concerns the risk assessment of individuals with psychiatric disorders given community sentences. The periodical risk assessment with specialised tools enables consistent monitoring of an individual's risk levels. This monitoring allows a probation officer or mental health practitioner to respond to changing circumstances of their client quicker. The specialised prediction tools for risk monitoring have to be derived and validated specifically for this purpose. One example of a risk assessment tool suitable for continuous monitoring of criminal recidivism risk is OxMore, developed in the current thesis. Actuarial risk assessment tools that were not developed, validated, and calibrated for dynamic risk monitoring will tend to overestimate the risk of criminal recidivism in sentenced individuals with long offence-free time. Therefore, such instruments should only be used for baseline risk assessments, even if they incorporate dynamic risk factors.

I argue that the inclusion of offence-free time or other equivalent measures should be standard practice for risk assessment instruments. Including offence-free time improves the prediction accuracy of actuarial tools by correcting the risk overestimation. In addition, the risk assessment instruments with high predictive validity that also account for desistance promote more informed and responsive risk management.

I would also like to highlight the availability of new methods for prediction model development and deployment. For countries and territories with high informatisation levels, a potentially promising approach to increasing the predictive and ecological validity of criminal recidivism risk assessment tools is to use their data infrastructure fully. Modern approaches to prediction model development allow the risk assessment tools to be deployed as a continuous delivery application (Google, 2021). The retraining on the most recent data and regular republishing of the risk assessment formulas enable a faster reaction to any new developments in policy and practice while also retaining the interpretability of the models. Of course, it would be impossible to do without first establishing the data collection and storing infrastructure which is regularly updated and available for researchers.

Implication for research

Compared to the considerable body of research in custodial populations, the research into the mental health of individuals given community sentences is surprisingly limited. Convenience sampling may be one reason for this disproportionality in the number of published studies. Current and released prisoners are more closely supervised and likely constitute a more accessible population for researchers. Another reason is a higher public and governmental concern associated with released prisoners, which leads to more funding for corresponding research. However, given the results of this thesis, this lack of research in community sentenced populations seems problematic. There are several directions of potential inquiry that, in my opinion, could address this research gap.

The first direction of research is the association between substance use and other psychiatric disorders in the community sentenced population. The cohort studies from different countries using comparable outcomes are required to estimate the effect of psychiatric disorders on recidivism across legal and correctional settings. Another direction of research is the longitudinal studies of criminal behaviour in the general population. This type of research is possible to do with population registers. It can be used to examine the causal effect of psychiatric disorders on criminal behaviour while controlling for temporality and early childhood factors.

The third direction of research is the intervention studies in individuals given community sentences. In my opinion, the lack of studies on interventions that aim to reduce recidivism in community sentences populations is the most significant gap in the current research on the topic. In the UK, the increased use of mental health treatment requirements is considered a promising approach to reducing recidivism in individuals with psychiatric disorders given community sentences. Although there have been published studies indicating improved clinical outcomes in individuals receiving mental health requirements (Long, Dolley, & Hollin, 2018), no published studies or reports have examined the effectiveness of mental health treatment requirements in reducing recidivism. In the US, specialty mental health probation has demonstrated promising results in reducing violations and rearrests as well as improving clinical outcomes in individuals with psychiatric disorders (Skeem et al., 2017; Wolff et al., 2014). However, specialty mental health probation programmes vary widely in their implementation and are markedly understudied (van Deinse et al., 2019). The lack of major published studies could be one of the barriers to the wider implementation of these potentially promising approaches to managing individuals with psychiatric disorders given community sentences. Probation agencies and relevant state bodies should consider prioritising this line of research.

There is also a lack of research into intervention programmes for reducing self-harm in individuals with substance use. Such intervention programmes could benefit individuals given community sentences, considering the significant association between substance use and suicide in this population.

Another important line of potential investigation is examining the effects of adverse life events on the immediate criminal recidivism risk and other adverse outcomes, such as self-harm and suicide. Some researchers argue that predictive models of recidivism have reached their performance ceiling (Mills, 2017). Adding more predictors to the models does not increase their performance due to high collinearity between the predictors. Consequently, the most commonly used recidivism risk assessment instruments have comparable performance (Singh et al., 2010). The risk assessment instrument that I developed in this thesis is no exception. However, there might be significant room for improvement in the temporal precision. The risk monitoring instruments that account for sudden and clinically relevant changes in symptoms are used in general clinical practice to dynamically prioritise patient workload (Goodday et al., 2020). However, the research into the acute factors of recidivism is limited and needs to be expanded. The same line of reasoning can also be applied to the research of acute risk factors for the immediate risk of suicide and self-harm.

Conclusions

I reported on the role of psychiatric disorders as factors for criminal recidivism and mortality in individuals given community sentences and created a simple, scalable tool for criminal recidivism risk monitoring in this population. Substance use and other psychiatric disorders were associated with general and violent reoffending as well as all-cause and external-cause mortality. Given their high prevalence, psychiatric disorders should be considered important targets for intervention during post-sentence supervision. Mitigating the effects of substance use and other psychiatric disorders can lead to a lower number of new criminal offences and preventable deaths. The use of dynamic risk assessment tools developed specifically for risk monitoring may enable more informed and flexible risk management of individuals given community sentences, especially those with psychiatric disorders. By providing an extensive overview of the association between psychiatric disorders with criminal recidivism and mortality, this thesis hopes to encourage further research in community sentences populations, which are heavily understudied at the moment.

Future research in community sentenced populations should focus on improving comparability of recidivism studies, potentially effective interventions for substance use and psychiatric disorders, etiological pathways for violence and criminality, the effects of policies on recidivism and general clinical outcomes, acute dynamic risk factors for adverse outcomes, and factors promoting desistance from crime.

REFERENCES

Adams, S., Bostwick, L., & Campbell, R. (2011). *Examining Illinois probationer characteristics and outcomes*. Illinois Criminal Justice Information Authority Chicago.

Aebi, M. F., Hashimoto, Y. Z., & Tiago, M. M. (2019). *Probation and Prisons in Europe, 2018: Key Findings of the SPACE reports*. Council of Europe.

Allison, P. D. (2009). *Fixed effects regression models*. SAGE Publications.

Altman, D. G. (1990). *Practical statistics for medical research*. CRC press.

Andersen, S. N., & Skardhamar, T. (2017). Pick a number: Mapping recidivism measures and their consequences. *Crime & Delinquency*, *63*(5), 613–635.

Andrews, I., & Kasy, M. (2019). Identification of and correction for publication bias. *American Economic Review*, *109*(8), 2766–2794.

Arnau-Soler, A., Adams, M. J., Clarke, T.-K., MacIntyre, D. J., Milburn, K., Navrady, L., Generation Scotland, Consortium, M. D. D. W. G. of the P. G., Hayward, C., McIntosh, A., & Thomson, P. A. (2019). A validation of the diathesis-stress model for depression in Generation Scotland. *Translational Psychiatry*, *9*(1), 25. https://doi.org/10.1038/s41398-018-0356-7

Arseneault, L., Moffitt, T. E., Caspi, A., Taylor, P. J., & Silva, P. A. (2000). Mental disorders and violence in a total birth cohort: results from the Dunedin Study. *Archives of General Psychiatry*, *57*(10), 979–986. https://doi.org/10.1001/archpsyc.57.10.979

Bales, W. D., & Piquero, A. R. (2012). Assessing the impact of imprisonment on recidivism. *Journal of Experimental Criminology*, *8*(1), 71–101.

Barnard, J., & Rubin, D. B. (1999). Miscellanea. Small-sample degrees of freedom with multiple imputation. *Biometrika*, *86*(4), 948–955.

Bartels, L. (2009). The Weight of the Sword of Damocles: A Reconviction Analysis of Suspended Sentences in Tasmania. *Australian & New Zealand Journal of Criminology*, *42*(1), 72–100. https://doi.org/10.1375/acri.42.1.72

Bauermeister, S. D., & Gallacher, J. (2018). *The effect of childhood adversity on adult behavioural, psychological, cognitive and health outcomes: A UK Biobank and DPUK cross-cohort investigation*. https://doi.org/10.13140/RG.2.2.18859.54565

Benekos, P. J., & Merlo, A. V. (2014). *Controversies in juvenile justice and delinquency*. Routledge.

Bonta, J., & Andrews, D. A. (2007). Risk-need-responsivity model for offender assessment and rehabilitation. *Rehabilitation*, *6*(1), 1–22.

Bonta, J., Blais, J., & Wilson, H. A. (2014). A theoretically informed meta-analysis of the risk for general and violent recidivism for mentally disordered offenders. *Aggression and Violent Behavior*, *19*(3), 278–287.

Borschmann, R., Tibble, H., Spittal, M. J., Preen, D., Pirkis, J., Larney, S., Rosen, D. L., Young, J. T., Love, A. D., Altice, F. L., Binswanger, I. A., Bukten, A., Butler, T., Chang, Z., Chen, C.-Y., Clausen, T., Christensen, P. B., Culbert, G. J., Degenhardt, L., … Kinner, S. A. (2020). The Mortality After Release from Incarceration Consortium (MARIC): Protocol for a multi-national, individual participant data meta-analysis. *International Journal of Population Data Science*, *5*(1), 1145. https://doi.org/10.23889/ijpds.v5i1.1145

Bosly, S., Flore, D., Honhon, A., & Maggio, J. (2012). *Probation measures and alternative sanctions in the European Union*. Intersentia Publishing Limited.

Brennan, G. M., Hyde, L. W., & Baskin-Sommers, A. R. (2017). Antisocial pathways associated with substance use disorders: characterizing etiological underpinnings and implications for treatment. *Current Opinion in Behavioral Sciences*, *13*, 124–129. https://doi.org/https://doi.org/10.1016/j.cobeha.2016.11.014

Brooker, C., Sirdifield, C., Blizard, R., Denney, D., & Pluck, G. (2012). Probation and mental illness. *The Journal of Forensic Psychiatry & Psychology*, *23*(4), 522–537. https://doi.org/10.1080/14789949.2012.704640

Bushway, S. D., Nieuwbeerta, P., & Blokland, A. (2011). The predictive value of criminal background checks: Do age and criminal history affect time to redemption? *Criminology*, *49*(1), 27–60. https://doi.org/https://doi.org/10.1111/j.1745-9125.2010.00217.x

Canada, K. E., Trawver, K. R., & Barrenger, S. (2020). Deciding to participate in mental health court: Exploring participant perspectives. *International Journal of Law and Psychiatry*, *72*, 101628. https://doi.org/https://doi.org/10.1016/j.ijlp.2020.101628

Caudy, M. S., Tillyer, M. S., & Tillyer, R. (2018). Jail Versus Probation: A Gender-Specific Test of Differential Effectiveness and Moderators of Sanction Effects. *Criminal Justice and Behavior*, *45*(7), 949–968. https://doi.org/10.1177/0093854818766375

Central Statistics Office. (2016). *Probation recidivism 2010 cohort*. http://www.cso.ie/en/releasesandpublications/er/pror/probationrecidivism2010cohort/

Chang, C.-K., Hayes, R. D., Broadbent, M., Fernandes, A. C., Lee, W., Hotopf, M., & Stewart, R. (2010). All-cause mortality among people with serious mental illness (SMI), substance use disorders, and depressive disorders in southeast London: a cohort study. *BMC Psychiatry*, *10*, 77. https://doi.org/10.1186/1471-244X-10-77

Chang, Z., Larsson, H., Lichtenstein, P., & Fazel, S. (2015). Psychiatric disorders and violent reoffending: a national cohort study of convicted prisoners in Sweden. *The Lancet. Psychiatry*, *2*(10), 891–900. https://doi.org/10.1016/S2215-0366(15)00234-5

Chang, Z., Lichtenstein, P., Larsson, H., & Fazel, S. (2015). Substance use disorders, psychiatric disorders, and mortality after release from prison: a nationwide longitudinal cohort study. *The Lancet. Psychiatry*, *2*(5), 422–430. https://doi.org/10.1016/S2215-0366(15)00088-7

Chasiropoulou, C., Siouti, N., Mougiakos, T., & Dimitrakopoulos, S. (2019). The diathesis-stress model in the emergence of major psychiatric disorders during military service. *Psychiatrike = Psychiatriki*, *30*(4), 291–298. https://doi.org/10.22365/jpsych.2019.304.291

Chesney, E., Goodwin, G. M., & Fazel, S. (2014). Risks of all-cause and suicide mortality in mental disorders: a meta-review. *World Psychiatry*, *13*(2), 153–160. https://doi.org/https://doi.org/10.1002/wps.20128

Clark, V. A. (2014). Predicting Two Types of Recidivism Among Newly Released Prisoners: First Addresses as “Launch Pads” for Recidivism or Reentry Success. *Crime & Delinquency*, *62*(10), 1364–1400. https://doi.org/10.1177/0011128714555760

Clarke, M. C., Peterson-Badali, M., & Skilling, T. A. (2017). The Relationship Between Changes in Dynamic Risk Factors and the Predictive Validity of Risk Assessments Among Youth Offenders. *Criminal Justice and Behavior*, *44*(10), 1340–1355. https://doi.org/10.1177/0093854817719915

Collins, G. S., de Groot, J. A., Dutton, S., Omar, O., Shanyinde, M., Tajar, A., Voysey, M., Wharton, R., Yu, L.-M., Moons, K. G., & Altman, D. G. (2014). External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Medical Research Methodology*, *14*(1), 40. https://doi.org/10.1186/1471-2288-14-40

Commons Chamber. (2020). *Sentencing White Paper (parliament debate)*. https://www.parallelparliament.co.uk/debate/commons/2020-09-16/debates/2FEDDE2C-FF94-424A-8B97-7CBA9760C5EB/SentencingWhitePaper

Dahlqwist, E., & Sjolander, A. (2019). *Package ‘AF.’* https://cran.r-project.org/web/packages/AF/AF.pdf

de Andrade, D., Elphinston, R. A., Quinn, C., Allan, J., & Hides, L. (2019). The effectiveness of residential treatment services for individuals with substance use disorders: A systematic review. *Drug and Alcohol Dependence*, *201*, 227–235. https://doi.org/https://doi.org/10.1016/j.drugalcdep.2019.03.031

de Vogel, V., de Vries Robbé, M., de Ruiter, C., & Bouman, Y. H. A. (2011). Assessing protective factors in forensic psychiatric practice: Introducing the SAPROF. *International Journal of Forensic Mental Health*, *10*(3), 171–177.

Dean, K., Singh, S., & Soon, Y.-L. (2020). Decriminalizing severe mental illness by reducing risk of contact with the criminal justice system, including for forensic patients. *CNS Spectrums*, *25*(5), 687–700. https://doi.org/DOI: 10.1017/S109285292000125X

Department of Justice. (2011). *Adult Reconviction in Northern Ireland 2005*. https://www.justice-ni.gov.uk/sites/default/files/publications/doj/adult-reconviction-in-northern-ireland-2005.pdf

D'Onofrio, B. M., Turkheimer, E., Emery, R. E., Slutske, W. S., Heath, A. C., Madden, P. A., & Martin, N. G. (2006). A genetically informed study of the processes underlying the association between parental marital instability and offspring adjustment. *Developmental Psychology*, *42*(3), 486.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj*, *315*(7109), 629–634.

Ersche, K. D., Turton, A. J., Chamberlain, S. R., Müller, U., Bullmore, E. T., & Robbins, T. W. (2012). Cognitive dysfunction and anxious-impulsive personality traits are endophenotypes for drug dependence. *The American Journal of Psychiatry*, *169*(9), 926–936. https://doi.org/10.1176/appi.ajp.2012.11091421

European Monitoring Centre for Drugs and Drug Addiction. (2017). *Drug trafficking penalties across the European Union: a survey of expert opinion*. https://www.emcdda.europa.eu/publications/technical-reports/trafficking-penalties_en

Evans, E., Li, L., Urada, D., & Anglin, M. D. (2014). Comparative Effectiveness of California's Proposition 36 and Drug Court Programs Before and After Propensity Score Matching. *Crime and Delinquency*, *60*(6), 909–938. https://doi.org/10.1177/0011128710382342

Fazel, S., Chang, Z., Fanshawe, T., Långström, N., Lichtenstein, P., Larsson, H., & Mallett, S. (2016). Prediction of violent reoffending on release from prison: derivation and external validation of a scalable tool. *The Lancet Psychiatry*, *3*(6), 535–543. https://doi.org/10.1016/S2215-0366(16)00103-6

Fazel, S., Gulati, G., Linsell, L., Geddes, J. R., & Grann, M. (2009). Schizophrenia and violence: systematic review and meta-analysis. *PLoS Medicine*, *6*(8), e1000120. https://doi.org/10.1371/journal.pmed.1000120

Fazel, S., & Runeson, B. (2020). Suicide. *New England Journal of Medicine*, *382*(3), 266–274. https://doi.org/10.1056/NEJMra1902944

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24 827 people: systematic review and meta-analysis. *BMJ : British Medical Journal*, *345*, e4692. https://doi.org/10.1136/bmj.e4692

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

Fazel, S., Wolf, A., Larsson, H., Mallett, S., & Fanshawe, T. R. (2019). The prediction of suicide in severe mental illness: development and validation of a clinical prediction rule (OxMIS). *Translational Psychiatry*, *9*(1), 1–10.

Fazel, S., Wolf, A., Vazquez-Montes, M. D. L. A., & Fanshawe, T. R. (2019). Prediction of violent reoffending in prisoners and individuals on probation: a Dutch validation study (OxRec). *Scientific Reports*, *9*(1), 1–9.

Fazel, S., Wolf, A., & Yukhnenko, D. (2019). Recidivism reporting checklist. *Open Science Framework*, *12*, 10–17605.

Federal Penitentiary Service of Russia. (2018). *Characteristics of persons registered in penitentiary inspectorates [Характеристика лиц, состоящих на учёте в уголовно-исполнительных инспекциях]*. http://fsin.su/structure/inspector/iao/statistika/Xar-ka v YII/

Federal Statistical Office. (2015). *Statistique des condamnations pénales 1984–2014*. https://www.bfs.admin.ch/bfs/en/home/statistics/crime-criminal-justice/recidivism/analysis.assetdetail.350339.html

Figma Design. (2017). Figma: the collaborative interface design tool. *Retrieved September*, *17*, 2017.

Fink, S. A., & Brown Jr, R. S. (2006). Survival Analysis. *Gastroenterology & Hepatology*, *2*(5), 380–383.

Fisher, D. (2019). *ADMETAN: Stata module to provide comprehensive meta-analysis*.

Fitton, L., Yu, R., & Fazel, S. (2020). Childhood Maltreatment and Violent Outcomes: A Systematic Review and Meta-Analysis of Prospective Studies. *Trauma, Violence & Abuse*, *21*(4), 754–768. https://doi.org/10.1177/1524838018795269

Flinchum, T., Hevener, H., Hall, M., & Wesoloski, J. (2016). *Correctional program evaluation: offenders placed on probation or released from prison in FY 2013.*

Flores, A. W., Holsinger, A. M., Lowenkamp, C. T., & Cohen, T. H. (2016). Time-Free Effects in Predicting Recidivism Using Both Fixed and Variable Follow-Up Periods: Do Different Methods Produce Different Results. *Criminal Justice and Behavior*, *44*(1), 121–137. https://doi.org/10.1177/0093854816678649

Flores, A. W., Holsinger, A. M., Lowenkamp, C. T., & Cohen, T. H. (2017). Time-free effects in predicting recidivism using both fixed and variable follow-up periods: Do different methods produce different results. *Criminal Justice and Behavior*, *44*(1), 121–137.

Gelsthorpe, L., Padfield, N., & Phillips, J. (2012). *Deaths on probation: an analysis of data regarding people dying under probation supervision; a report for the Howard league for penal reform*. https://howardleague.org/wp-content/uploads/2016/05/Deaths-on-probation.pdf

Gendreau, P., Little, T., & Goggin, C. (1996). A meta-analysis of the predictors of adult offender recidivism: What works! *Criminology*, *34*(4), 575–608.

Goodday, S. M., Atkinson, L., Goodwin, G., Saunders, K., South, M., Mackay, C., Denis, M., Hinds, C., Attenburrow, M.-J., Davies, J., Welch, J., Stevens, W., Mansfield, K., Suvilehto, J., & Geddes, J. (2020). The True Colours Remote Symptom Monitoring System: A Decade of Evolution. *J Med Internet Res*, *22*(1), e15188. https://doi.org/10.2196/15188

Google. (2021). *MLOps: Continuous delivery and automation pipelines in machine learning*. https://cloud.google.com/architecture/mlops-continuous-delivery-and-automation-pipelines-in-machine-learning#continuous_delivery

Graham, L., Fischbacher, C. M., Stockton, D., Fraser, A., Fleming, M., & Greig, K. (2015). Understanding extreme mortality among prisoners: a national cohort study in Scotland using data linkage. *European Journal of Public Health*, *25*(5), 879–885. https://doi.org/10.1093/eurpub/cku252

Grann, M., Danesh, J., & Fazel, S. (2008). The association between psychiatric diagnosis and violent re-offending in adult offenders in the community. *BMC Psychiatry*, *8*, 92. https://doi.org/10.1186/1471-244X-8-92

Grzymala-Busse, A. (2011). Time will tell? Temporality and the analysis of causal mechanisms and processes. *Comparative Political Studies*, *44*(9), 1267–1297.

Hanson, R. K. (2018). Long-term recidivism studies show that desistance is the norm. *Criminal Justice and Behavior*, *45*(9), 1340–1346.

Hanson, R. K., Harris, A. J. R., Letourneau, E., Helmus, L. M., & Thornton, D. (2018). Reductions in risk based on time offense-free in the community: Once a sexual offender, not always a sexual offender. *Psychology, Public Policy, and Law*, *24*(1), 48.

Hanson, R. K., & Morton-Bourgon, K. E. (2005). The characteristics of persistent sexual offenders: a meta-analysis of recidivism studies. *Journal of Consulting and Clinical Psychology*, *73*(6), 1154.

Harding, D. J., Morenoff, J. D., Nguyen, A. P., & Bushway, S. D. (2017). Short- and long-term effects of imprisonment on future felony convictions and prison admissions. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(42), 11103–11108. https://doi.org/10.1073/pnas.1701544114

Harrell, F. E. J., Lee, K. L., & Mark, D. B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, *15*(4), 361–387. https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4

Harris, P. M. (2011). The first-time adult-onset offender: findings from a community corrections cohort. *International Journal of Offender Therapy and Comparative Criminology*, *55*(6), 949–981. https://doi.org/10.1177/0306624X10372110

Hazewinkel, A.-D. (2018). *Prediction models with survival data: a comparison between machine learning and a Cox proportional hazard regression model*. https://www.universiteitleiden.nl/binaries/content/assets/science/mi/scripties/statscience/2018-2019/hazewinkel_master_thesis.pdf

Heard, C. (2015). *Community sentences since 2000: How they work–and why they have not cut prisoner numbers*. https://www.crimeandjustice.org.uk/publications/community-sentences-2000

Heard, C. (2016). *Alternatives to imprisonment in Europe: A handbook of good practice*. http://www.prisonobservatory.org/upload/Good%20practice%20handbook%20AS.pdf

Heinze, G., & Dunkler, D. (2017). Five myths about variable selection. *Transplant International : Official Journal of the European Society for Organ Transplantation*, *30*(1), 6–10. https://doi.org/10.1111/tri.12895

Heinze, G., Wallisch, C., & Dunkler, D. (2018). Variable selection – A review and recommendations for the practicing statistician. *Biometrical Journal*, *60*(3), 431–449. https://doi.org/https://doi.org/10.1002/bimj.201700067

Hendricks, P. S., Crawford, M. S., Cropsey, K. L., Copes, H., Sweat, N. W., Walsh, Z., & Pavela, G. (2017). The relationships of classic psychedelic use with criminal behavior in the United States adult population. *Journal of Psychopharmacology*, *32*(1), 37–48. https://doi.org/10.1177/0269881117735685

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (2019). *Cochrane handbook for systematic reviews of interventions*. John Wiley & Sons.

Hoke, S. (2015). Mental illness and prisoners: Concerns for communities and healthcare providers. *Online Journal of Issues in Nursing*, *20*(1).

Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software*, *45*(1), 1–47.

Howard, P. D., & Dixon, L. (2012). The Construction and Validation of the OASys Violence Predictor: Advancing Violence Risk Assessment in the English and Welsh Correctional Services. *Criminal Justice and Behavior*, *39*(3), 287–307. https://doi.org/10.1177/0093854811431239

Howerton, A., Byng, R., Campbell, J., Hess, D., Owens, C., & Aitken, P. (2007). Understanding help seeking behaviour among male offenders: qualitative interview study. *BMJ*, *334*(7588), 303. https://doi.org/10.1136/bmj.39059.594444.AE

Huebner, B. M., & Cobbina, J. (2007). The Effect of Drug Use, Drug Treatment Participation, and Treatment Completion on Probationer Recidivism. *Journal of Drug Issues*, *37*(3), 619–641. https://doi.org/10.1177/002204260703700307

Humphrey, J. A., Burford, G., & Dye, M. H. (2012). A longitudinal analysis of reparative probation and recidivism. *Criminal Justice Studies*, *25*(2), 117–130.

Hyatt, J. M., & Barnes, G. C. (2017). An experimental evaluation of the impact of intensive supervision on the recidivism of high-risk probationers. *Crime & Delinquency*, *63*(1), 3–38.

Iba, K., Shinozaki, T., Maruo, K., & Noma, H. (2021). Re-evaluation of the comparative effectiveness of bootstrap-based optimism correction methods in the development of multivariable clinical prediction models. *BMC Medical Research Methodology*, *21*(1), 9. https://doi.org/10.1186/s12874-020-01201-w

IBM. (2021). *Cox Regression Analysis*. https://www.ibm.com/docs/en/spss-statistics/24.0.0?topic=option-cox-regression-analysis

Jacobs, L. A., & Gottlieb, A. (2020). The Effect of Housing Circumstances on Recidivism: Evidence From a Sample of People on Probation in San Francisco. *Criminal Justice and Behavior*, *47*(9), 1097–1115. https://doi.org/10.1177/0093854820942285

Jolliffe, D., & Hedderman, C. (2015). Investigating the impact of custody on reoffending using propensity score matching. *Crime & Delinquency*, *61*(8), 1051–1077.

Karlsson, L. (2016). *An Evaluation of Methods for Assessing the Functional Form of Covariates in the Cox Model* [Master Thesis, Uppsala University]. http://www.diva-portal.org/smash/get/diva2:942232/FULLTEXT01.pdf

Kassambara, A., Kosinski, M., Biecek, P., & Fabian, S. (2017). Package ‘survminer.’ *Drawing Survival Curves Using ‘Ggplot2’(R Package Version 03 1)*.

Keaney, F., Gossop, M., Dimech, A., Guerrini, I., Butterworth, M., Al-Hassani, H., & Morinan, A. (2011). Physical health problems among patients seeking treatment for substance use disorders: A comparison of drug dependent and alcohol dependent patients. *Journal of Substance Use*, *16*(1), 27–37. https://doi.org/10.3109/14659890903580474

Khalifeh, H., Johnson, S., Howard, L. M., Borschmann, R., Osborn, D., Dean, K., Hart, C., Hogg, J., & Moran, P. (2015). Violent and non-violent crime against adults with severe mental illness. *British Journal of Psychiatry*, *206*(4), 275–282. https://doi.org/DOI: 10.1192/bjp.bp.114.147843

Ķipēna, K., Zavackis, A., & Ņikišins, J. (2013). Sodu izcietušo personu noziedzīgo nodarījumu recidīvs. *Jurista Vārds.*, *35*(786), 12–17.

Kriminalvården. (2020). *Kriminalvård och Statistik*. https://www.kriminalvarden.se/globalassets/publikationer/kartlaggningar-och-utvarderingar/kos-2020---kriminalvard-och-statistik.pdf

Lahey, B. B., & D'Onofrio, B. M. (2010). All in the family: Comparing siblings to test causal hypotheses regarding environmental influences on behavior. *Current Directions in Psychological Science*, *19*(5), 319–323.

Landenberger, N. A., & Lipsey, M. W. (2005). The positive effects of cognitive–behavioral programs for offenders: A meta-analysis of factors associated with effective treatment. *Journal of Experimental Criminology*, *1*(4), 451–476.

Lapointe-Shaw, L., Bouck, Z., Howell, N. A., Lange, T., Orchanian-Cheff, A., Austin, P. C., Ivers, N. M., Redelmeier, D. A., & Bell, C. M. (2018). Mediation analysis with a time-to-event outcome: a review of use and reporting in healthcare research. *BMC Medical Research Methodology*, *18*(1), 118. https://doi.org/10.1186/s12874-018-0578-7

Lawrence, D., Kisely, S., & Pais, J. (2010). The Epidemiology of Excess Mortality in People with Mental Illness. *The Canadian Journal of Psychiatry*, *55*(12), 752–760. https://doi.org/10.1177/070674371005501202

Leonardi, F. (2007). Le misure alternative alla detenzione tra reinserimento sociale e abbattimento della recidiva. *Rassegna Penitenziaria e Criminologica*, *2*(7). http://www.rassegnapenitenziaria.it/cop/4825.pdf

Liu, N. H., Daumit, G. L., Dua, T., Aquila, R., Charlson, F., Cuijpers, P., Druss, B., Dudek, K., Freeman, M., Fujii, C., Gaebel, W., Hegerl, U., Levav, I., Munk Laursen, T., Ma, H., Maj, M., Elena Medina-Mora, M., Nordentoft, M., Prabhakaran, D., … Saxena, S. (2017). Excess mortality in persons with severe mental disorders: a multilevel intervention framework and priorities for clinical practice, policy and research agendas. *World Psychiatry : Official Journal of the*

World Psychiatric Association (WPA), *16*(1), 30–40. https://doi.org/10.1002/wps.20384

Liu, S., Tian, L., Lee, S., & Xie, M. (2018). Exact inference on meta-analysis with generalized fixed-effects and random-effects models. *Biostatistics & Epidemiology*, *2*(1), 1–22.

Long, C. G., Dolley, O., & Hollin, C. (2018). The use of the mental health treatment requirement (MHTR): clinical outcomes at one year of a collaboration. *Journal of Criminal Psychology*, *8*(3), 215–233. https://doi.org/10.1108/JCP-01-2018-0003

Lowenkamp, C. T., Holsinger, A., Robinson, C. R., & Alexander, M. (2014). Diminishing or durable treatment effects of STARR? A research note on 24-month re-arrest rates. *Journal of Crime and Justice*, *37*(2), 275–283.

Ludvigsson, J. F., Andersson, E., Ekbom, A., Feychting, M., Kim, J.-L., Reuterwall, C., Heurgren, M., & Olausson, P. O. (2011). External review and validation of the Swedish national inpatient register. *BMC Public Health*, *11*(1), 450. https://doi.org/10.1186/1471-2458-11-450

Ludvigsson, J. F., Otterblad-Olausson, P., Pettersson, B. U., & Ekbom, A. (2009). The Swedish personal identity number: possibilities and pitfalls in healthcare and medical research. *European Journal of Epidemiology*, *24*(11), 659–667. https://doi.org/10.1007/s10654-009-9350-y

Ludvigsson, J. F., Svedberg, P., Olén, O., Bruze, G., & Neovius, M. (2019). The longitudinal integrated database for health insurance and labour market studies (LISA) and its use in medical research. *European Journal of Epidemiology*, *34*(4), 423–437. https://doi.org/10.1007/s10654-019-00511-8

Lurigio, A. J., Cho, Y. I., Swartz, J. A., Johnson, T. P., Graf, I., & Pickup, L. (2003). Standardized assessment of substance-related, other psychiatric, and comorbid disorders among probationers. *International Journal of Offender Therapy and Comparative Criminology*, *47*, 630–652.

Mackenzie, J. C., Cartwright, T., & Borrill, J. (2017). Exploring suicidal behaviours by probation clients—a qualitative near-lethal study. *Journal of Public Health*, *40*(1), 146–153. https://doi.org/10.1093/pubmed/fdx005

Maliek, N. A. (2017). *An Empirical Assesment of the Direct and Indirect Effects of Mental Health Disorders on Probation Outcomes* [Master Thesis]. The University of Texas at San Antonio.

Mann, R., & Bermingham, R. (2020). *Non-custodial sentences (research briefing)*. https://post.parliament.uk/research-briefings/post-pn-0613/

Marriam-Webster. (2021). *Dictionary*. https://www.merriam-webster.com/dictionary

Maruna, S., & Mann, R. (2019). Reconciling 'desistance'and 'what works.' *Academic Insights*, *1*, 3–10.

Mews, A., Hillier, J., McHugh, M., & Coxon, C. (2015). *The impact of short custodial sentences, community orders and suspended sentence orders on re-offending*. https://www.bl.uk/britishlibrary/~/media/bl/global/social-welfare/pdfs/non-secure/i/m/p/impact-of-short-custodial-sentences-community-orders-and-suspended-sentence-orders-on-reoffending.pdf

Mills, J. F. (2017). Violence risk assessment: A brief review, current issues, and future directions. *Canadian Psychology/Psychologie Canadienne*, *58*(1), 40.

Ministère de la Sécurité publique. (2015). *Projet: Enquête sur la récidive/reprise de la clientèle confiée aux Services correctionnels du Québec*. https://www.securitepublique.gouv.qc.ca/services-correctionnels/publications-et-statistiques/enquete-sur-la-recidivereprise.html

Ministry of Justice. (2018a). *Proven reoffending statistics quarterly: January 2016 to March 2016.* https://www.gov.uk/government/statistics/proven-reoffending-statistics-january-2016-to-march-2016

Ministry of Justice. (2018b). *Vulnerable offenders steered towards treatment (press release)*. https://www.gov.uk/government/news/vulnerable-offenders-steered-towards-treatment

Ministry of Justice. (2020a). *Criminal justice system statistics quarterly: March 2020*. https://www.gov.uk/government/statistics/criminal-justice-system-statistics-quarterly-march-2020

Ministry of Justice. (2020b). *Deaths of Offenders in the Community, England and Wales, 2019/20*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/att achment_data/file/980144/Deaths_of_Offenders_in_the_Community_2019-20_bulletin.pdf

Ministry of Justice. (2021a). *Criminal justice system statistics quarterly: December 2020*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/987892/criminal-justice-statistics-dec-2020.pdf

Ministry of Justice. (2021b). *Offender management statistics quarterly: October to December 2020 and annual 2020*. https://www.gov.uk/government/statistics/offender-management-statistics-quarterly-october-to-december-2020/offender-management-statistics-quarterly-october-to-december-2020-and-annual-2020--2

Ministry of Justice. (2021c). *Strengthening probation, building confidence*. https://www.gov.uk/guidance/strengthening-probation-building-confidence

Minor, K. I., Wells, J. B., & Sims, C. (2003). Recidivism among federal probationers-predicting sentence violations. *Fed. Probation*, *67*, 31.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *BMJ (Clinical Research Ed.)*, *339*, b2535. https://doi.org/10.1136/bmj.b2535

Molero, Y., Zetterqvist, J., Binswanger, I. A., Hellner, C., Larsson, H., & Fazel, S. (2018). Medications for Alcohol and Opioid Use Disorders and Risk of Suicidal Behavior, Accidental Overdoses, and Crime. *American Journal of Psychiatry*, *175*(10), 970–978. https://doi.org/10.1176/appi.ajp.2018.17101112

Nagin, D. S., & Snodgrass, G. M. (2013). The effect of incarceration on re-offending: Evidence from a natural experiment in Pennsylvania. *Journal of Quantitative Criminology*, *29*(4), 601–642.

National Cancer Institute. (2021). *NCI's Dictionary of Cancer Terms*. https://www.cancer.gov/publications/dictionaries/cancer-terms

National Centre for Suicide Research and Prevention. (2021). *Suicide in Sweden*. https://ki.se/en/nasp/suicide-in-sweden

National Institute of Mental Health. (2021). *Mental Illness*. https://www.nimh.nih.gov/health/statistics/mental-illness

Nieuwbeerta, P., Nagin, D. S., & Blokland, A. A. J. (2009). Assessing the impact of first-time imprisonment on offenders' subsequent criminal career development: A matched samples comparison. *Journal of Quantitative Criminology*, *25*(3), 227–257.

Norman, E., Starkey, N., & Polaschek, D. (2021). The association between self-reported traumatic brain injury, neuropsychological function, and compliance among people serving community sentences. *Brain Impairment*, 1–17. https://doi.org/10.1017/BrImp.2021.15

North Carolina Sentencing & Advisory Commission. (2018). *Correctional program evaluation. Offenders placed on probation or released from prison in FY 2015*. https://www.nccourts.gov/assets/documents/publications/recidivism_2018.pdf

O'Connor, R. C., Rasmussen, S., & Hawton, K. (2010). Predicting depression, anxiety and self-harm in adolescents: The role of perfectionism and acute life stress. *Behaviour Research and Therapy*, *48*(1), 52–59. https://doi.org/https://doi.org/10.1016/j.brat.2009.09.008

Olson, D. E., Alderden, M., & Lurigio, A. J. (2003). Men are from Mars, women are from Venus, but what role does gender play in probation recidivism? *Justice Research and Policy*, *5*(2), 33–54.

Olson, D. E., & Lurigio, A. J. (2000). Predicting Probation Outcomes: Factors Associated with Probation Rearrest, Revocations, and Technical Violations during Supervision. *Justice Research and Policy*, *2*(1), 73–86. https://doi.org/10.3818/JRP.2.1.2000.73

Olsson, M. O., Öjehagen, A., Brådvik, L., & Håkansson, A. (2015). Predictors of Psychiatric Hospitalization in Ex-Prisoners With Substance Use Problems: A Data-Linkage Study. *Journal of Drug Issues*, *45*(2), 202–213. https://doi.org/10.1177/0022042615575374

Olver, M. E., Stockdale, K. C., & Wormith, J. S. (2014). Thirty years of research on the Level of Service Scales: A meta-analytic examination of predictive accuracy and sources of variability. *Psychological Assessment*, *26*(1), 156.

Oram, S., Trevillion, K., Khalifeh, H., Feder, G., & Howard, L. M. (2014). Systematic review and meta-analysis of psychiatric disorder and the perpetration of partner violence. *Epidemiology and Psychiatric Sciences*, *23*(4), 361–376. https://doi.org/10.1017/S2045796013000450

Owens, G. P., Rogers, S. M., & Whitesell, A. A. (2011). Use of Mental Health Services and Barriers to Care for Individuals on Probation or Parole. *Journal of Offender Rehabilitation*, *50*(1), 37–47. https://doi.org/10.1080/10509674.2011.536721

Pearson, D. A. S., McDougall, C., Kanaan, M., Torgerson, D. J., & Bowles, R. A. (2014). Evaluation of the Citizenship Evidence-Based Probation Supervision Program Using a Stepped Wedge Cluster Randomized Controlled Trial. *Crime & Delinquency*, *62*(7), 899–924. https://doi.org/10.1177/0011128714530824

Peillard, A. M. M., Correa, N. M., Chahuán, G. W., & Lacoa, J. F. (2012). La reincidencia en el sistema penitenciario chileno. *Santiago*. https://pazciudadana.cl/biblioteca/documentos/la-reincidencia-en-el-sistema-penitenciario-chileno/

Petersen, A. H., & Lange, T. (2020). What Is the Causal Interpretation of Sibling Comparison Designs? *Epidemiology (Cambridge, Mass.)*, *31*(1), 75–81. https://doi.org/10.1097/EDE.0000000000001108

Peterson, J., Welch, V., Losos, M., & Tugwell, P. (2011). The Newcastle-Ottawa scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. *Ottawa: Ottawa Hospital Research Institute*, 1–12.

Phillips, J., Gelsthorpe, L., & Padfield, N. (2017). Non-custodial deaths: Missing, ignored or unimportant? *Criminology & Criminal Justice*, *19*(2), 160–178. https://doi.org/10.1177/1748895817745939

Phillips, J., Padfield, N., & Gelsthorpe, L. (2018). Suicide and community justice. *Health & Justice*, *6*(1), 14. https://doi.org/10.1186/s40352-018-0072-7

Porporino, F. J. (2018). Developments and challenges in probation practice: Is there a way forward for establishing effective and sustainable probation systems? *European Journal of Probation*, *10*(1), 76–95. https://doi.org/10.1177/2066220318764713

Putter, H., & Putter, M. H. (2011). *Package ‘dynpred.’* https://cran.r-project.org/web/packages/dynpred/dynpred.pdf

Putter, H., & van Houwelingen, H. C. (2017). Understanding Landmarking and Its Relation with Time-Dependent Cox Regression. *Statistics in Biosciences*, *9*(2), 489–503. https://doi.org/10.1007/s12561-016-9157-9

R Core Team. (2013). *R: A language and environment for statistical computing*.

Rasmussen, L. J. H., Moffitt, T. E., Arseneault, L., Danese, A., Eugen-Olsen, J., Fisher, H. L., Harrington, H., Houts, R., Matthews, T., Sugden, K., Williams, B., & Caspi, A. (2020). Association of Adverse Experiences and Exposure to Violence in Childhood and Adolescence With Inflammatory Burden in Young People. *JAMA Pediatrics*, *174*(1), 38–47. https://doi.org/10.1001/jamapediatrics.2019.3875

Reising, K. (2021). *Criminal Offending and Mental Disorders: Long-term bidirectional and intergenerational effects between mental health problems and offending behaviour*. University of Cambridge.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*(1), 1–8.

Robinson, G., & McNeill, F. (2015). *Community punishment: European perspectives*. Routledge.

Rosenthal, R., & DiMatteo, M. R. (2001). Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, *52*(1), 59–82.

Royal College of Psychiatrists. (2021). *Mental health treatment requirements*. https://www.rcpsych.ac.uk/docs/default-source/improving-care/better-mh-policy/position-statements/ps04_21---mental-health-treatment-requirements.pdf

RStudio Team. (2020). *RStudio: Integrated Development for R*. http://www.rstudio.com/

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, *1*(5), 206–215.

Ryan, C., Nielssen, O., Paton, M., & Large, M. (2010). Clinical Decisions in Psychiatry Should Not Be Based On Risk Assessment. *Australasian Psychiatry*, *18*(5), 398–403. https://doi.org/10.3109/10398562.2010.507816

Ryland, H., Forrester, A., Exworthy, T., Gallagher, S., Ramsay, L., & Khan, A. A. (2021). Liaison and diversion services in South East London: Referral patterns over a 25-year period. *Medico-Legal Journal*, *89*(3), 166–172. https://doi.org/10.1177/00258172211010558

Saha, S., Chant, D., Welham, J., & McGrath, J. (2005). A Systematic Review of the Prevalence of Schizophrenia. *PLOS Medicine*, *2*(5), e141.

Sampson, R. J., & Laub, J. H. (1997). A life-course theory of cumulative disadvantage and the stability of delinquency. *Developmental Theories of Crime and Delinquency*, *7*, 133–161.

Sampson, R. J., & Laub, J. H. (2005). A Life-Course View of the Development of Crime. *The ANNALS of the American Academy of Political and Social Science*, *602*(1), 12–45. https://doi.org/10.1177/0002716205280075

Sariaslan, A., Lichtenstein, P., Larsson, H., & Fazel, S. (2016). Triggers for Violent Criminality in Patients With Psychotic Disorders. *JAMA Psychiatry*, *73*(8), 796–803. https://doi.org/10.1001/jamapsychiatry.2016.1349

SAS Institute Inc. (2013). *SAS 9.4*. SAS Institute Inc.

Sattar, G. (2001). *Rates and causes of death among prisoners and offenders under community supervision*. Citeseer. https://www.ojp.gov/ncjrs/virtual-library/abstracts/rates-and-causes-death-among-prisoners-and-offenders-under

Schisterman, E. F., Cole, S. R., & Platt, R. W. (2009). Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology (Cambridge, Mass.)*, *20*(4), 488.

Schweder, A. E. (2003). *Predictors of deliquent and aggressive behavior problems in maltreated children: A test of the diathesis-stress model*. Yale University.

Scott, D. A., McGilloway, S., Dempster, M., Browne, F., & Donnelly, M. (2013). Effectiveness of Criminal Justice Liaison and Diversion Services for Offenders With Mental Disorders: A Review. *Psychiatric Services*, *64*(9), 843–849. https://doi.org/10.1176/appi.ps.201200144

Selzam, S., Coleman, J. R. I., Caspi, A., Moffitt, T. E., & Plomin, R. (2018). A polygenic p factor for major psychiatric disorders. *Translational Psychiatry*, *8*(1), 205. https://doi.org/10.1038/s41398-018-0217-4

Sentencing Academy. (2021). *The Suspended Sentence Order in England and Wales*. https://sentencingacademy.org.uk/2021/09/the-suspended-sentence-order-in-england-and-wales/

Sentencing Council. (2020). *Sentencing offenders with mental disorders, developmental disorders, or neurological impairments*. https://www.sentencingcouncil.org.uk/overarching-guides/magistrates-court/item/sentencing-offenders-with-mental-disorders-developmental-disorders-or-neurological-impairments/

Sentencing Council. (2021). *Breach of a community order*. https://www.sentencingcouncil.org.uk/offences/magistrates-court/item/breach-of-a-community-order-2018/

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *Bmj*, *349*.

Simpson, A. I., Grimbos, T., Chan, C., & Penney, S. R. (2015). Developmental typologies of serious mental illness and violence: Evidence from a forensic psychiatric setting. *Australian & New Zealand Journal of Psychiatry*, *49*(11), 1048–1059.

Sims, B., & Jones, M. (1997). Predicting success or failure on probation: Factors associated with felony probation outcomes. *Crime & Delinquency*, *43*(3), 314–327.

Singh, J., Grann, M., & Fazel, S. (2010). A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychology Review*, *31*, 499–513. https://doi.org/10.1016/j.cpr.2010.11.009

Singh, J. P., Fazel, S., Gueorguieva, R., & Buchanan, A. (2014). Rates of violence in patients classified as high risk by structured risk assessment instruments. *The British Journal of Psychiatry*, *204*(3), 180–187.

Sinha, R. (2008). Chronic stress, drug use, and vulnerability to addiction. *Annals of the New York Academy of Sciences*, *1141*, 105–130. https://doi.org/10.1196/annals.1441.030

Sirdifield, C. (2012). The prevalence of mental health disorders amongst offenders on probation: A literature review. *Journal of Mental Health*, *21*(5), 485–498. https://doi.org/10.3109/09638237.2012.664305

Sirdifield, C., & Brooker, C. (2020). *Maximising positive mental health outcomes for people under probation supervision*. HMIP. https://www.justiceinspectorates.gov.uk/hmiprobation/wp-content/uploads/sites/5/2020/08/Maximising-positive-mental-health-outcomes-for-people-under-probation-supervision.pdf

Sirdifield, C., Marples, R., Brooker, C., & Denney, D. (2019). Improving the Health and Quality of Healthcare for People on Probation. *Probation Quarterly*.

Sjölander, A., & Zetterqvist, J. (2017). Confounders, Mediators, or Colliders: What Types of Shared Covariates Does a Sibling Comparison Design Control For? *Epidemiology*, *28*(4). https://journals.lww.com/epidem/Fulltext/2017/07000/Confounders,_Mediators,_or_Colliders__What_Types.11.aspx

Skeem, J. L., Manchak, S., & Montoya, L. (2017). Comparing Public Safety Outcomes for Traditional Probation vs Specialty Mental Health Probation. *JAMA Psychiatry*, *74*(9), 942–948. https://doi.org/10.1001/jamapsychiatry.2017.1384

Skinner, G. C. M., & Farrington, D. P. (2020). A systematic review and meta-analysis of premature mortality in offenders. *Aggression and Violent Behavior*, *53*, 101431. https://doi.org/https://doi.org/10.1016/j.avb.2020.101431

Sodhi-Berry, N., Knuiman, M., Alan, J., Morgan, V. A., & Preen, D. B. (2015). Pre-sentence mental health service use predicts post-sentence mortality in a population cohort of first-time adult offenders. *Social Psychiatry and Psychiatric Epidemiology*, *50*(1), 109–124. https://doi.org/10.1007/s00127-014-0919-8

Spittal, M. J., Forsyth, S., Borschmann, R., Young, J. T., & Kinner, S. A. (2019). Modifiable risk factors for external cause mortality after release from prison: a nested case–control study. *Epidemiology and Psychiatric Sciences*, *28*(2), 224–233. https://doi.org/DOI: 10.1017/S2045796017000506

Spjeldnes, S., & Goodkind, S. (2009). Gender differences and offender reentry: A review of the literature. *Journal of Offender Rehabilitation*, *48*(4), 314–335.

Stahler, G. J., Mennis, J., Belenko, S., Welsh, W. N., Hiller, M. L., & Zajac, G. (2013). Predicting Recidivism for Released State Prison Offenders: Examining the Influence of Individual and Neighborhood Characteristics and Spatial Contagion on the Likelihood of Reincarceration. *Criminal Justice and Behavior*, *40*(6), 690–711. https://doi.org/10.1177/0093854812469609

StataCorp. (2017). *Stata Statistical Software: Release 15*. StataCorp LLC.

Steel, Z., Marnane, C., Iranpour, C., Chey, T., Jackson, J. W., Patel, V., & Silove, D. (2014). The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *International Journal of Epidemiology*, *43*(2), 476–493. https://doi.org/10.1093/ije/dyu038

Stępniak, M., Pritchard, J. P., Geurs, K. T., & Goliszek, S. (2019). The impact of temporal resolution on public transport accessibility measurement: Review and case study in Poland. *Journal of Transport Geography*, *75*, 8–24. https://doi.org/https://doi.org/10.1016/j.jtrangeo.2019.01.007

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, *338*, b2393. https://doi.org/10.1136/bmj.b2393

Sun, X., Zhou, X., Yu, Y., & Liu, H. (2018). Exploring reporting quality of systematic reviews and Meta-analyses on nursing interventions in patients with Alzheimer’s disease before and after PRISMA introduction. *BMC Medical Research Methodology*, *18*(1), 154. https://doi.org/10.1186/s12874-018-0622-7

Swann, A. C., Lijffijt, M., Lane, S. D., Kjome, K. L., Steinberg, J. L., & Moeller, F. G. (2011). Criminal conviction, impulsivity, and course of illness in bipolar disorder. *Bipolar Disorders*, *13*(2), 173–181. https://doi.org/https://doi.org/10.1111/j.1399-5618.2011.00900.x

Swedish National Council for Crime Prevention. (2017). *Recidivism*. https://www.bra.se/bra-in-english/home/crime-and-statistics/crime-statistics/recidivism.html

Swedish National Council for Crime Prevention. (2021a). *All conviction decisions, by principal offence and year, 1975–2019*. https://www.bra.se/bra-in-english/home/crime-and-statistics/crime-statistics.html

Swedish National Council for Crime Prevention. (2021b). *Crime statistics*. https://www.bra.se/bra-in-english/home/crime-and-statistics/crime-statistics.html

Swedish Prison and Probation Service. (2021). *Probation*. https://www.kriminalvarden.se/swedish-prison-and-probation-service/probation/

Tableau Software. (2021). *Tableau Software Website*. https://www.tableau.com

Théau, J. (2008). *Temporal Resolution BT - Encyclopedia of GIS* (S. Shekhar & H. Xiong, Eds.; pp. 1150–1151). Springer US. https://doi.org/10.1007/978-0-387-35973-1_1376

Therneau, T., & Grambsch, P. (2000). Modeling Survival Data: Extending The Cox Model. In *Stat Med* (Vol. 48). https://doi.org/10.1007/978-1-4757-3294-8

Therneau, T. M. (2021). *A Package for Survival Analysis in R*. https://cran.r-project.org/package=survival

Tonry, M. (2017). Community punishments. *Reforming Criminal Justice: A Report of the Academy for Justice on Bridging the Gap between Scholarship and Reform*, *4*, 187–204.

Trevena, J., & Weatherburn, D. (2015). Does the first prison sentence reduce the risk of further offending? *BOCSAR NSW Crime and Justice Bulletins*, 16.

van Deinse, T. B., Bunger, A., Burgin, S., Wilson, A. B., & Cuddeback, G. S. (2019). Using the Consolidated Framework for Implementation Research to examine implementation determinants of specialty mental health probation. *Health & Justice*, *7*(1), 17. https://doi.org/10.1186/s40352-019-0098-5

van Houwelingen, H. C. (2007). Dynamic prediction by landmarking in event history analysis. *Scandinavian Journal of Statistics*, *34*(1), 70–85.

van Houwelingen, H., & Putter, H. (2011). *Dynamic Prediction in Clinical Survival Analysis* (1st Edition). CRC Press.

Viglione, J., & Taxman, F. S. (2018). Low Risk Offenders Under Probation Supervision: Risk Management and the Risk-Needs-Responsivity Framework. *Criminal Justice and Behavior*, *45*(12), 1809–1831. https://doi.org/10.1177/0093854818790299

Visher, C. A., Winterfield, L., & Coggeshall, M. B. (2005). Ex-offender employment programs and recidivism: A meta-analysis. *Journal of Experimental Criminology*, *1*(3), 295–316.

von Elm, E., Altman, D., Egger, M., Pocock, S., Gotzsche, P., & Vandenbroucke, J. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies. *Lancet*, *370*(9596), 1453–1457.

Vygotsky, L. S. (2012). *Thought and language*. MIT press.

Walter, F., Carr, M. J., Mok, P. L. H., Antonsen, S., Pedersen, C. B., Appleby, L., Fazel, S., Shaw, J., & Webb, R. T. (2019). Multiple adverse outcomes following first discharge from inpatient psychiatric care: a national cohort study. *The Lancet Psychiatry*, *6*(7), 582–589. https://doi.org/https://doi.org/10.1016/S2215-0366(19)30180-4

Warwickshire and West Mercia CRC. (2021). *Community Order*. https://www.wwmcrc.co.uk/page.php

Wermink, H., Apel, R., Nieuwbeerta, P., & Blokland, A. A. J. (2013). The incapacitation effect of first-time imprisonment: a matched samples comparison. *Journal of Quantitative Criminology*, *29*(4), 579–600.

Wermink, H., Blokland, A., Nieuwbeerta, P., Nagin, D., & Tollenaar, N. (2010). Comparing the effects of community service and short-term imprisonment on recidivism: a matched samples approach. *Journal of Experimental Criminology*, *6*(3), 325–349.

Wickham, H., Chang, W., & Wickham, M. H. (2016). Package ‘ggplot2.’ *Create Elegant Data Visualisations Using the Grammar of Graphics. Version*, *2*(1), 1–189.

Wildeman, C., Goldman, A. W., & Wang, E. A. (2019). Age-Standardized Mortality of Persons on Probation, in Jail, or in State Prison and the General Population, 2001-2012. *Public Health Reports*, *134*(6), 660–666. https://doi.org/10.1177/0033354919879732

Wolf, A., Fanshawe, T. R., Sariaslan, A., Cornish, R., Larsson, H., & Fazel, S. (2018). Prediction of violent crime on discharge from secure psychiatric hospitals: A clinical prediction rule (FoVOx). *European Psychiatry : The Journal of the Association of European Psychiatrists*, *47*, 88–93. https://doi.org/10.1016/j.eurpsy.2017.07.011

Wolff, N., Epperson, M., Shi, J., Huening, J., Schumann, B. E., & Sullivan, I. R. (2014). Mental health specialized probation caseloads: Are they effective? *International Journal of Law and Psychiatry*, *37*(5), 464–472. https://doi.org/https://doi.org/10.1016/j.ijlp.2014.02.019

Wong, K., & Horan, R. (2021). *Needs assessment: risk, desistance and engagement*.

Wood, A. M., White, I. R., & Royston, P. (2008). How should variable selection be performed with multiply imputed data? *Statistics in Medicine*, *27*(17), 3227–3246. https://doi.org/10.1002/sim.3177

Wood, M., Cattell, J., Hales, G., Lord, C., Kenny, T., & Capes, T. (2015). Re-offending by offenders on community orders: Results from the Offender Management Community Cohort Study. In *London: Ministry of Justice*.

World Health Organization. (2021). *Global Health Estimates: Life expectancy and leading causes of death and disability*. https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates

World Prison Brief. (2018). *Highest to lowest: prison population total*. https://www.prisonstudies.org/highest-to-lowest/prison-population-total

World Psychiatric Association. (2011). A conceptual framework for the revision of the ICD-10 classification of mental and behavioural disorders. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, *10*(2), 86–92. https://doi.org/10.1002/j.2051-5545.2011.tb00022.x

Wynants, L., van Smeden, M., McLernon, D. J., Timmerman, D., Steyerberg, E. W., Van Calster, B., & initiative, on behalf of the T. G. ‘Evaluating diagnostic tests and prediction models' of the S. (2019). Three myths about risk thresholds for prediction models. *BMC Medicine*, *17*(1), 192. https://doi.org/10.1186/s12916-019-1425-3

Yee, N., Matheson, S., Korobanova, D., Large, M., Nielssen, O., Carr, V., & Dean, K. (2020). A meta-analysis of the relationship between psychosis and any type of criminal offending, in both men and women. *Schizophrenia Research*, *220*, 16–24. https://doi.org/https://doi.org/10.1016/j.schres.2020.04.009

Yukhnenko, D., Blackwood, N., & Fazel, S. (2020). Risk factors for recidivism in individuals receiving community sentences: a systematic review and meta-analysis. *CNS Spectrums*, *25*(2), 252–263. https://doi.org/DOI: 10.1017/S1092852919001056

Yukhnenko, D., Wolf, A., Blackwood, N., & Fazel, S. (2019). Recidivism rates in individuals receiving community sentences: A systematic review. *PLOS ONE*, *14*(9), e0222495.

Zhang, Z., Reinikainen, J., Adeleke, K. A., Pieterse, M. E., & Groothuis-Oudshoorn, C. G. M. (2018). Time-varying covariates and coefficients in Cox regression models. *Annals of Translational Medicine*, *6*(7).

Zhong, S., Senior, M., Yu, R., Perry, A., Hawton, K., Shaw, J., & Fazel, S. (2021). Risk factors for suicide in prisons: a systematic review and meta-analysis. *The Lancet Public Health*, *6*(3), e164–e174. https://doi.org/https://doi.org/10.1016/S2468-2667(20)30233-4

APPENDIX A1. TERMS AND SEARCH CONDITIONS USED FOR SYSTEMATIC SEARCH IN PUBLICATION DATABASES

Search on MEDLINE, 1946 to March 16, 2018, and PsychINFO, 1806 to March 16, 2018, with no language restrictions:

(recidivism OR reconviction OR reoffending) AND (USA OR "United States" OR China OR Russia* OR Brazil OR India OR Thailand OR Mexico OR Iran OR Indonesia OR "South Africa" OR Turkey OR Vietnam OR Colombia OR Philippines OR Ethiopia OR Ukraine OR "United Kingdom" OR England OR Poland OR Pakistan OR Morocco) AND (probation OR "community service" OR "community order" OR "community correction" OR "community sentence")

Search on SAGE Journals, 1902 to March 16, 2018, with no language restrictions:

[[All recidivism] OR [All reconviction] OR [All reoffending]] AND [[All usa] OR [All "united states"] OR [All china] OR [All russia*] OR [All brazil] OR [All india] OR [All thailand] OR [All mexico] OR [All iran] OR [All indonesia] OR [All "south africa"] OR [All turkey] OR [All vietnam] OR [All colombia] OR [All philippines] OR [All ethiopia] OR [All ukraine] OR [All "united kingdom"] OR [All england] OR [All poland] OR [All pakistan] OR [All morocco]] AND [[All probation] OR [All "community service"] OR [All "community order"] OR [All "community correction"] OR [All "community sentence"]]

APPENDIX A2. PRISMA 2009 CHECKLIST (CRIMINAL RECIDIVISM RATES IN COMMUNITY SENTENCED POPULATIONS)

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	13
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	13
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	14
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	16
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	17
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	18
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	17
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Apndx. A1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	18
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	18-19

Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	17-18
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	n/a
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	n/a
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	n/a
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	n/a
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	n/a
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	19
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Apndx. A4
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	n/a
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Apndx. A4, 21-25
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	n/a
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	n/a
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	n/a
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	27-31
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	31

Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	32
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	n/a

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Appendix A3. Identified studies and reports that fit the inclusion criteria

Only reports containing the most recent data were included for given territory or country. The data were mostly reported by governmental agencies; however, four identified papers were published in scientific journals (Bartels, 2009; Flores et al., 2017; Ķipēna et al., 2013; Leonardi, 2007). Several sources (Department of Correctional Services, 2014; Department of Corrections, 2016, 2017; Ķipēna, Zavackis, & Ņikišins, 2013) did not report sizes of the cohorts. Data for Denmark and Oregon, USA were obtained through online data tools on governmental agencies' websites.

Country	Selection period	Length of follow-up	Cohort size	Outcomes	Source
Europe					
Nordic countries					
Denmark	2013	6 months – 2 years	6,501	Reconviction	Statistics Denmark, 2018
Finland	2005	2 years	3,767	Reconviction	Graunbøl et al., 2010
Iceland	2005	2 years	73	Reconviction	Graunbøl et al., 2010
Norway	2005	2 years	2,839	Reconviction	Graunbøl et al., 2010
Sweden	2008	1, 2, 3 years	22,306	Reconviction (after "initial event")	Swedish National Council for Crime Prevention, 2017
The United Kingdom					
England and Wales	2015/2016	1 year	139,617	Proven reoffending	Ministry of Justice, 2018
Northern Ireland	2014/2015	1-12 months	17,560	Proven reoffending	Duncan & Damkat, 2017
Northern Ireland	2005	2 years	19,047	Reconviction	Department of Justice, 2011

Scotland	2014/2015	1 year	21,733	Reconviction	Scottish Government, 2017
Other					
Austria	2004	3 years	93,073	Reconviction	Rebuplik Österreich Parlament, 2012
Germany	2010	3 years	96,521	Reconviction	Hans-Jörg, A. & Jörg-Martin, J., 2014
Italy	1998	7 years	8,817	Reconviction	Leonardi, 2007
Ireland, Republic of	2010	6 months – 3 years	3,698	Reconviction	Central Statistics Office, 2016
Latvia	2009	29 months	n/a	Reconviction (or initiation of proceedings)	Ķipēna, Zavackis, & Ņikišins, 2013
Netherlands	2003	2 years	38,530	Reconviction (or initiation of proceedings)	Wartna & Tollenaar, 2006
North America					
Canada					
Ontario	2013/2014	2 years	35,561	Reconviction (after completing a sentence)	Ontario Ministry of Community Safety and Correctional Services, 2017
Quebec	2007/2008	1 month – 2 years	4,851	Reconviction (after completing a sentence) Reincarceration (after completing a sentence)	Ministère de la Sécurité publique, 2015
USA					
USA (federal)	2004/2005	1-9 years	13,504	Rearrest	Flores, Holsinger, Lowenkamp, & Cohen, 2017
Illinois	2006	5 years	2,770	Rearrest	Illinois Criminal Justice Information Authority, 2011
North Carolina	2013	2 years	35,103	Reconviction Reincarceration	North Carolina Sentencing and Policy Advisory Commission, 2016

				Rearrest	
New York State	2002	5 years	31,267	Reconviction	The Council of State Governments, 2013
Oregon	2014	1, 2, 3 years	4,403	Reconviction Reincarceration Rearrest	State of Oregon Criminal Justice Commission, 2018
<u>South America</u>					
Chile	2007	3 years	23,736	Reconviction Rearrest	Peillard, Correa, Chahuán, & Lacoa, 2012
<u>Oceania</u>					
Australia					
Australia (federal)	2012/2013	2 years	n/a	Reconviction	Department of Correctional Services, 2014
New South Wales	2015	1 year	16,907	Reconviction	Bureau of Crime Statistics and Research, 2017
Tasmania	2002/2004	2 years	416	Reconviction	Bartels, 2009
West Australia	2012/2014	2 years	n/a	Reconviction	Department of Correctional Services, 2014
New Zealand					
New Zealand	2014/2015	1, 2 years	n/a	Reconviction Reincarceration	Department of Corrections, 2017 Department of Corrections, 2016

APPENDIX A4. DESCRIPTION OF DATA EXTRACTED FROM THE STUDIES

Country	Study	Description of outcomes	Description of follow-up	Sentences	Notes and exclusions
Australia – New South Wales	Bureau of Crime Statistics and Research, 2017	Reconviction The crime and conviction should both happen during a follow-up to be counted as recidivism. Technical violations are not automatically counted as reoffence.	1 year. Starts with an imposition of a sentence.	Any sentence other than imprisonment.	A single offender is counted only once.
Australia – Tasmania	Bartels, 2009	Reconviction The crime and conviction should both happen during a follow-up to be counted as recidivism. Technical violations are not automatically counted as reoffence.	2 years. Starts with an imposition of a sentence.	Wholly suspended sentence Noncustodial order	Excludes pseudo-reconvictions. A single offender is counted only once.
Australia – West Australia	Department of Correctional Services, 2014	Reconviction (return) Return of an individual to Corrective Services during a follow-up period.	2 years. Starts with an end of a sentence.	Any sentence that results in an offender entering Community Corrections.	One offender may be counted several times, if he is sentenced, released and reconvicted again during a follow-up period. Excludes returns to community correction Work and Development orders and Bail orders. Excludes fines.
Australia (federal)	Department of Correctional Services, 2014	Reconviction (return) Return of an individual to Corrective Services during a follow-up period.	2 years. Starts with an end of a sentence.	Any sentence that results in an offender entering Community Corrections.	One offender may be counted several times, if he is sentenced, released and reconvicted again during a follow-up period. Excludes fines.
Canada – Ontario	Ontario Ministry of Community Safety and Correctional Services, 2017	Reconviction Return to a provincial correctional supervision after committing an offence during the time of follow-up	2 years. Start with the end of a sentence.	Community supervision	Excludes individuals sentenced to federal prisons
Canada – Quebec	Ministère de la Sécurité publique, 2015	Reconviction The crime and conviction should both happen during a follow-up to be counted as recidivism. Technical violations are not automatically counted as reoffence	2 years. Start with the end of a sentence.	Probation	Two types of reconviction rates are provided in the report: for a period during supervision (i.e., revocation) and during a follow-up after the end of a sentence. Taking into account

					recidivism during serving a sentence increases the 2-year reconviction rate to 52%
Chile	Peillard, Correa, Chahuán, & Lacoa, 2012	Rearrest Any rearrest during a follow-up period. In case of a successful restitution, rearrest in not counted.	3 years. Starts with an imposition of a sentence.	Conditional sentence Probation Night detention	
Denmark	Statistics Denmark, 2018	Reconviction 3 years after follow-up ends, an individual can be sentenced for an offence committed during the follow-up period.	2 years, broken down into 6 monthly periods. Starts with an imposition of a sentence	Suspended sentence Community service Withdrawal of charges with conditions	Individuals of age 19 and younger are excluded. Online tool is provided.
Finland	Graunbøl et al., 2010	Reconviction The offence and conviction should both happen during a follow-up to be counted as recidivism	2 years. Starts with an imposition of a sentence	Probation with supervision Probation with treatment Community service	
France	Ministère de la Justice, 2013	Reconviction The offence and conviction should both happen during a follow-up to be counted as recidivism.	5 years. Starts with an end of a sentence. Broken down by 1-year periods.	Conditional sentence	Two types of reconviction rates are provided in the report: for a period during supervision (i.e., revocation) and during a follow-up after the end of a sentence. Only 2.2% of a cohort were reconvicted while serving a sentence.
Germany	Albrecht & Jehle, 2014	Reconviction The offence and conviction should both happen during a follow-up to be counted as recidivism.	3 years. Starts with an imposition of a sentence	Suspended sentence	The rates are estimated from graphs.
Iceland	Graunbøl et al., 2010	Reconviction The offence and conviction should both happen during a follow-up to be counted as recidivism	2 years. Starts with an imposition of a sentence	Probation with supervision Probation with treatment	
Ireland, Republic of	Central Statistics Office, 2016	Reconviction To be counted as a recidivism event, an offence should occur within a follow-up period and a	6 months – 3 years. Broken down into multiple intervals.	Probation orders Community service orders	Sex offenders and individuals who committed certain road offences are not included in a sample.

		conviction should happen within two years after the offence.			
Italy	Leonardi, 2007	Reconviction Operationalisation is unclear.	7 years	Community sanctions. The exact sentences are unclear	
Latvia	Ķipēna, Zavackis, & Ņikišins, 2013	Reconviction (or initiation of proceedings) A new criminal charge that did not results in acquittal or another technical dismissal during a follow-up period.	29 months. Starts with an imposition of a sentence.	Community service Probation	The report provides two definitions of recidivism (reconviction and initiation of legal proceedings).
Netherlands	Wartna & Tollenaar, 2006	Reconviction (or the initiation of proceedings) A new criminal charge that did not results in acquittal or another technical dismissal during a follow-up period	2 years. Starts with an imposition of a sentence.	Community service Training order Conditional sentence Discretionary dismissal	Rates for fines are reported separately (Fig. 3).
New Zealand	Department of Corrections, 2017 Department of Corrections, 2016	Reconviction The crime and conviction should both happen during a follow-up to be counted as recidivism	1, 2 years. Starts with an end of a sentence	Community sentence	Cohort sizes are not provided. Some offenders may be double-counted if they participate in multiple rehabilitation programmes.
Norway	Graunbøl et al., 2010	Reconviction The offence and conviction should both happen during a follow-up to be counted as recidivism	2 years. Starts with an imposition of a sentence	Probation with supervision Probation with treatment Community service	
Sweden	Swedish National Council for Crime Prevention, 2017	Reconviction The offence and conviction should both happen during a follow-up period to be counted as recidivism	2 years. Starts with an imposition of a sentence	Intensive supervision with electronic monitoring Probation (incl. with community service) Suspended sentence (incl. with community service)	One offender can be counted multiple times. For intensive supervision, follow-up starts at the end of a sentence.

UK – England & Wales	Ministry of Justice, 2018	Proven reoffending 6 months after observational period ends, an individual can be sentenced for an offence committed during this period.	1-year observational period. Starts with an imposition of a sentence	Pre CJA orders Community orders Suspended sentence order	Cautions and discharges are not included. Rates for fines are reported separately (Fig. 3).
UK – N. Ireland	Duncan & Damkat, 2017	Proven reoffending 6 months after observational period ends, an individual can be sentenced for an offence committed during this period. Technical violations are not counted as reoffence	1-year observational period. Starts with an imposition of a sentence	Community supervision Community other	Diversionary disposals are not included. Rates for fines are reported separately (Fig. 3).
UK – N. Ireland	Department of Justice, 2011	Reoffending The offence and conviction should both happen during a follow-up period to be counted as recidivism	2 years. Starts with an imposition of a sentence	Community service order Probation order Combination order Bound over Conditional discharge Suspended prison Other	Fines as an index offence are excluded in extracted overall general recidivism rate (Appendix 4). Rates reported by gender include fines as an index offence (Table 1).
UK – Scotland	Scottish Government, 2017	Reconviction The conviction should happen during a follow-up period to be counted as recidivism	1-year. Starts with an imposition of a sentence	Restriction of liberty order Community payback order Drug treatment and testing orders	Rates for fines are reported separately (Fig. 3). An offence against public justice is not counted as an index offence.
USA – Illinois	Illinois Criminal Justice Information Authority, 2011	Rearrest An arrest during a follow-up period.	5 years. Starts with an imposition of a sentence.	Probation	Data are reported separately for a period during probation and after probation. Combined rate was extracted. Minor traffic offences are excluded.
USA – Michigan	Harding et al., 2013	Reincarceration New sentence that leads to imprisonment during a follow-up period	1, 3, 5 years. Starts with an imposition of a sentence.	Probation	The cohort includes individuals sentenced for a felony. Separate data on reincarceration resulting from

		Reconviction New sentence for a felony during a follow-up period			technical violations and from a new sentence are provided.
USA – New York State	The Council of State Governments, 2013	Reconviction Operationalisation is unclear	5 years. Starts with an imposition of a sentence.	Probation	For all felony probation sentences in NYS, 5 years is also a supervision period.
USA – North Carolina	North Carolina Sentencing and Policy Advisory Commission, 2016	Rearrest Reconviction Reincarceration An occurrence of a respective event during a follow-up period. Arrests for technical violations are not counted as recidivism. Only rearrests when an individual was fingerprinted	1, 2 years. Starts with an imposition of a sentence.	Probation	Excludes offenders with serious DWI, serious misdemeanour traffic offences.
USA – Oregon	State of Oregon Criminal Justice Commission, 2018	Reconviction Rearrest Reincarceration An occurrence of a respective event during a follow-up period. Only rearrests when an individual was fingerprinted.	1, 2, 3 years. Starts with an imposition of a sentence.	Probation	Online tool is provided. One offender can be counted multiple times since each new admission is considered a separate case. The sample does not include those sentenced to felony bench or court probation.
USA (federal)	Flores, Holsinger, Lowenkamp, & Cohen, 2017	Rearrest An arrest during a follow-up period.	1-9 years. Broken down into multiple intervals.	Probation	Rates are based on random samples that consist of offenders that had undertaken PCRA assessment

List of identified sources

Albrecht, H.-J. & Jehle, J.-M. eds., 2014. National Reconviction Statistics and Studies in Europe = Nationale Rückfallstatistiken und-untersuchungen in Europa. Göttinger Studien zu den Kriminalwissenschaften. Available from: http://dx.doi.org/10.17875/gup2014-791.

Bartels, L. (2009). The weight of the sword of Damocles: A reconviction analysis of suspended sentences in Tasmania. The Australian and New Zealand Journal of Criminology 42(1): 72--100.

Bureau of Crime Statistics and Research. (2017). Re-offending statistics for NSW. Retrieved 15 Jun 2018. http://www.bocsar.nsw.gov.au/Pages/bocsar_pages/Re-offending.aspx

Central Statistics Office. (2016). Probation recidivism 2010 cohort. Retrieved 1 May 2018 from: http://www.cso.ie/en/releasesandpublications/er/pror/probationrecidivism2010cohort/

Department of Correctional Services. (2014). Recidivism trends in Western Australia with comparison to national trends. Retrieved 15 Jun 2018: https://www.correctiveservices.wa.gov.au/_files/about-us/statistics-publications/statistics/DCS-recidivism-trends-WA-October2014.pdf

Department of Corrections. (2016). Annual report: 1 July 2015 - 30 June 2016. Retrieved 15 June 2018 from: http://www.corrections.govt.nz/__data/assets/pdf_file/0010/857737/Annual_report_201516.pdf

Department of Corrections. (2017). Annual report: 1 July 2016 - 30 June 2017. Retrieved 15 June 2018 from: http://www.corrections.govt.nz/__data/assets/pdf_file/0006/898629/Annual_Report_2016-17.pdf

Department of Justice. (2011). Adult reconviction in Northern Ireland 2005, Statistics and Research Branch, Department of Justice, Belfast. Retrieved 1 May 2018 from: https://www.justice-ni.gov.uk/sites/default/files/publications/doj/adult-reconviction-in-northern-ireland-2005.pdf

Duncan, L., and Damkat, I. (2017). Adult and youth reoffending in Northern Ireland (2014/15 Cohort), Analytical Service Group, Department of Justice, Belfast. Retrieved from: https://www.justice-ni.gov.uk/sites/default/files/publications/justice/r-bulletin-29-2017-adult-and-youth-reoffending-northern-ireland-201415-cohort.pdf

Flinchum, T., Hevener, H., Hall, M., & Wesoloski, J. (2016). Correctional Program Evaluation: Offenders Placed on Probation or Released from Prison in FY 2013, North Carolina Sentencing and Policy Advisory Commission, Raleigh, NC. Retrieved 1 May 2018 from: https://www.nccourts.gov/assets/documents/publications/recidivism_2016.pdf

Flores, A. W., Holsinger, A.M., Lowenkamp, C.T., and Cohen, T.H. (2017). Time-free effects in predicting recidivism using both fixed and variable follow-up periods: Do different methods produce different results. Crim Just Beh 44(1): 121--137.

Graunbøl H.M., Kielstrup B., Muiluvuori M.-L., Tyni S., Baldursson E.S., Gudmundsdottir H., et al. (2010). Retur: en nordisk undersøgelse af recidiv blant klienter i kriminalforsorgen Oslo: Kriminalomsorgens utdanningssenter. Retrieved 15 April 2018 from: http://www.kriminalforsorgen.dk/Files/Filer/Statistik/Retur_-_nordisk_recidiv_maj_2010.pdf

Harding, D.J., Morenoff, J.D., Nguyen, A.P., and Bushway, S.D. (2017). Short- and long-term effects of imprisonment on future felony convictions and prison admissions. Proceedings of the National Academy of Sciences, 114(42): 11103--11108. doi:10.1073/pnas.1701544114

Illinois Criminal Justice Information Authority. (2011). Examining Illinois probationer characteristics and outcomes. Retrieved 1 May 2018 from: http://www.icjia.state.il.us/assets/pdf/ResearchReports/Examining_IL_probationer_characteristics_and_outcomes_092011.pdf

Ķipēna, K., Zavackis A., and Ņikišins J. (2013). Sodu izcietušo personu noziedzīgo nodarījumu recidīvs. Jurista Vārds, 35(786): 12--17.

Leonardi, F. (2007). Le misure alternative alla detenzione tra reinserimento sociale e abbattimento della recidiva. Rassegna penitenziaria e criminological, 2: 7--26.

Ministère de la Justice (2013). Mesurer la récidive: Contribution à la conférence de consensus de prévention de la récidive. Retrieved 1 June 2018 from: http://www.justice.gouv.fr/art_pix/stat_recidive_2013.pdf

Ministère de la Sécurité publique. (2015). Projet: Enquête sur la récidive/reprise de la clientèle confiée aux Services correctionnels du Québec. Retrieved 1 May 2018 from: https://www.securitepublique.gouv.qc.ca/services-correctionnels/publications-et-statistiques/enquete-sur-la-recidivereprise.html

Ministry of Justice. (2018). Proven reoffending statistics quarterly: January 2016 to March 2016. Retrieved 1 May 2018 from: https://www.gov.uk/government/statistics/proven-reoffending-statistics-january-2016-to-march-2016

Ontario Ministry of Community Safety and Correctional Services. (2017). Rates of recidivism (re-conviction) in Ontario. Retrieved 15 May 2018 from: https://www.mcscs.jus.gov.on.ca/english/Corrections/RatesRecidivism.html

Peillard, A.M.M., Correa, N.M., Chahuán, G.W., and Lacoa, J.F. (2012). La Reincidencia en el Sistema Penitenciario Chileno, Santiago. Retrieved 1 June 2018 from: http://www.pensamientopenal.com.ar/system/files/2017/09/doctrina45811.pdf

Scottish Government. (2017). Reconviction Rates in Scotland: 2014-15 Offender Cohort. Retrieved 1 May 2018 from: https://www.gov.scot/Publications/2017/05/8362/0

State of Oregon Criminal Justice Commission. (2018). Recidivism. Retrieved 1 May 2018 from: http://www.oregon.gov/cjc/data/Pages/recidivism.aspx

Statistics Denmark. (2018). Recidivism. Retrieved 1 May 2018 from: https://www.dst.dk/en/Statistik/emner/levevilkaar/kriminalitet/tilbagefald-til-kriminalitett

Swedish National Council for Crime Prevention. (2017). Recidivism. Retrieved 1 May 2018 from: https://www.bra.se/bra-in-english/home/crime-and-statistics/crime-statistics/recidivism.html

The Council of State Governments. (2013). Improving probation and alternatives to incarceration in New York State: Increasing public safety and reducing spending on prisons and jails. Retrieved 1 May 2018 from: https://csgjusticecenter.org/wp-content/uploads/2013/03/122112_Probation-ATI-Recs_BRIEF_for-NYSAC.pdf

Wartna, B.S.J., and Tollenaar, N. (2006). Recidive 1997-2003: Ontwikkelingen in het niveau van de strafrechtelijke recidive van jeugdige en volwassen daders, Wetenschappelijk Onderzoeken Documentatiecentrum, Den Haag.

APPENDIX A5. RECENT STUDIES IN COMMUNITY SENTENCED POPULATIONS THAT UTILISED ADVANCED RESEARCH DESIGNS

Search on SAGE Journals (January 1, 2014 to July 20, 2019) with no language restrictions. References of screened-in papers were scanned. We included randomised trials and matched cohort studies in adult individuals receiving community sentences. The studies based on self-reported outcomes were excluded.

Search terms: [[All 'community sentence'] OR [All probation]] AND [All recidivism] AND [[All matching] OR [All randomi*]]

Publication	Research question	Sample	Research design	Primary outcome	Main results
Bales & Piqeuro, 2012	The impact of custodial sentences on reoffending	144,416 offenders sentenced to prison or intensive supervision in Florida, USA	Propensity score matching, exact score matching	Reconviction for a felony during 3 years after release (prisoners) or receiving a sentence (community orders)	*3-year reconviction rates (precision/exact matching)* *Note: rates are depended on matching models* Released from custody: 18.2% - 49.1% Community orders: 15.4% - 38.4% *3-year reconviction rates (propensity score matching)* *Note: rates are depended on matching models* Released from custody: 42.9% - 53.8% Community orders: 31.3% - 34.8%
DeVall et al., 2017	The impact of the Swift and Sure Sanctions Probation Programme (SSSPP) on reoffending	758 offenders sentenced to either regular probation or SSSPP in Michigan, USA	Propensity score matching	Charge for a new offence after the allocation to the programme (maximum 23 months of flexible follow-up)	*New charge rates* SSSPP group: 37.7% Comparison group: 46.7%
Evans et al., 2014	The impact of the Proposition 36 programme on reoffending	29,321 drug offenders referred to drug courts or Proposition 36 programme in California, USA. Includes individuals on parole	Propensity score matching	Posttreatment re-arrest during 12 months after the completion of the treatment	*1-year re-arrests rates from (unmatched cohorts)* Prop 36: 48.0% Drug courts: 44.0% *1-year rearrests rates from (matched cohorts)* Prop 36: 49.7%

					Drug courts: 43.1%
Hyatt & Barnes, 2014	The impact of Intensive Supervision Probation (ISP) on probationer recidivism	832 high-risk offenders under the community supervision in Philadelphia, USA	Random forest forecasting	Charge for a new offence during the 12 months following the allocation to the programme	*New charge rates* ISP group: 40,5% Comparison group: 41.6% *Detected violations* ISP group: 43,0% Comparison group: 27.0%
Jollife and Hedderman, 2015	The impact of custodial sentences on reoffending	5,500 male offenders from the United Kingdom	Nearest neighbour matching and stratification based on propensity scores	Reconviction after 12 months after release (prisoners) or receiving a sentence (community orders)	*1-year reconviction rates in matched samples* Released from custody: 51.1% Community orders: 44.5%
Lowenkamp et al., 2014	The impact of STARR training programme for probation officers	999 offenders on post-conviction supervision from the USA (unspecified state)	Random assignment of probation officer to training programmes	Re-arrest during 24 months following the initial conviction	*2-year re-arrest rates for all offenders* STARR trained: 43.0% Control: 48.0% *2-year re-arrest rates for moderate risk offenders* STARR trained: 41.0% Control: 28.0% *2-year re-arrest rates for high risk offenders* STARR trained: 55.0% Control: 55.0%
Pearson et al., 2016	The impact of the 'Citizenship' probation supervision program on recidivism	1,091 offenders with medium to high risk of reconviction from England and Wales	Stepped wedge cluster randomisation	Time to reconviction during the period of 6-18 months from the allocation.	*1-year reoffending rates estimated from survival* curves ≈46.0% in both groups. *Note: no crude rates reported. Some reduction in adjusted reoffending rates in the intervention group (no statistically significant results).*

Quinn and Quinn, 2015	The impact of cognitive-behavioural therapy programme on reoffending	286 defendants on probation supervision with histories of repeated driving while intoxicated offences from New York State, USA	Samples matched by distribution of demographic characteristics	Charge for a serious traffic offence during 3 years from the allocation	*New charge rates* CBT group: 11.0% Comparison group: 24.0%
Sorsby et al., 2017	The impact of Skills for Effective Engagement and Development (SEED) training programme for case officers on community orders compliance	931 individuals receiving community orders from 3 different probation trusts in England and Wales	Regression adjustment based on 1. treatment covariates; 2. propensity scores.	Compliance with community orders (completion)	*Unadjusted non-completion rates* SEED trained: 25,1% Control: 27,7% No matching or randomisation were used during the allocation.
Trevena and Weatherburn, 2015	The impact of short custodial sentences on reoffending	7,920 individuals sentenced to short-term (up to 12 months) imprisonment or receiving suspended sentences (up to 2 years) in New South Wales, Australia	Propensity score matching	Time to reconviction during the period of up to 6 years from the allocation.	*3-year reconviction rates (unmatched)* Released from custody: 45.6% Suspended sentence: 39.7% *3-year reconviction rates (propensity score matching)* Released from custody: 43.4% Suspended sentence: 42.3%

List of identified sources

Bales, W.D. & Piquero, A.R. (2012). Assessing the impact of imprisonment on recidivism. Journal of Experimental Criminology, 8:71-101. https://doi.org/10.1007/s11292-011-9139-3

Christopher T. Lowenkamp, Alexander Holsinger, Charles R. Robinson & Melissa Alexander (2014) Diminishing or durable treatment effects of STARR? A research note on 24-month re-arrest rates, Journal of Crime and Justice, 37:2, 275-283, DOI: 10.1080/0735648X.2012.753849

DeVall, K. E., Lanier, C., Hartmann, D. J., Williamson, S. H., & Askew, L. N. (2017). Intensive Supervision Programs and Recidivism: How Michigan Successfully Targets High-Risk Offenders. The Prison Journal, 97(5), 585–608. https://doi.org/10.1177/0032885517728876

Evans, E., Li, L., Urada, D., & Anglin, M. D. (2014). Comparative Effectiveness of California's Proposition 36 and Drug Court Programs Before and After Propensity Score Matching. Crime & Delinquency, 60(6), 909–938. https://doi.org/10.1177/0011128710382342

Hyatt, J. M., & Barnes, G. C. (2017). An Experimental Evaluation of the Impact of Intensive Supervision on the Recidivism of High-Risk Probationers. Crime & Delinquency, 63(1), 3–38. https://doi.org/10.1177/0011128714555757

Jolliffe, D., & Hedderman, C. (2015). Investigating the Impact of Custody on Reoffending Using Propensity Score Matching. Crime & Delinquency, 61(8), 1051–1077. https://doi.org/10.1177/0011128712466007

Pearson, D. A. S., McDougall, C., Kanaan, M., Torgerson, D. J., & Bowles, R. A. (2016). Evaluation of the Citizenship Evidence-Based Probation Supervision Program Using a Stepped Wedge Cluster Randomized Controlled Trial. Crime & Delinquency, 62(7), 899–924. https://doi.org/10.1177/0011128714530824

Quinn, T. P., & Quinn, E. L. (2015). The Effect of Cognitive-Behavioral Therapy on Driving While Intoxicated Recidivism. Journal of Drug Issues, 45(4), 431–446. https://doi.org/10.1177/0022042615603390

Sorsby, A, Shapland, J, Robinson, G (2017) Using compliance with probation supervision as an interim outcome measure in evaluating a probation initiative. Criminology and Criminal Justice 17(1): 40–61.

Trevena, J., & Weatherburn, D. (2015). Does the first prison sentence reduce the risk of further offending? Contemporary Issues in Crime and Justice, 187. Available from: https://www.bocsar.nsw.gov.au/Documents/CJB/Report-2015-Does-the-first-prison-sentence-reduce-the-risk-of-further-offending-cjb187.pdf

APPENDIX B1. PRISMA 2009 CHECKLIST (RISK FACTORS FOR RECIDIVISM IN COMMUNITY SENTENCED POPULATIONS)

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	33
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	33
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	34-35
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	35
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	37
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	38-39
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	37
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	37
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	38-39
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	39-40

Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	39-40
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	40-41
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	40
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	41
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	41
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	41
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	Figure 2-1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	Table 2-1
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	Table 2-1
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	Apndx. B2
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	Figures 4-1 - 4-7, Table 4-2, Apndx. B3
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	46, 50
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	46, 48
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	51-54

Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	54-55
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	55
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	n/a

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

APPENDIX B2. DESCRIPTION OF DATA USED IN A META-ANALYSIS BY RISK FACTOR DOMAINS.

Some groups do not sum up to the cohort size because of missing records. OASys, TRAS, RNA – standardized recidivism risk assessment tools.

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Gender					
Adams et al., 2011	USA	Individuals sentenced to probation sampled from several counties of Illinois. One individual can be counted several times (2006, 2770, 82%, 31).	Re-arrest (during supervision, av. 19.4 months)	Male (records, 2215, 40%)	Female (542, 32%)
Caudy, 2018	USA	Individuals sentenced to probation from one urban county of an unnamed southwestern state (2011-2013, 10642, 76%, 34)	Re-arrest (fixed end date, mean 1 year)	Male (records, 8076, 12%)	Female (2566, 11%)
Department of Justice, 2011	UK - N. Ireland	National cohort of individuals receiving non-custodial sentences (2005, 19047, 85%, ≈33).	Reconviction (starts with a sentence, 2 years)	Male (records, 16233, 21%)	Female (2814, 10%)
Harris, 2011	USA	A cohort of felony probationers from an unnamed southern central state (1993, 3598, 78%, 29.3)	Re-arrest, excluding arrests for technical violations (starts with a sentence, 3 years)	Male (records, 2800, 46%) *group size approximated from provided ORs	Female (788, 34%) *group size approximated from provided ORs
Huebner & Cobbina, 2007	USA	Individuals sentenced to probation sampled from several counties of Illinois (2000, 3017, 80%, 30.8) *same dataset as in Olson, 2003	Re-arrest (starts with an end of a sentence, 4 years)	Male (records, 2414, 46%)	Female (603, 40%)
Humphrey et al., 2012	USA	Individuals sampled from cohort of standard and reparative probationers from Vermont (1998-2000, 4792, 73.2%, 28.1)	Reconviction (starts with a sentence, 5 years)	Male (records, 3508, 47%)	Female (1284, 38%)

Maliek, 2017	USA	Individuals sentenced to probation from Texas (2014-2017, 10127, 73%, 34.6)	Probation revocation (during supervision, unknown)	Male (records, 7376, 29%)	Female (2751, 26%)
Minor et al., 2003	USA	Individuals sentenced to probation from Kentucky (1996-1999, 200, 68%, median 40.38)	Probation violations (starts with a sentence, 2 years)	Male (records, 136, 28%)	Female (64, 36%)
N. Carolina S. & Ad. Comm., 2018	USA	Individuals sentenced to probation from N. Carolina (2015, 32537, 72%, 32)	Re-arrest (starts with a sentence, 2 years)	Male (RNA, 23427, 40%)	Female (9110, 30%)
Olson & Lurigio, 2000	USA	Individuals sentenced to probation sampled from several counties of Illinois (1997, 2438, 80%, ≈30)	Re-arrest (during supervision, unspecified)	Male (records, 1958, 33%)	Female (480, 27%)
Peilard et al., 2012	Chile	National cohort on individuals receiving non-custodial sentences (2007, 23736, 86%, ≈33)	Re-arrest (starts with a sentence, 3 years)	Male (records, 20389, 27%)	Female (3347, 28%)
Sims & Jones, 1997	USA	Individuals sentenced to probation from North Carolina (1993, 2850, 83%, 27)	Probation failure/revocation (starts with a sentence, 30 months on average) *cohort is selected based on release date	Male (records, 2365, 46%)	Female (485, 31%)
Wood et al., 2015	UK – England & Wales	National cohort of individuals sampled from different probation trusts, excluding Tier 1 probationers (2009-2010, 125718, 84%, ≈32).	Proven reoffending (starts with a sentence, 1 year for an offence to happen + 6 months for conviction)	Male (records, 106082, 36%)	Female (19636, 29%)

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Age					
Adams et al., 2011	USA	Individuals sentenced to probation sampled from several counties of Illinois. One individual can be counted several times (2006, 2770, 82%, 31).	Re-arrest (during supervision, av. 19.4 months)	Younger than 21 (records, 539, 53%)	21 years old and older (2198, 34%)
		Individuals sentenced to probation sampled from several counties. One individual can be sentenced several times (2006, 1762, 80%, 32).	Revocation (during supervision, av. 19.4 months)	Younger than 21 (records, 323, 55%)	21 years old and older (1413, 42%)
Department of Justice, 2011	UK - N. Ireland	National cohort of individuals receiving non-custodial sentences (2005, 19047, 85%, ≈33).	Reconviction (starts with a sentence, 2 years)	Younger than 21 (records, 2573, 32%)	21 years old and older (16474, 17%)
N. Carolina S. & Ad. Comm., 2018	USA	Individuals sentenced to probation from N. Carolina (2015, 32537, 72%, 32)	Re-arrest (starts with a sentence, 2 years)	Younger than 21 (RNA, 4471, 49%) *approximated from provided data	21 years old and older (25180, 34%) *approximated from provided data
Olson & Lurigio, 2000	USA	Individuals sentenced to probation sampled from several counties of Illinois (1997, 2438, 80%, ≈30)	Re-arrest (during supervision, unspecified)	Younger than 21 (records, 310, 45%)	21 years old and older (2128, 30%)
Wood et al., 2015	UK – England & Wales	National cohort of individuals sampled from different probation trusts, excluding Tier 1 probationers (2009-2010, 125718, 84%, ≈32).	Proven reoffending (starts with a sentence, 1 year for an offence to happen + 6 months for conviction)	Younger than 21 (survey, 21102, 43%)	21 years old and older (104616, 33%)

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Marital status					
Adams et al., 2011	USA	Individuals sentenced to probation sampled from several counties of Illinois. One individual can be counted several times (2006, 1762, 80%, 32).	Revocation (during supervision, av. 19.4 months)	Single (records, 1318, 46%)	Married (records, 343, 34%)
N. Carolina S. & Ad. Comm., 2018	USA	Individuals sentenced to probation from N. Carolina (2015, 32537, 72%, 32)	Re-arrest (starts with a sentence, 2 years)	Single (records, 28307, 39%)	Married (4230, 27%)
Olson et al., 2003	USA	Individuals sentenced to probation from Illinois (2000, 3325, 79%, 31)	Re-arrest (during supervision, unspecified)	Single (records, 2580, 29%)	Married (745, 23%)
Sims & Jones, 1997	USA	Individuals sentenced to probation from North Carolina (1993, 2850, 83%, 27)	Probation failure/revocation (starts with a sentence, 30 months on average) *cohort is selected based on release date	Single (records, 2337, 45%)	Married (513, 35%)

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Ethnicity					
Adams et al., 2011	USA	Individuals sentenced to probation sampled from several counties of Illinois. One individual can be counted several times (2006, 1762, 80%, 32).	Revocation (during supervision, av. 19.4 months)	Non-white (records, 411, 63%)	White (1240, 40%)
Maliek, 2017	USA	Individuals sentenced to probation from Texas (2014-2017, 10127, 73%, 34.6)	Probation revocation (during supervision, unknown)	Non-white (records, 8132, 29%)	White (1995, 26%)

Minor et al., 2003	USA	Individuals sentenced to probation from Kentucky (1996-1999, 200, 68%, median 40.38)	Probation violations (starts with a sentence, 2 years)	Non-white (records, 22, 59%)	White (178, 27%)
N. Carolina S. & Ad. Comm., 2018	USA	Individuals sentenced to probation from N. Carolina (2015, 32537, 72%, 32)	Re-arrest (starts with a sentence, 2 years)	Non-white (RNA, 16594, 39%)	White (15943, 35%)
Olson & Lurigio, 2000	USA	Individuals sentenced to probation sampled from several counties of Illinois (1997, 2438, 80%, ≈30)	Revocation (during supervision, unspecified)	Non-white (records, 961, 18%)	White (1477, 11%)
Olson et al., 2003	USA	Individuals sentenced to probation from Illinois (2000, 3325, 79%, 31)	Re-arrest (during supervision, unspecified)	Non-white (records, 1633, 29%)	White (1692, 26%)
Sims & Jones, 1997	USA	Individuals sentenced to probation from North Carolina (1993, 2850, 83%, 27)	Probation failure/revocation (starts with a sentence, 30 months on average) *cohort is selected based on release date	Non-white (records, 1655, 54%)	White (1192, 28%)

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Criminal history					
Adams et al., 2011	USA	Individuals sentenced to probation sampled from several counties of Illinois. One individual can be counted several times (2006, 2770, 82%, 31).	Re-arrest (during supervision, av. 19.4 months)	More than two prior arrests (records, 1736, 47%)	Two or less prior arrests (1020, 22%)
Harris, 2011	USA	A cohort of felony probationers from an unnamed southern central state (1993, 3598, 78%, 29.3)	Re-arrest, excluding arrests for technical violations (starts with a sentence, 3 years)	Having prior criminal record (records, 2229, 50%)	First time adult offenders (1369, 32%)

Humphrey et al., 2012	USA	Individuals sampled from cohort of standard and reparative probationers from Vermont (1998-2000, 4792, 73.2%, 28.1)	Reconviction (starts with a sentence, 5 years)	Having prior criminal record (records, 1917, 55%)	No prior criminal record (2875, 38%)
Maliek, 2017	USA	Individuals sentenced to probation from Texas (2014-2017, 10127, 73%, 34.6)	Probation revocation (during supervision, unknown)	High risk criminal history (TRAS, 932, 62.9%)	Low/moderate risk criminal history (9195, 24.8%)
Minor et al., 2003	USA	Individuals sentenced to probation from Kentucky (1996-1999, 200, 68%, median 40.38)	Probation violations (starts with a sentence, 2 years)	Having prior convictions (records, 91, 42%)	No prior convictions (109, 21%)
N. Carolina S. & Ad. Comm., 2018	USA	Individuals sentenced to probation from N. Carolina (2015, 32537, 72%, 32)	Re-arrest (starts with a sentence, 2 years)	One or more prior arrests (RNA, 25053, 41%)	No prior arrests (7484, 23%)
Olson & Lurigio, 2000	USA	Individuals sentenced to probation sampled from several counties of Illinois (1997, 2438, 80%, ≈30)	Re-arrest (during supervision, unspecified)	Prior adult convictions (records, 1009, 43%)	No prior adult convictions (1429, 24%)
Olson et al., 2003	USA	Individuals sentenced to probation from Illinois (2000, 3325, 79%, 31)	Re-arrest (during supervision, unspecified)	Prior convictions (records, 1679, 35%)	No prior convictions (1646, 20%)
Wood et al., 2015	UK – England & Wales	National cohort of individuals sampled from different probation trusts, excluding Tier 1 probationers (2009-2010, 125718, 84%, ≈32).	Proven reoffending (starts with a sentence, 1 year for an offence to happen + 6 months for conviction)	Previous convictions (records, 107616, 38%)	No previous convictions (18102, 8%)

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Educational problems					
Adams et al., 2011	USA	Individuals sentenced to probation sampled from several counties of Illinois. One individual can be counted several times (2006, 1762, 82%, 32).	Revocation (during supervision, av. 19.4 months)	No high school diploma (records, 283, 57%)	High school diploma or higher (1473, 42%)

Harris, 2011	USA	A cohort of felony probationers from an unnamed southern central state (1993, 3598, 78%, 29.3)	Re-arrest, excluding arrests for technical violations (starts with a sentence, 3 years)	Low level of academic skills that impairs functioning (records, 1155, 43%)	Normal level of academic skills (2443, 44%)
Maliek, 2017	USA	Individuals sentenced to probation from Texas (2014-2017, 10127, 73%, 34.6)	Probation revocation (during supervision, unknown)	No high school diploma (records, 3113, 40%)	High school diploma or higher (7014, 23%)
Minor et al., 2003	USA	Individuals sentenced to probation from Kentucky (1996-1999, 200, 68%, median 40.38)	Probation violations (starts with a sentence, 2 years)	No high school diploma (records, 56, 36%)	High school diploma or higher (142, 29%)
N. Carolina S. & Ad. Comm., 2018	USA	Individuals sentenced to probation from N. Carolina (2015, 32537, 72%, 32)	Re-arrest (starts with a sentence, 2 years)	High school dropout/GED (RNA, 16919, 43%)	High school graduate (15618, 31%)
Olson & Lurigio, 2000	USA	Individuals sentenced to probation sampled from several counties of Illinois (1997, 2438, 80%, ≈30)	Revocation (during supervision, unspecified)	No high school diploma (records, 741, 19%)	High school diploma or higher (1697, 10%)
Olson et al., 2003	USA	Individuals sentenced to probation from Illinois (2000, 3325, 79%, 31)	Re-arrest (during supervision, unspecified)	No high school diploma (records, 998, 34%)	High school diploma or higher (2328, 25%)
			Violation (during supervision, unspecified)	No high school diploma (records, 998, 49%)	High school diploma or higher (2328, 39%)
Sims & Jones, 1997	USA	Individuals sentenced to probation from North Carolina (1993, 2850, 83%, 27)	Probation failure/revocation (during supervision, av. 30 months) *cohort is selected based on release date	No high school diploma (records, 1396, 62%)	High school diploma or higher (1454, 41%)
Wood et al., 2015	UK – England & Wales	National cohort of individuals sampled from different probation trusts, excluding Tier 1 probationers (2009-2010, 1496, 84%, ≈32).	Proven reoffending (starts with a sentence, 1 year for an offence to happen + 6 months for conviction)	Identified educational needs (survey, 430, 40%)	No identified educational needs (1066, 32%)

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Substance misuse					
Grann et al., 2008	Sweden	National cohort of individuals receiving non-custodial sentences (1993-2001, 4828, 35.7, 91%	Reconviction for a violent crime (starts with a sentence + fixed end date, mean 4.8 years)	Diagnosed with substance use disorder (DSM-III/DSM-IV, 2336, 36%)	No diagnosed disorder (159, 22%)
Harris, 2011	USA	A cohort of felony probationers from an unnamed southern central state (1993, 3598, 78%, 29.3)	Re-arrest, excluding arrests for technical violations (starts with a sentence, 3 years)	Drug usage related to criminal activity (records, 1611, 45%) *group size approximated from provided ORs	No drug usage or unrelated to criminal activity (1987, 43%) *group size approximated from provided ORs
				Alcohol usage related to criminal activity (records, 1689, 46%) *group size approximated from provided ORs	No alcohol usage or unrelated to criminal activity (1909, 41%) *group size approximated from provided ORs
Huebner & Cobbina, 2007	USA	Individuals sentenced to probation sampled from several counties of Illinois (2000, 3017, 80%, 30.8)	Re-arrest (starts with an end of a sentence, 4 years)	History of drug use (records, 1934, 48%)	No history of drug use (1083, 44%)
Maliek, 2017	USA	Individuals sentenced to probation from Texas (2014-2017, 10127, 73%, 34.6)	Probation revocation (during supervision, unknown)	High level of substance abuse need (TRAS, 416, 61%)	Low/level of substance abuse need (911, 27%)
N. Carolina S. & Ad. Commission 2018	USA	Individuals sentenced to probation from N. Carolina (2015, 32537, 72%, 32)	Re-arrest (starts with a sentence, 2 years)	Indicated substance abuse (RNA, 21474, 39%)	No indicated substance abuse (11063, 30%)
Olson & Lurigio, 2000	USA	Individuals sentenced to probation sampled from several counties of Illinois (1997, 2438, 80%, ≈30)	Re-arrest (during supervision, unspecified)	History of drug abuse (records, 868, 44.9%)	No history of drug use (1570, 24.1%)

Sims & Jones, 1997	USA	Individuals sentenced to probation from North Carolina (1993, 2850, 83%, 27)	Probation failure/revocation (during supervision, av. 30 months) *cohort is selected based on release date	Identified drug use problem (records, 826, 48%)	No identified drug use problem (2024, 41%)
				Identified alcohol consumption problem (records, 826, 43%)	No identified drug use problem (2024, 41%)
Wood et al., 2015	UK – England & Wales	National cohort of individuals sampled from different probation trusts, excluding Tier 1 probationers (2009-2010, 1496, 84%, ≈32).	Proven reoffending (starts with a sentence, 1 year for an offence to happen + 6 months for conviction)	Identified drug misuse (OASys, 339, 55%)	No identified drug misuse (860, 24%)
				Identified alcohol misuse (OASys, 419, 35%)	No identified alcohol misuse (629, 36%)

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Mental health					
Grann et al., 2008	Sweden	National cohort of individuals receiving non-custodial sentences (1993-2001, 4828, 35.7, 91%	Reconviction for a violent crime (starts with a sentence + fixed end date, average 4.8 years)	Diagnosed with schizophrenia (DSM-III/DSM-IV, 248, 23%)	No diagnosed disorder (159, 22%)
				Diagnosed with depression (DSM-III/DSM-IV, 308, 60%)	No diagnosed disorder (159, 22%)
				Diagnosed with any personality disorder (DSM-III/DSM-IV, 2159, 35%)	No diagnosed disorder (159, 22%)
				Diagnosed with schizophrenia, depression or any personality disorder (DSM-III/DSM-IV, 2715, 32%)	No diagnosed disorder (159, 22%)
Harris, 2011	USA	A cohort of felony probationers from an unnamed southern central state (1993, 3598, 78%, 29.3)	Re-arrest, excluding arrests for technical violations (starts with a sentence, 3 years)	Mental impairment. Poor motor skills or lower IQ (records, 233, 48%) *group size approximated from provided ORs	Normal IQ and motor skills (3365, 43%)
				Emotional instability. Depression, anxiety, anger, impulsivity limit functioning (records, 832, 46%) *group size approximated from provided ORs	Symptoms absent or do not limit functioning (2766, 43%)

Maliek, 2017	USA	Individuals sentenced to probation from Texas (2014-2017, 10127, 73%, 34.6)	Probation revocation (during supervision, unknown)	Enrolled to mental health supervision (records, 424, 36%)	Other probationers (9703, 28%)
Wood et al., 2015	UK – England & Wales	National cohort of individuals sampled from different probation trusts, excluding Tier 1 probationers (2009-2010, 1496, 84%, ≈32).	Proven reoffending (starts with a sentence, 1 year for an offence to happen + 6 months for conviction)	Identified mental health need (survey, 304, 41%)	No identified mental health needs (1192, 33%)

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Association with antisocial peers					
Adams et al., 2011	USA	Individuals sentenced to probation sampled from several counties of Illinois. One individual can be counted several times (2006, 2770, 82%, 31).	Re-arrest (during supervision, av. 19.4 months)	Known gang affiliation (records, 106, 62%)	No known gang affiliation (1947, 34%)
Harris, 2011	USA	A cohort of felony probationers from an unnamed southern central state (1993, 3598, 78%, 29.3)	Re-arrest, excluding arrests for technical violations (starts with a sentence, 3 years)	Deviant companions (records, 2251, 45%) *group size approximated from provided ORs	No known deviant companions (1347, 42%)
Maliek, 2017	USA	Individuals sentenced to probation from Texas (2014-2017, 10127, 73%, 34.6)	Probation revocation (during supervision, unknown)	High risk negative peer association (TRAS, 566, 61.8%)	Low/moderate risk negative peer association (9561, 26.3%)
Olson et al., 2003	USA	Individuals sentenced to probation from Illinois (2000, 3325, 79%, 31)	Re-arrest (during supervision, unspecified)	Gang affiliation (records, 216, 47%)	No known gang affiliation (3109, 26%)
Sims & Jones, 1997	USA	Individuals sentenced to probation from North Carolina (1993, 2850, 83%, 27)	Probation failure/revocation (during supervision, av. 30 months) *cohort is selected based on release date	Negative friends (records, 1111, 46%)	Positive friends (1739, 41%)

Wood et al., 2015	UK – England & Wales	National cohort of individuals sampled from different probation trusts, excluding Tier 1 probationers (2009-2010, 1496, 84%, ≈32).	Proven reoffending (starts with a sentence, 1 year for an offence to happen + 6 months for conviction)	Reckless lifestyle or deviant associates (OASys, 569, 43%)	No reckless lifestyle or deviant associates (479, 23%)

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Employment problems					
Adams et al., 2011	USA	Individuals sentenced to probation sampled from several counties of Illinois. One individual can be counted several times (2006, 2770, 82%, 31).	Re-arrest (during supervision, av. 19.4 months)	Unemployed (records, 1484, 46%)	Employed (1174, 28%)
Harris, 2011	USA	A cohort of felony probationers from an unnamed southern central state (1993, 3598, 78%, 29.3)	Re-arrest, excluding arrests for technical violations (starts with a sentence, 3 years)	Unemployment or unsatisfactory employment (records, 1572, 47%) *group size approximated from provided ORs	Employed (2026, 41%)
Maliek, 2017	USA	Individuals sentenced to probation from Texas (2014-2017, 10127, 73%, 34.6)	Probation revocation (during supervision, unknown)	Unemployed (TRAS, 1856, 49.2%)	Employed full-time (5972, 21.1%)
Huebner & Cobbina, 2007	USA	Individuals sentenced to probation sampled from several counties of Illinois (2000, 3017, 80%, 30.8) *same dataset as in Olson, 2003	Re-arrest (starts with an end of a sentence, 4 years)	Unemployed (records, 1237, 55.4%)	Employed (1780, 38%)
Minor et al., 2003	USA	Individuals sentenced to probation from Kentucky (1996-1999, 200, 68%, median 40.38)	Probation violations (starts with a sentence, 2 years)	Unemployed (records, 75, 31%)	Employed (126, 30%)
N. Carolina S. & Ad. Comm., 2018	USA	Individuals sentenced to probation from N. Carolina (2015, 32537, 72%, 32)	Re-arrest (starts with a sentence, 2 years)	Unemployed (RNA, 15943, 39%)	Employed (16594, 34%)

Sims & Jones, 1997	USA	Individuals sentenced to probation from North Carolina (1993, 2850, 83%, 27)	Probation failure/revocation (during supervision, av. 30 months) *cohort is selected based on release date	Unstable employment history (records, 1738, 47%)	Stable employment history (1111, 36%)
Wood et al., 2015	UK – England & Wales	National cohort of individuals sampled from different probation trusts, excluding Tier 1 probationers (2009-2010, 1496, 84%, ≈32).	Proven reoffending (starts with a sentence, 1 year for an offence to happen + 6 months for conviction)	Identified need of employment (survey, 924, 27%)	No identified need of employment (572, 45%)

Study	Country	Cohort (selection years, n, % male, mean age)	Outcome (type and length of follow-up)	Exposure (source of data, n, % with outcome)	Comparison (n, % with outcome)
Low income					
Adams et al., 2011	USA	Individuals sentenced to probation sampled from several counties of Illinois. One individual can be counted several times (2006, 2770, 82%, 31).	Re-arrest (during supervision, av. 19.4 months)	Income less than $10,000 per year (records, 1122, 50%)	Income more than $10,000 per year (554, 24%)
Harris, 2011	USA	A cohort of felony probationers from an unnamed southern central state (1993, 3598, 78%, 29.3)	Re-arrest, excluding arrests for technical violations (starts with a sentence, 3 years)	Difficulties meeting financial obligations (records, 2523, 43%) *group size approximated from provided ORs	No significant financial difficulties (1075, 43%) *group size approximated from provided ORs
Olson & Lurigio, 2000	USA	Individuals sentenced to probation sampled from several counties of Illinois (1997, 2438, 80%, ≈30)	Re-arrest (during supervision, unspecified)	Income less than $15,000 per year (records, 1563, 39%)	Income less than $15.000 per year (875, 20%)

Wood et al., 2015	UK – England & Wales	National cohort of individuals sampled from different probation trusts, excluding Tier 1 probationers (2009-2010, 1496, 84%, ≈32).	Proven reoffending (starts with a sentence, 1 year for an offence to happen + 6 months for conviction)	Identified financial needs (survey, 420, 45%)	No identified financial needs (1076, 30%)

APPENDIX B3. ASSOCIATION BETWEEN STATIC (NON-MODIFIABLE) RISK FACTORS AND RECIDIVISM IN INDIVIDUALS GIVEN COMMUNITY SENTENCES

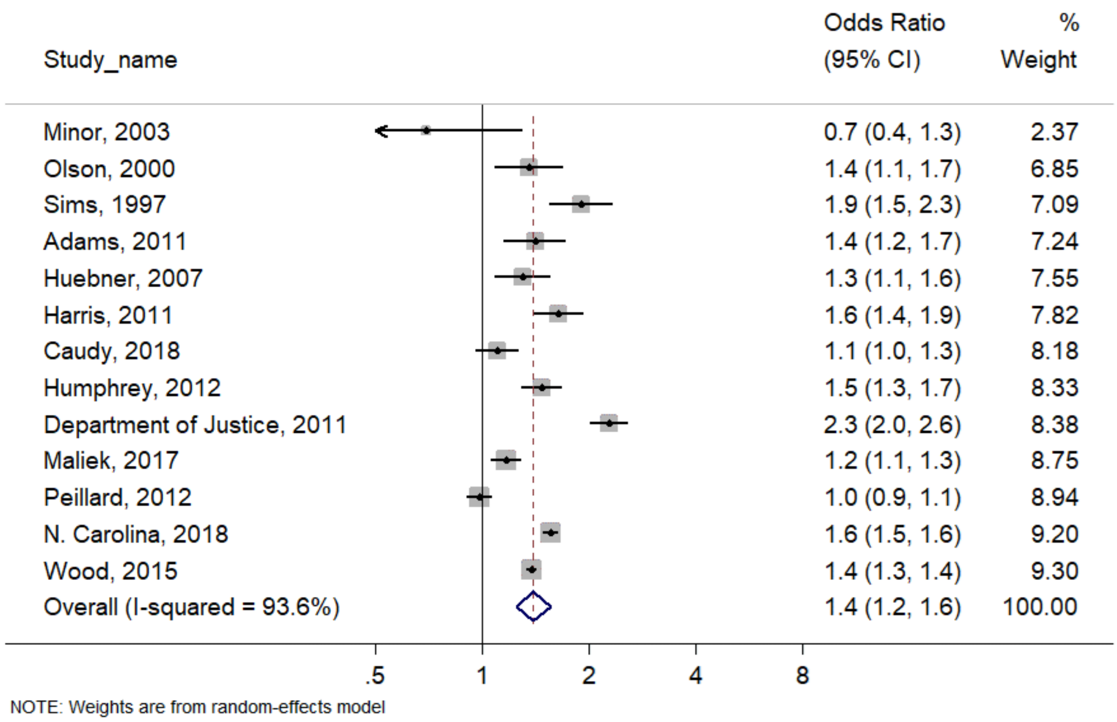


Figure 5-18. Odds ratios (ORs) for the association between gender (being male) and the risk of recidivism in community sentenced populations

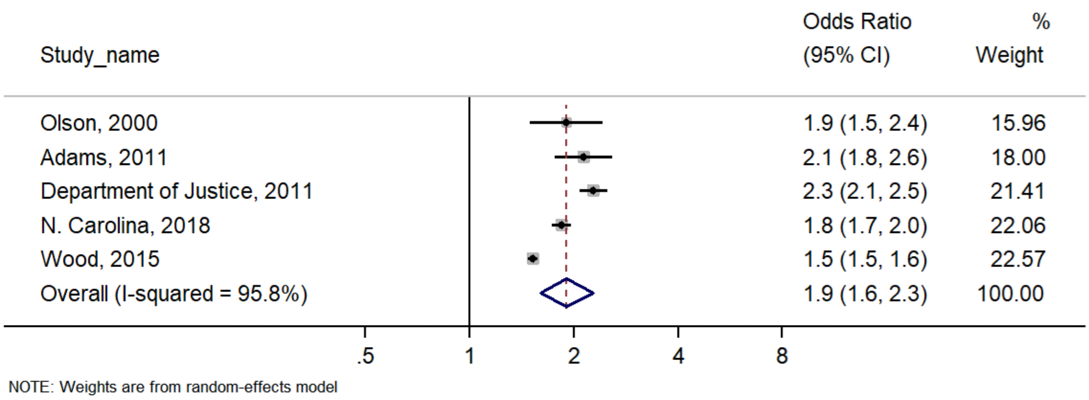


Figure 5-19. Odds ratios (ORs) for the association between age (being younger than 21 years old) and the risk of recidivism in community sentenced populations

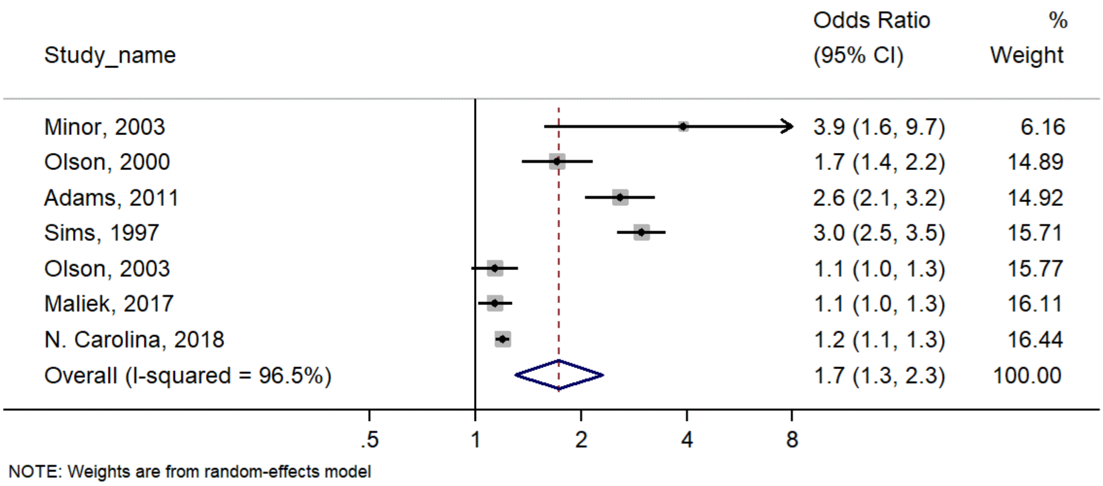


Figure 5-20. Odds ratios (ORs) for the association between ethnicity (being non-white) and the risk of recidivism in community sentenced populations

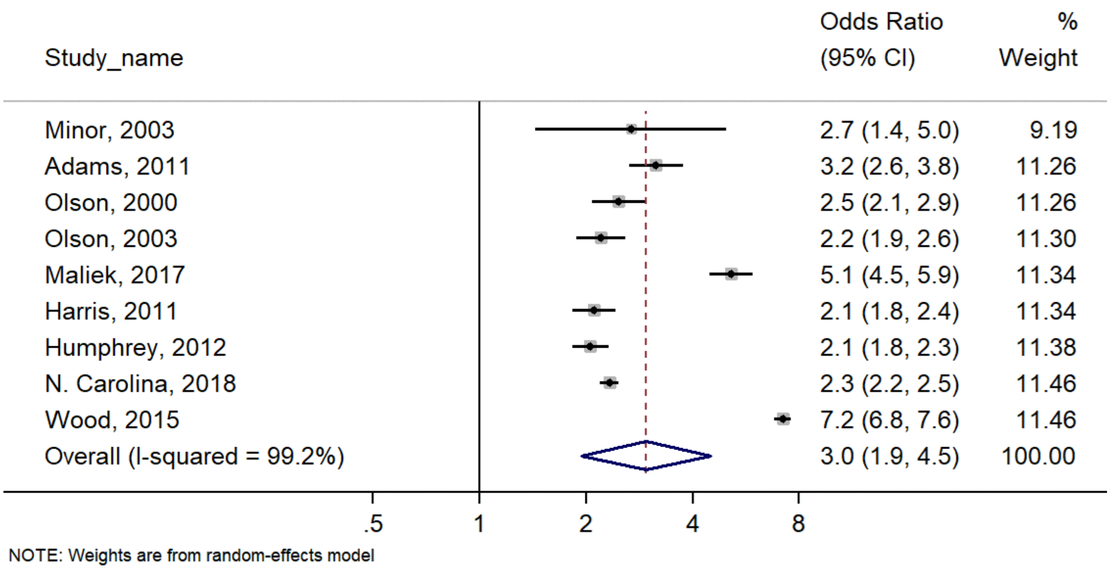


Figure 5-21. Odds ratios (ORs) for the association between criminal history (having a prior arrest or conviction) and the risk of recidivism in community sentenced populations

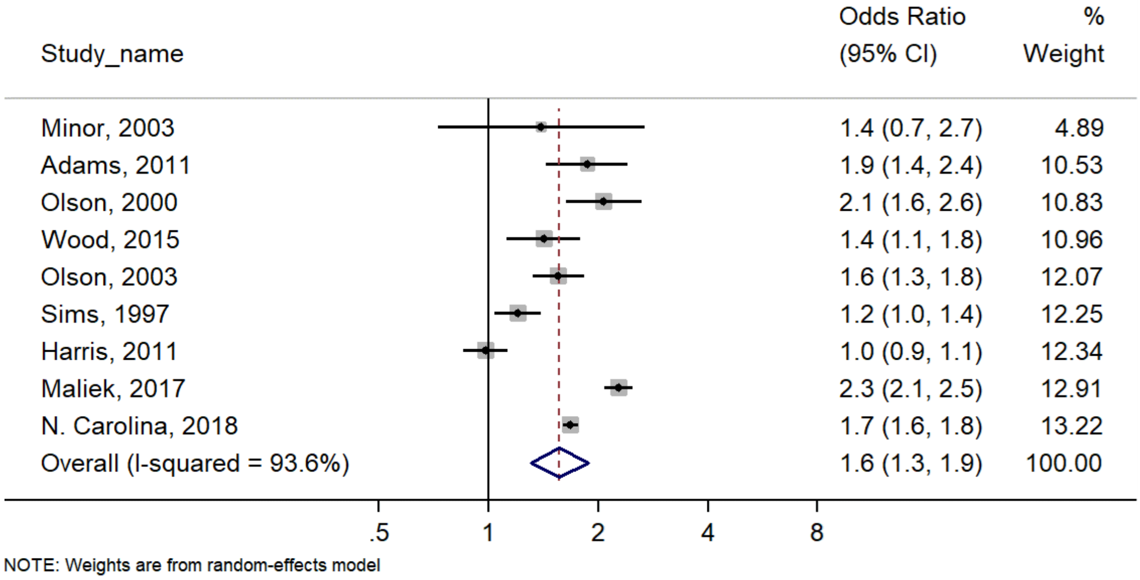


Figure 5-22. Odds ratios (ORs) for the association between educational problems (not graduating high school or having educational needs identified by standardised assessment tools) and the risk of recidivism in community sentenced populations

APPENDIX C1. STROBE CHECKLIST (RECIDIVISM)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

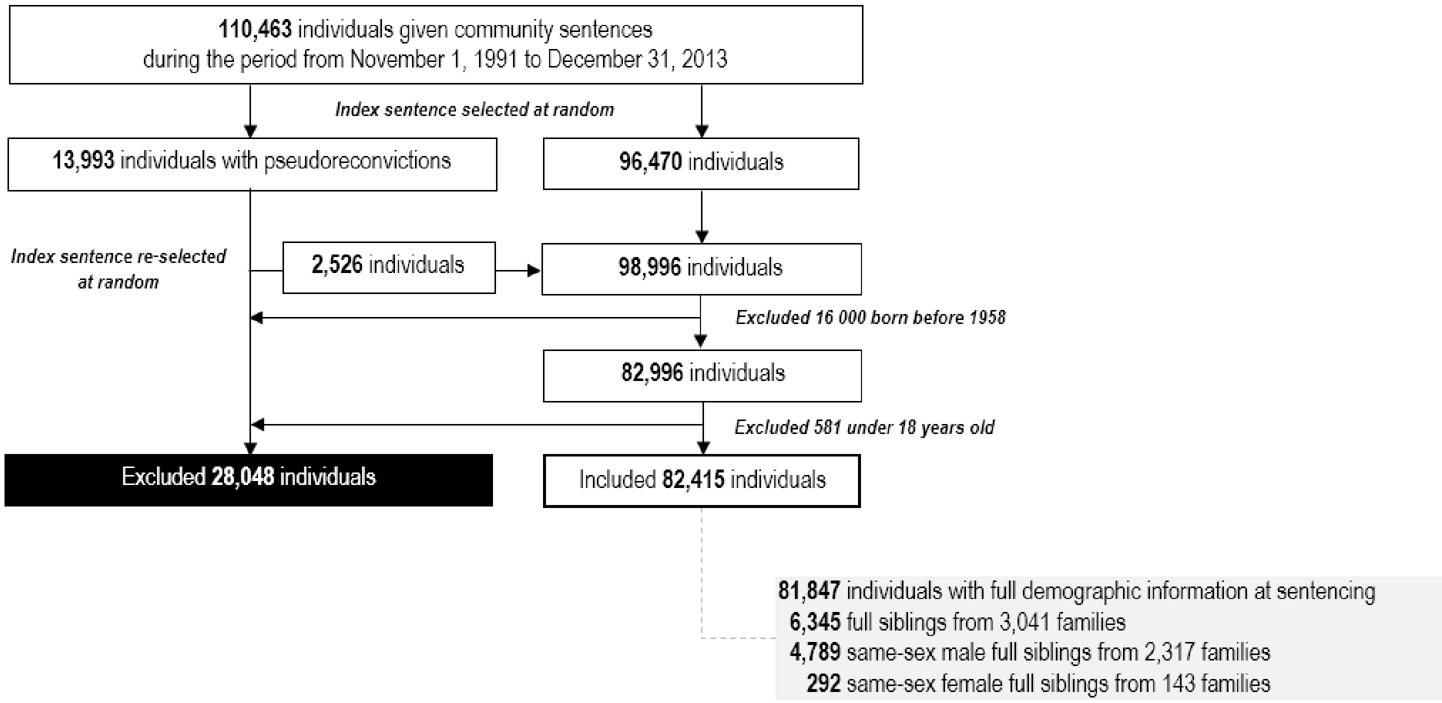
	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	57
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	57
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	58-59
Objectives	3	State specific objectives, including any prespecified hypotheses	59
Methods			
Study design	4	Present key elements of study design early in the paper	61, 66
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	61
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	62-63
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	63,-65
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	61
Bias	9	Describe any efforts to address potential sources of bias	66
Study size	10	Explain how the study size was arrived at	62
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	63-64
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	66-67
		(*b*) Describe any methods used to examine subgroups and interactions	n/a
		(*c*) Explain how missing data were addressed	64
		(*d*) If applicable, explain how loss to follow-up was addressed	62-63
		(*e*) Describe any sensitivity analyses	64
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Table 3-1
		(b) Give reasons for non-participation at each stage	Table 3-1
		(c) Consider use of a flow diagram	Apndx. C2

Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Table 3-1
		(b) Indicate number of participants with missing data for each variable of interest	Table 3-1
		(c) Summarise follow-up time (eg, average and total amount)	Table 3-1, Figure 3.-2
Outcome data	15*	Report numbers of outcome events or summary measures over time	Table 3-1, Figure 3.-2
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Apndx. C4
		(*b*) Report category boundaries when continuous variables were categorized	Table 3-1
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Table 3-1
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	84
Discussion			
Key results	18	Summarise key results with reference to study objectives	84-85
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	89
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	85-88
Generalisability	21	Discuss the generalisability (external validity) of the study results	89-90
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	n/a

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

APPENDIX C2. SELECTION PROCESS FOR THE RECIDIVISM ANALYSIS COHORT



APPENDIX C3. ICD CODES FOR EXTRACTED VARIABLES.

Codes for prior psychiatric diagnoses and self-harm used in the mortality study.

Diagnosis	Codes
Any psychiatric	ICD-8 codes: 290-315 ICD-9 codes: 290-319 ICD-10 codes: F00-F99
Any psychiatric (excluding substance use diagnoses)	-‖- excluding codes for substance use
Schizophrenia spectrum disorder	ICD-8 codes: 295, 297, 298.1-9, 299 ICD-9 codes: 295, 297, 298 (ex. A), 299 ICD-10 codes: F20-F29
Bipolar disorder	ICD-8 codes: 296.1, 296.3, 296.8 ICD-9 codes: 296A, 296C-E, 296W ICD-10 codes: F30-F31
Depressive disorder	ICD-8 codes: 296.2, 296.9, 298.0, 300.4 ICD-9 codes: 296B, 296X, 298A, 300E, 311 ICD-10 codes: F32-F39
Anxiety disorder	ICD-8 codes: 300 (ex. .4), 305, 307 ICD-9 codes: 300 (ex. E), 306, 308, 309 ICD-10 codes: F40-F48
Alcohol use disorder	ICD-8 codes: 291, 303 ICD-9 codes: 291, 303, 305A ICD-10 codes: F10
Drug use disorder	ICD-8 codes: 304 ICD-9 codes: 292, 304, 305 (ex. A) ICD-10 codes: F11-F19
Personality disorder	ICD-8 codes: 301 ICD-9 codes: 301 (ex. B) ICD-10 codes: F60-F61
Attention-deficit and hyperactivity disorder	ICD-9 codes: 314 ICD-10 codes: F90
Other developmental disorder	ICD-8 codes: 308 ICD-9 codes: 299A, 312, 313, 315 ICD-10 codes: F80-F98 (ex. F90)
Self-harm	ICD-10 codes: X60-X69, X70-X79, X81-X84, Y10-Y19, Y20-Y29, Y30-Y34

Appendix C4. Univariate association between baseline sociodemographic/clinical factors and criminal recidivism in individuals given community sentences (by sex)

Supplementary table. Unadjusted association between baseline sociodemographic/clinical factor and reoffending (general and violent) in men individuals given community sentences.

510 men have missing values for marital status, employment, and income support. 2,912 men have missing values for education.

Men (N = 70,643)		**General reoffending**		**Violent reoffending**	
	N of individuals	N with outcome	Hazard ratio (95% CI)	N with outcome	Hazard ratio (95% CI)
Age					
18-24 years	28,299	15,560 (55.0%)	1	5,400 (19.1%)	1
25-39 years	27,468	13,101 (47.7%)	0.78 (0.76-0.80)	3,892 (14.2%)	0.70 (0.67-0.73)
≥ 40 years	14,876	5,113 (34.4%)	0.55 (0.54-0.57)	1,299 (8.7%)	0.46 (0.43-0.49)
Civil status					
Single	7,453	2,464 (33.1%)	0.59 (0.57-0.62)	719 (9.6%)	0.6 (0.56-0.65)
Married	62,680	31,100 (49.6%)	1	9,824 (15.7%)	1
Highest education					
< 9 yr	2,322	1,237 (53.3%)	1	360 (15.5%)	1
9-12 yr	60,649	29,594 (48.8%)	0.78 (0.74-0.83)	320 (6.7%)	0.35 (0.30-0.41)
≥ 12 yr	4,760	1,153 (24.2%)	0.33 (0.30-0.36)	9,240 (15.2%)	0.84 (0.75-0.93)
Employed					
Yes	31,153	10,625 (34.1%)	0.45 (0.44-0.46)	3,240 (10.4%)	0.49 (0.47-0.51)
No	38,980	22,939 (58.8%)	1	7,303 (18.7%)	1
Recipient of income support					
Yes	24,367	15,896 (65.2%)	2.21 (2.16-2.26)	5,126 (21.0%)	1.97 (1.90-2.05)
No	45,766	17,668 (38.6%)	1	5,417 (11.8%)	1
Prior criminal history					
Any prior convictions	54,395	29,681 (54.6%)	2.72 (2.63-2.81)	9,369 (17.2%)	2.48 (2.34-2.63)
No prior convictions	16,248	4,093 (25.2%)	1	1,222 (7.5%)	1
Prior violent crime					
Prior conviction for a violent offence	27,222	16,600 (61.0%)	1.9 (1.86-1.94)	6,147 (22.6%)	2.54 (2.45-2.64)
No convictions for a violent offence	43,421	17,174 (39.6%)	1	4,444 (10.2%)	1

Prior prison					
Prior imprisonment	15,755	10,149 (64.4%)	1.9 (1.86-1.95)	2,906 (18.4%)	1.54 (1.47-1.60)
No prior imprisonment	54,888	23,625 (43.0%)	1	7,685 (14.0%)	1
Index violent crime					
Yes	32,941	13,943 (42.3%)	0.77 (0.76-0.79)	5,881 (17.9%)	1.54 (1.48-1.60)
No	37,702	19,831 (52.6%)	1	4,710 (12.5%)	1
Any psychiatric disorder					
Yes	27,138	14,281 (52.6%)	1.46 (1.43-1.5)	4,536 (16.7%)	1.43 (1.38-1.49)
No	43,505	19,493 (44.8%)	1	6,055 (13.9%)	1
Any psychiatric disorder (other than substance use)					
Yes	18,047	8,717 (48.3%)	1.19 (1.16-1.22)	2,964 (16.4%)	1.34 (1.28-1.40)
No	52,596	25,057 (47.6%)	1	7,627 (14.5%)	1
Schizophrenia spectrum disorder					
Yes	2,032	1,039 (51.1%)	1.21 (1.14-1.29)	425 (20.9%)	1.55 (1.40-1.70)
No	68,611	32,735 (47.7%)	1	10,166 (14.8%)	1
Bipolar disorder					
Yes	690	266 (38.6%)	0.90 (0.80-1.01)	82 (11.9%)	0.95 (0.76-1.18)
No	69,953	33,508 (47.9%)	1	10,509 (15.0%)	1
Depressive disorder					
Yes	5,447	2,308 (42.4%)	0.95 (0.91-0.99)	720 (13.2%)	0.98 (0.91-1.05)
No	65,196	31,466 (48.3%)	1	9,871 (15.1%)	1
Anxiety disorder					
Yes	5,604	2,699 (48.2%)	1.12 (1.08-1.17)	875 (15.6%)	1.17 (1.09-1.25)
No	65,039	31,075 (47.8%)	1	9,716 (14.9%)	1
Alcohol use disorder					
Yes	11,569	6,101 (52.7%)	1.32 (1.28-1.36)	2,091 (18.1%)	1.44 (1.37-1.51)
No	59,074	27,673 (46.8%)	1	8,500 (14.4%)	1
Drug use disorder					
Yes	11,864	7,607 (64.1%)	2.12 (2.06-2.17)	2,144 (18.1%)	1.57 (1.50-1.65)
No	58,779	26,167 (44.5%)	1	8,447 (14.4%)	1
Substance use disorder (alcohol or drug use disorder)					
Yes	18,680	10,777 (57.7%)	1.71 (1.67-1.75)	3,272 (17.5%)	1.49 (1.43-1.56)
No	51,963	22,997 (44.3%)	1	7,319 (14.1%)	1
Personality disorder					
Yes	2,671	1,578 (59.1%)	1.52 (1.45-1.6)	606 (22.7%)	1.78 (1.64-1.93)
No	67,972	32,196 (47.4%)	1	9,985 (14.7%)	1
Attention deficit hyperactivity disorder					

Yes	3,370	1,594 (47.3%)	1.45 (1.38-1.52)	568 (16.9%)	1.75 (1.61-1.9)
No	67,273	32,180 (47.8%)	1	10,023 (14.9%)	1
Other developmental or childhood disorder					
Yes	3,246	1,673 (51.5%)	1.37 (1.31-1.44)	684 (21.1%)	1.83 (1.7-1.98)
No	67,397	32,101 (47.6%)	1	9,907 (14.7%)	1
Prior self-harm					
Yes	5,811	2,951 (50.8%)	1.26 (1.22-1.31)	979 (16.8%)	1.34 (1.25-1.43)
No	64,832	30,823 (47.5%)	1	9,612 (14.8%)	1

Supplementary table. Unadjusted association between baseline sociodemographic/clinical factor and reoffending (general and violent) in women given community sentences.

58 women have missing values for marital status, employment, and income support. 423 women have missing values for education.

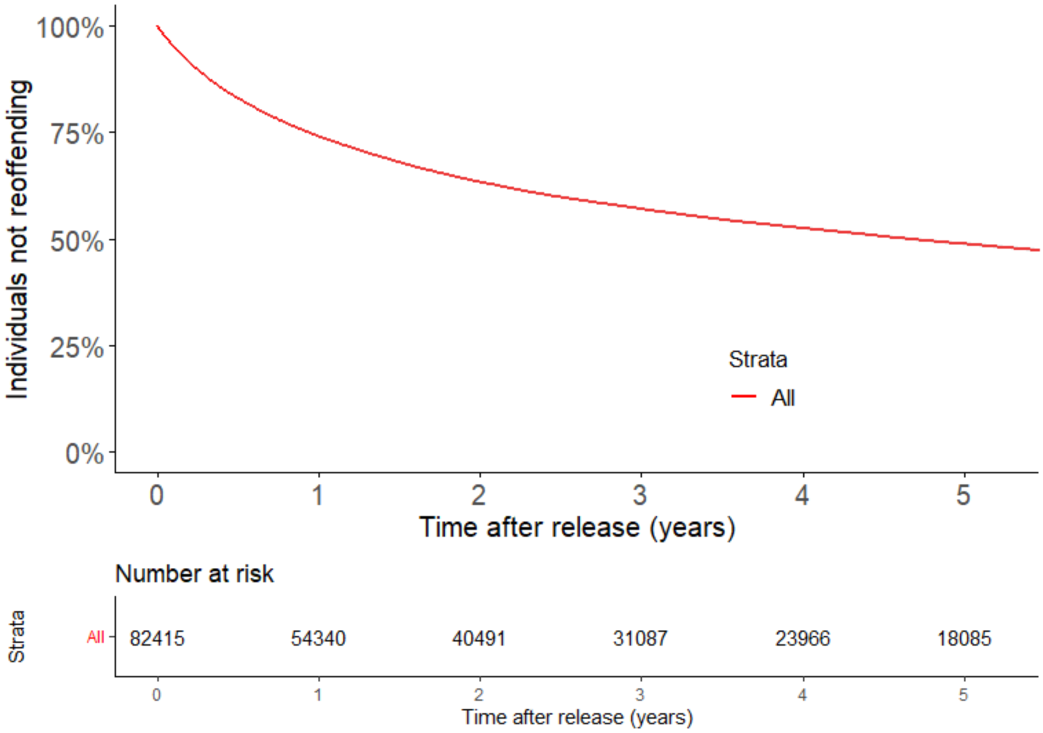
<u>Women (N = 11,772)</u>		General reoffending		Violent reoffending	
	N of individuals	N with outcome	Hazard ratio (95% CI)	N with outcome	Hazard ratio (95% CI)
Age					
18-24 years	3,616	1,311 (36.3%)	1	322 (8.9%)	1
25-39 years	4,815	2,068 (42.9%)	1.22 (1.13-1.30)	390 (8.1%)	0.89 (0.77-1.03)
≥ 40 years	3,341	1,055 (31.6%)	0.89 (0.82-0.96)	156 (4.7%)	0.57 (0.47-0.70)
Civil status					
Single	1,642	557 (33.9%)	0.85 (0.78-0.93)	106 (6.5%)	0.85 (0.69-1.04)
Married	10,072	3,856 (38.3%)	1	759 (7.5%)	1
Highest education					
< 9 yr	500	231 (46.2%)	1	50 (10.0%)	1
9-12 yr	9,705	3,720 (38.3%)	0.73 (0.64-0.83)	721 (7.4%)	0.67 (0.51-0.90)
≥ 12 yr	1,144	265 (23.2%)	0.41 (0.35-0.49)	46 (4.0%)	0.38 (0.26-0.57)
Employed					
Yes	3,909	826 (21.1%)	0.38 (0.35-0.41)	151 (3.9%)	0.41 (0.34-0.49)
No	7,805	3,587 (46.0%)	1	714 (9.1%)	1
Recipient of income support					
Yes	5,634	2,847 (50.5%)	2.3 (2.16-2.45)	562 (10.0%)	1.96 (1.7-2.25)
No	6,080	1,566 (25.8%)	1	303 (5.0%)	1
Prior criminal history					
Any prior convictions	7,832	3,728 (47.6%)	3.32 (3.06-3.6)	694 (8.9%)	2.04 (1.72-2.4)
No prior convictions	3,940	706 (17.9%)	1	174 (4.4%)	1

Prior violent crime					
Prior conviction for a violent offence	2,373	1,190 (50.1%)	1.69 (1.58-1.81)	345 (14.5%)	2.95 (2.58-3.38)
No convictions for a violent offence	9,399	3,244 (34.5%)	1	523 (5.6%)	1
Prior prison					
Prior imprisonment	1,392	873 (62.7%)	2.39 (2.22-2.58)	114 (8.2%)	1.24 (1.02-1.51)
No prior imprisonment	10,380	3,561 (34.3%)	1	754 (7.3%)	1
Index violent crime					
Yes	4,013	1,178 (29.4%)	0.69 (0.65-0.74)	461 (11.5%)	2.61 (2.28-2.99)
No	7,759	3,256 (42.0%)	1	407 (5.2%)	1
Any psychiatric disorder					
Yes	7,062	2,986 (42.3%)	1.68 (1.58-1.79)	632 (8.9%)	2.07 (1.78-2.40)
No	4,710	1,448 (30.7%)	1	236 (5.0%)	1
Any psychiatric disorder (other than substance use)					
Yes	5,486	2,150 (39.2%)	1.25 (1.18-1.33)	494 (9.0%)	1.81 (1.58-2.08)
No	6,286	2,284 (36.3%)	1	374 (5.9%)	1
Schizophrenia spectrum disorder					
Yes	563	247 (43.9%)	1.29 (1.13-1.46)	82 (14.6%)	2.2 (1.75-2.77)
No	11,209	4,187 (37.4%)	1	786 (7.0%)	1
Bipolar disorder					
Yes	340	100 (29.4%)	0.86 (0.71-1.05)	24 (7.1%)	1.23 (0.82-1.85)
No	11,432	4,334 (37.9%)	1	844 (7.4%)	1
Depressive disorder					
Yes	2,037	713 (35.0%)	1.01 (0.93-1.09)	138 (6.8%)	1.05 (0.87-1.26)
No	9,735	3,721 (38.2%)	1	730 (7.5%)	1
Anxiety disorder					
Yes	1,869	745 (39.9%)	1.15 (1.06-1.24)	178 (9.5%)	1.49 (1.26-1.75)
No	9,903	3,689 (37.3%)	1	690 (7.0%)	1
Alcohol use disorder					
Yes	2,961	1,169 (39.5%)	1.18 (1.10-1.26)	322 (10.9%)	2.01 (1.75-2.30)
No	8,811	3,265 (37.1%)	1	546 (6.2%)	1
Drug use disorder					
Yes	3,345	1,825 (54.6%)	2.39 (2.25-2.54)	336 (10.0%)	1.78 (1.55-2.04)
No	8,427	2,609 (31.0%)	1	532 (6.3%)	1
Substance use disorder (alcohol or drug use disorder)					
Yes	4,825	2,266 (47.0%)	1.88 (1.78-2.00)	479 (9.9%)	2.01 (1.76-2.30)
No	6,947	2,168 (31.2%)	1	389 (5.6%)	1
Personality disorder					

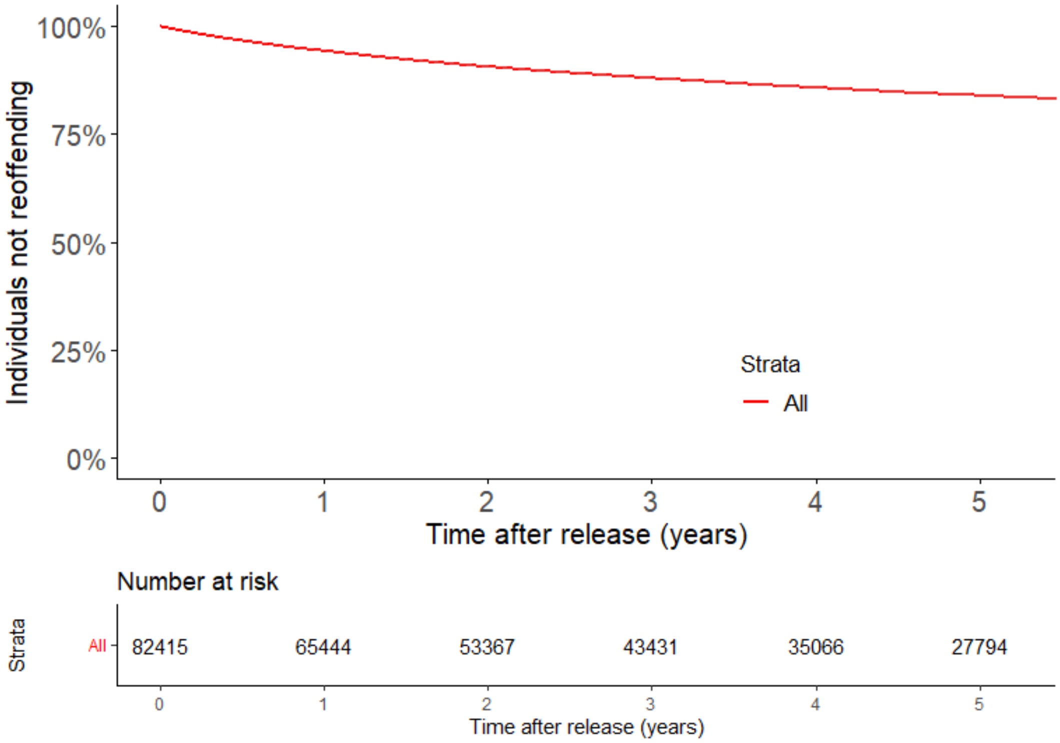
Yes	1,324	613 (46.3%)	1.5 (1.38-1.64)	184 (13.9%)	2.46 (2.09-2.90)
No	10,448	3,821 (36.6%)	1	684 (6.5%)	1
Attention deficit hyperactivity disorder					
Yes	608	210 (34.5%)	1.31 (1.14-1.50)	50 (8.2%)	1.82 (1.37-2.43)
No	11,164	4,224 (37.8%)	1	818 (7.3%)	1
Other developmental or childhood disorder					
Yes	777	352 (45.3%)	1.48 (1.33-1.65)	110 (14.2%)	2.57 (2.10-3.14)
No	10,995	4,082 (37.1%)	1	758 (6.9%)	1
Prior self-harm					
Yes	2,323	904 (38.9%)	1.16 (1.08-1.25)	218 (9.4%)	1.56 (1.34-1.82)
No	9,449	3,530 (37.4%)	1	650 (6.9%)	1

Appendix C5. Kaplan-Meier estimates for recidivism in the cohort of individuals given community sentences

General reoffending in individuals given community sentences.



Violent reoffending in individuals given community sentences.



APPENDIX C6. PAIRWISE COLLINEARITY BETWEEN BASELINE COVARIATES (STUDY 3)

Cramer's V does not show the direction of associations, only the magnitude.

No association is 0, full collinearity is 1.

	Any.disorder	Any.disorder (exl. substance use)	Drug.use	Alcohol.use	Depression	Anxiety	Employed	Income.support	Prior.prison	Age	ADHD	Other.dev	Prior.violence	Prior.crime	Sch.spectrum	Male	Bipolar	Education	Married
Any.disorder (exl. substance use)	0.75																		
Drug.use	0.56	0.33																	
Alcohol.use	0.55	0.28	0.29																
Depression	0.38	0.5	0.17	0.19															
Anxiety	0.37	0.5	0.14	0.12	0.1														
Employed	0.19	0.16	0.22	0.07	0.04	0.06													
Income.support	0.2	0.14	0.25	0.09	0.05	0.08	0.42												
Prior.prison	0.19	0.08	0.25	0.15	0.02	0.06	0.14	0.18											
Age	0.17	0.13	0.14	0.19	0.11	0.09	0.23	0.1	0.32										
ADHD	0.27	0.36	0.16	0.09	0.1	0.07	0.11	0.07	0.02	0.09									
Other.dev	0.27	0.36	0.1	0.07	0.09	0.07	0.1	0.06	0	0.08	0.3								
Prior.violence	0.15	0.09	0.17	0.11	0	0.04	0.14	0.16	0.38	0.07	0.07	0.08							
Prior.crime	0.15	0.09	0.19	0.1	0.02	0.04	0.15	0.18	0.29	0.1	0.06	0.05	0.43						
Sch.spectrum	0.21	0.29	0.18	0.11	0.06	0.06	0.1	0.06	0.05	0.1	0.06	0.1	0.06	0.04					
Male	0.15	0.16	0.1	0.08	0.12	0.1	0.08	0.1	0.09	0.09	0.01	0.03	0.13	0.09	0.04				
Bipolar	0.13	0.18	0.08	0.08	0.04	0.04	0.02	0.01	0.01	0.06	0.1	0.05	0	0.01	0.02	0.06			
Education	0.03	0.03	0.06	0.01	0.02	0.02	0.12	0.14	0.08	0.15	0.04	0.04	0.11	0.13	0.03	0.05	0.04		
Married	0.05	0.03	0.06	0.03	0	0.01	0.1	0.05	0.01	0.3	0.05	0.04	0.07	0.07	0.02	0.04	0	0.17	
Index.violent	0.03	0.02	0.09	0.01	0.01	0	0.04	0.04	0.12	0.11	0.03	0.04	0.04	0.12	0.03	0.09	0	0.01	0.01

APPENDIX C7. POPULATION ATTRIBUTABLE FRACTIONS

Population attributable fraction estimations for the recidivism outcomes during the first 2 years of the follow-up.

Population attributable fraction (% [95% CI])
estimated from Cox regression models adjusted for age

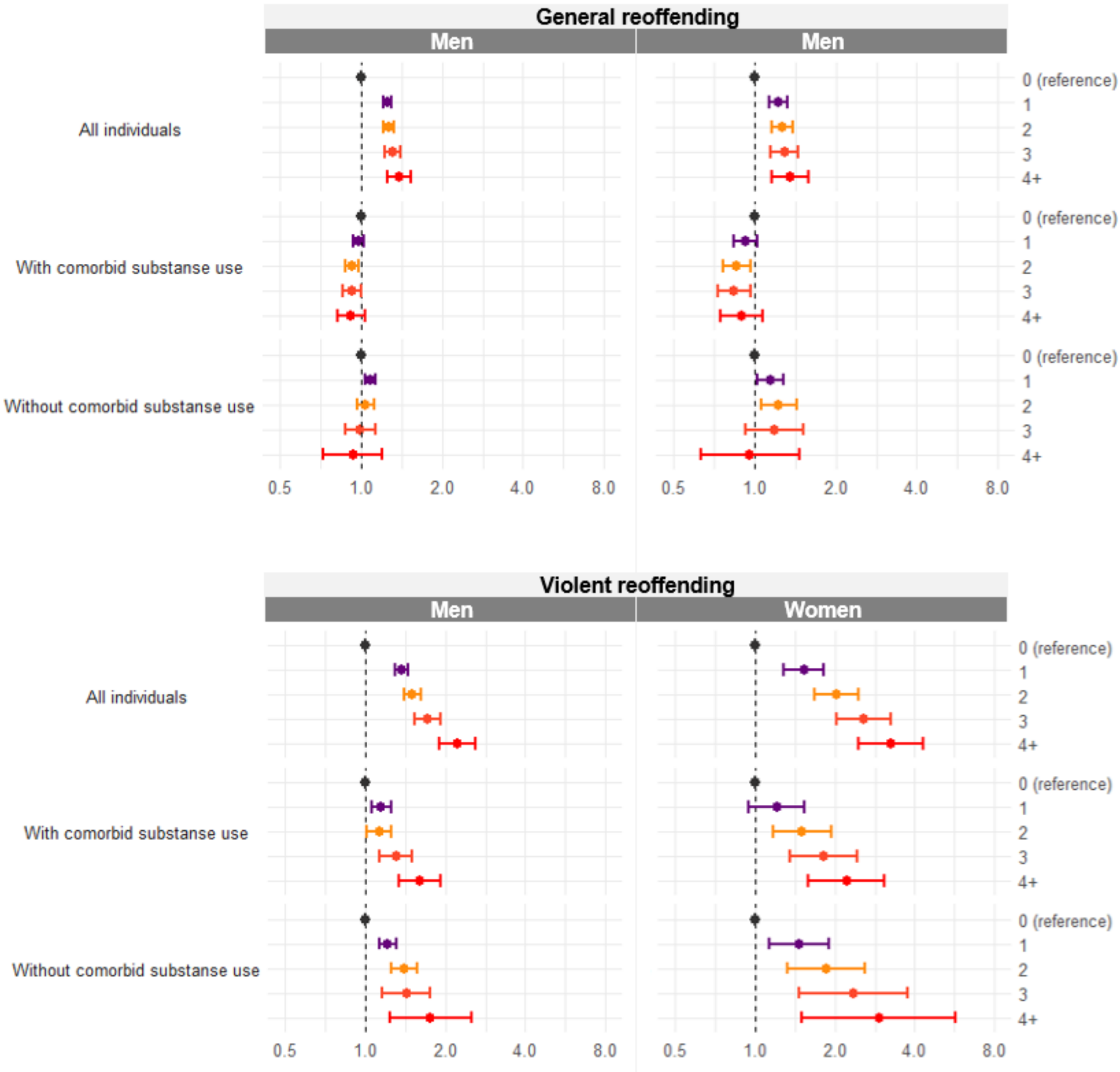
		Any psychiatric disorder (**including** substance use)	Any psychiatric disorder (**excluding** substance use)	Substance use
General reoffending	Men	13.6 (12.9-14.2)	4.8 (4.3-5.3)	13.3 (12.8-13.8)
	Women	25.4 (22.6-28.2)	9.3 (7.0-11.6)	22.4 (20.3-24.5)
Violent reoffending	Men	16.7 (15.3-18.1)	9.5 (8.3-10.6)	13.6 (12.4-14.8)
	Women	40.3 (33.8-46.9)	28.8 (23.1-34.5)	29.7 (24.4-35.0)

Population attributable fraction (% [95% CI])
estimated from Cox regression models adjusted for age, criminal history and sociodemographic factors

		Any psychiatric disorder (**including** substance use)	Any psychiatric disorder (**excluding** substance use)	Substance use
General reoffending	Men	5.5 (4.8-6.3)	0.7 (0.2-1.2)	6.7 (6.2-7.3)
	Women	15.7 (12.5-18.9)	4.8 (2.4-7.2)	14.4 (12.0-16.7)
Violent reoffending	Men	8.3 (6.6-10.0)	4.4 (3.1-5.6)	7.3 (5.9-8.7)
	Women	30.9 (22.7-39.0)	20.5 (13.7-27.2)	24.2 (18-30.4)

APPENDIX C8. NUMBER OF DIAGNOSES AND REOFFENDING

The association between number of different psychiatric diagnosis (other than substance use disorder; maximum of 7 different diagnoses) and recidivism outcome in individuals given community sentences. Individuals without psychiatric diagnosis prior to the start of their sentence serve as a reference group.



APPENDIX D1. STROBE CHECKLIST (MORTALITY)

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

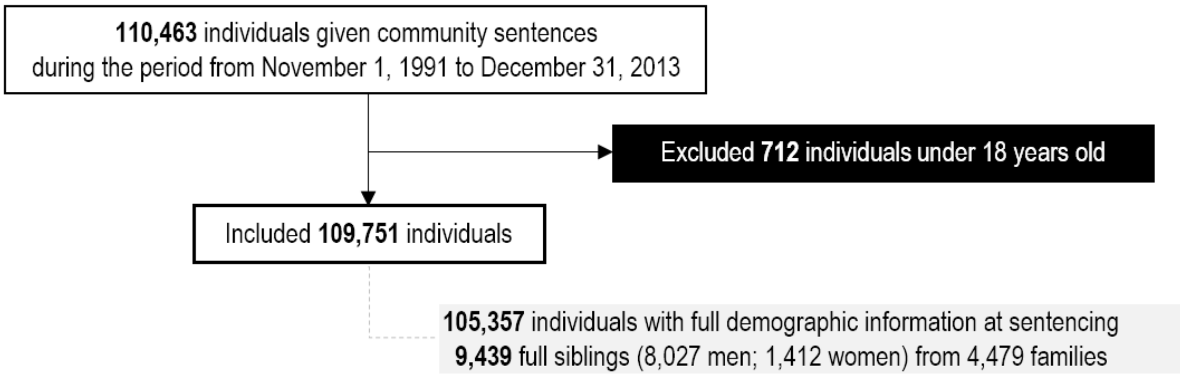
	Item No	Recommendation	Page No
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	91
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	91
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	92-95
Objectives	3	State specific objectives, including any prespecified hypotheses	94-95
Methods			
Study design	4	Present key elements of study design early in the paper	95-97
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	95
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up	96
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed	n/a
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	97-98
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	97, 99
Bias	9	Describe any efforts to address potential sources of bias	99
Study size	10	Explain how the study size was arrived at	96
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	97-98
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	99
		(*b*) Describe any methods used to examine subgroups and interactions	n/a
		(*c*) Explain how missing data were addressed	98
		(*d*) If applicable, explain how loss to follow-up was addressed	98
		(*e*) Describe any sensitivity analyses	99
Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	Table 4-1
		(b) Give reasons for non-participation at each stage	Table 4-1
		(c) Consider use of a flow diagram	Apndx. D2

Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	Table 4-1
		(b) Indicate number of participants with missing data for each variable of interest	Table 4-1
		(c) Summarise follow-up time (eg, average and total amount)	Table 4-1, Apndx. D3
Outcome data	15*	Report numbers of outcome events or summary measures over time	Figure 4-1, Figure 4-3, Figure 4-4
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	Apndx. C4, 101-116
		(*b*) Report category boundaries when continuous variables were categorized	Table 4-1
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Apndx. C4
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	n/a
Discussion			
Key results	18	Summarise key results with reference to study objectives	116
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	119
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	116-119
Generalisability	21	Discuss the generalisability (external validity) of the study results	119
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	n/a

*Give information separately for exposed and unexposed groups.

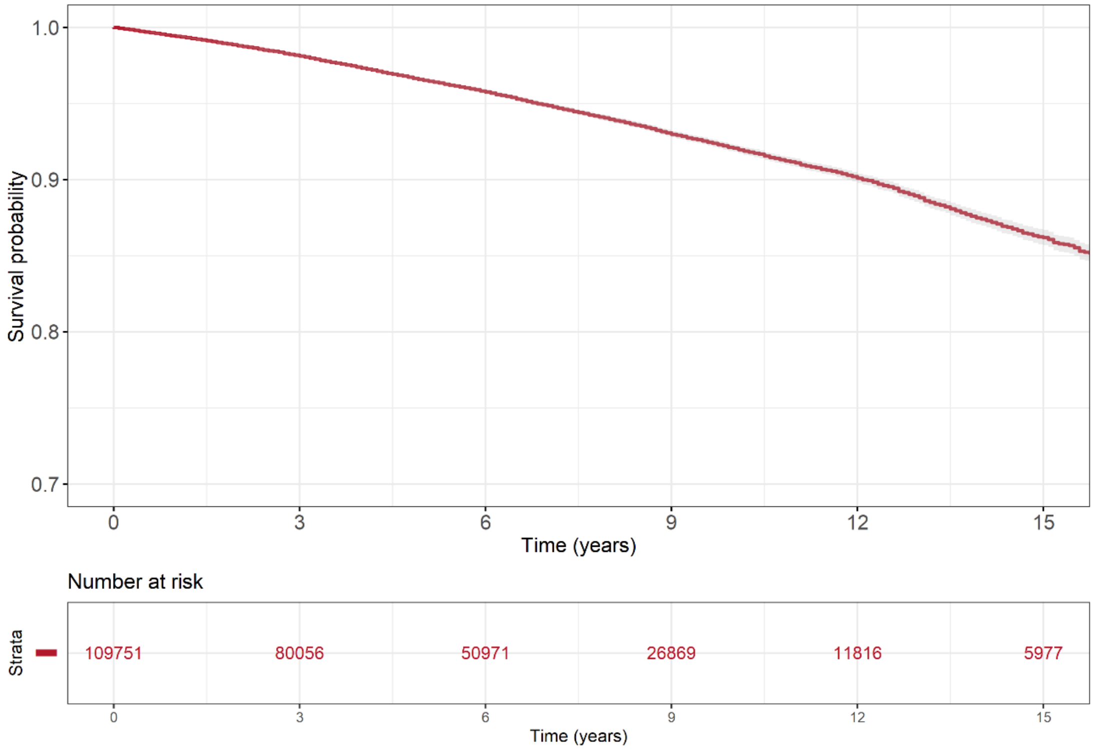
Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

APPENDIX D2. SELECTION PROCESS FOR THE ANALYSIS COHORT

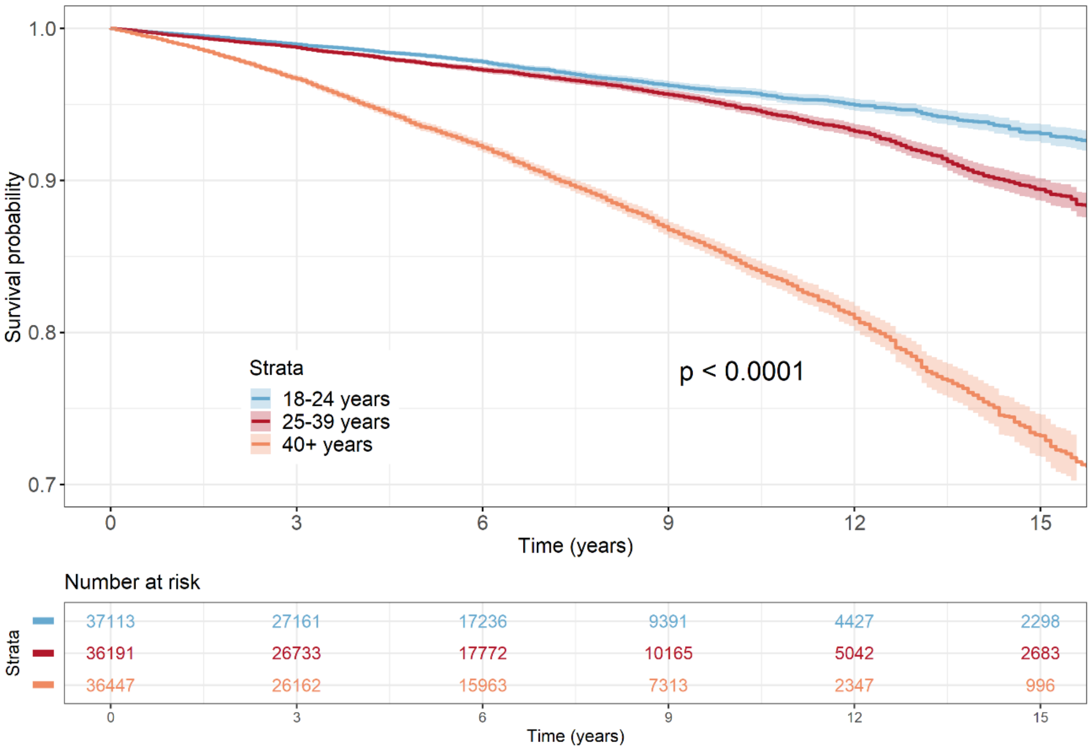


Appendix D3. Kaplan-Meier curves for all-cause mortality in the cohort of individuals given community sentences

Total cohort survival.



Survival by age group.


Survival probability
1.0
0.9
0.8
0.7
Strata
18-24 years
25-39 years
40+ years
p < 0.0001
0
3
6
9
12
15
Time (years)
Number at risk
Strata
37113 27161 17236 9391 4427 2298
36191 26733 17772 10165 5042 2683
36447 26162 15963 7313 2347 996
Time (years)

APPENDIX D4. ASSOCIATION BETWEEN PSYCHIATRIC DISORDERS AND MORTALITY IN INDIVIDUALS GIVEN COMMUNITY SENTENCES ESTIMATED BY A FIXED-EFFECTS SIBLING MODEL

In these analyses, individuals with a given psychiatric disorder were compared to their siblings discordant by the diagnosis. The models were additionally adjusted for age and sex. We reported sibling models' estimations only if there were at least 100 discordant siblings in the cohort.

	Sibling comparison	**HR (95% CI)**	
Previous psychiatric disorder	***Cases / disc. siblings***	***All-cause***	***External cause***
Any psychiatric diagnosis	2,106 / 2,143	1.63 (1.08-2.46)	2.19 (1.27-3.78)
Any psychiatric diagnosis (other than substance use)	1,765 / 1,857	1.27 (0.84-1.91)	1.43 (0.83-2.49)
Schizophrenia spectrum	265 / 313	1.09 (0.45-2.67)	0.98 (0.27-3.53)
Bipolar	85 / 96	-	-
Depression	665 / 743	0.84 (0.44-1.60)	1.18 (0.51-2.75)
Anxiety	799 / 895	1.84 (1.04-3.28)	1.99 (0.92-4.28)
Personality disorder	448 / 509	0.89 (0.46-1.71)	0.77 (0.32-1.85)
Attention-deficit hyperactivity	401 / 438	0.86 (0.29-2.50)	1.16 (0.36-3.75)
Other developmental or childhood	419 / 471	0.98 (0.39-2.42)	0.63 (0.12-3.17)
Substance use	1,830 / 1,910	2.01 (1.32-3.06)	2.82 (1.60-4.99)
Alcohol use	1,291 / 1,382	1.61 (1.05-2.47)	2.13 (1.22-3.71)
Drug use	1,477 / 1,571	2.42 (1.59-3.69)	3.43 (1.93-6.09)
Prior self-harm	913 / 1,023	2.55 (1.52-4.29)	3.46 (1.77-6.79)

APPENDIX D5. UNIVARIATE ASSOCIATION BETWEEN BASELINE SOCIODEMOGRAPHIC/CLINICAL FACTORS AND MORTALITY IN INDIVIDUALS GIVEN COMMUNITY SENTENCES

739 individuals have missing values for marital status, employment, and income support. 4,394 individuals have missing values for education. Hazard ratios were not adjusted.

Total cohort N = 109,751		All-cause mortality		External cause mortality	
	N of individuals	N with outcome	Hazard ratio (95% CI)	N with outcome	Hazard ratio (95% CI)
Age					
18-24 years	37,113	1,013 (2.7%)	1	861 (2.3%)	1
25-39 years	36,191	1,443 (4.0%)	1.36 (1.26-1.47)	968 (2.7%)	1.09 (0.99-1.19)
≥ 40 years	36,447	3,293 (9.0%)	3.93 (3.66-4.22)	880 (2.4%)	1.19 (1.09-1.31)
Sex					
Male	94,221	5,096 (5.4%)	1.18 (1.09-1.29)	2,396 (2.5%)	1.17 (1.04-1.32)
Female	15,530	653 (4.2%)	1	313 (2.0%)	1
Civil status					
Single	14,673	848 (5.8%)	1.27 (1.18-1.37)	230 (1.6%)	0.67 (0.59-0.77)
Married	94,339	4,875 (5.2%)	1	2,469 (2.6%)	1
Highest education					
< 9 yr	5,781	552 (9.5%)	1	142 (2.5%)	1
9-12 yr	8,884	444 (5.0%)	0.69 (0.61-0.78)	141 (1.6%)	0.83 (0.65-1.04)
> 12 yr	90,692	4,555 (5.0%)	0.53 (0.49-0.58)	2,307 (2.5%)	1.04 (0.88-1.23)
Employed					
Yes	44,184	1,479 (3.3%)	0.64 (0.61-0.68)	655 (1.5%)	0.58 (0.53-0.63)
No	64,828	4,244 (6.5%)	1	2,044 (3.2%)	1
Recipient of income support					
Yes	39,755	2,567 (6.5%)	0.99 (0.94-1.04)	1,449 (3.6%)	1.48 (1.37-1.60)
No	69,257	3,156 (4.6%)	1	1,250 (1.8%)	1
Prior criminal history					
Any prior convictions	85,999	5,093 (5.9%)	1.46 (1.34-1.58)	2,487 (2.9%)	2.20 (1.92-2.53)
No prior convictions	23,752	656 (2.8%)	1	222 (0.9%)	1
Prior violent crime					
Prior conviction for a violent offence	40,322	2,683 (6.7%)	1.19 (1.13-1.26)	1,344 (3.3%)	1.37 (1.27-1.48)
No prior convictions for a violent offence	69,429	3,066 (4.4%)	1	1,365 (2.0%)	1
Prior prison					

Prior imprisonment	28,123	2,581 (9.2%)	1.76 (1.67-1.86)	1,118 (4.0%)	1.56 (1.44-1.68)
No prior imprisonment	81,628	3,168 (3.9%)	1	1,591 (1.9%)	1
Index violent crime					
Yes	44,944	1,748 (3.9%)	0.79 (0.75-0.84)	877 (2.0%)	0.85 (0.79-0.92)
No	64,807	4,001 (6.2%)	1	1,832 (2.8%)	1
Has pseudoreconviction (an unprosecuted offence)					
Yes	13,942	955 (6.8%)	0.98 (0.92-1.05)	585 (4.2%)	1.40 (1.28-1.54)
No	95,809	4,794 (5.0%)	1	2,124 (2.2%)	1
Any psychiatric disorder					
Yes	48,346	3,881 (8.0%)	2.95 (2.79-3.12)	1,895 (3.9%)	3.29 (3.03-3.57)
No	61,405	1,868 (3.0%)	1	814 (1.3%)	1
Any psychiatric disorder (other than substance use)					
Yes	31,748	2,164 (6.8%)	1.92 (1.82-2.02)	1,161 (3.7%)	2.34 (2.17-2.52)
No	78,003	3,585 (4.6%)	1	1,548 (2.0%)	1
Schizophrenia spectrum disorder					
Yes	3,669	332 (9.0%)	1.91 (1.71-2.13)	164 (4.5%)	1.99 (1.70-2.34)
No	106,082	5,417 (5.1%)	1	2,545 (2.4%)	1
Bipolar disorder					
Yes	1,471	106 (7.2%)	2.16 (1.78-2.62)	57 (3.9%)	2.38 (1.83-3.09)
No	108,280	5,643 (5.2%)	1	2,652 (2.4%)	1
Depressive disorder					
Yes	10,348	690 (6.7%)	1.85 (1.7-2)	369 (3.6%)	2.07 (1.86-2.31)
No	99,403	5,059 (5.1%)	1	2,340 (2.4%)	1
Anxiety disorder					
Yes	9,883	646 (6.5%)	1.53 (1.41-1.66)	354 (3.6%)	1.79 (1.60-2.00)
No	99,868	5,103 (5.1%)	1	2,355 (2.4%)	1
Alcohol use disorder					
Yes	22,968	2,506 (10.9%)	3.14 (2.98-3.31)	1,017 (4.4%)	2.43 (2.25-2.63)
No	86,783	3,243 (3.7%)	1	1,692 (1.9%)	1
Drug use disorder					
Yes	21,264	2,128 (10.0%)	2.41 (2.29-2.55)	1,254 (5.9%)	3.56 (3.30-3.84)
No	88,487	3,621 (4.1%)	1	1,455 (1.6%)	1
Substance use disorder (alcohol or drug use disorder)					
Yes	34,918	3,432 (9.8%)	3.29 (3.12-3.47)	1,656 (4.7%)	3.49 (3.23-3.77)
No	74,833	2,317 (3.1%)	1	1,053 (1.4%)	1
Personality disorder					
Yes	5,552	594 (10.7%)	2.07 (1.9-2.25)	310 (5.6%)	2.33 (2.07-2.63)
No	104,199	5,155 (4.9%)	1	2,399 (2.3%)	1
Attention deficit hyperactivity disorder					
Yes	4,778	152 (3.2%)	1.42 (1.21-1.67)	118 (2.5%)	2.20 (1.83-2.65)
No	104,973	5,597 (5.3%)	1	2,591 (2.5%)	1

Other developmental or childhood disorder						
	Yes	4,880	242 (5.0%)	1.21 (1.06-1.38)	118 (2.5%)	2.20 (1.83-2.65)
	No	104,871	5,507 (5.3%)	1	2,591 (2.5%)	1
Prior self-harm						
	Yes	10,876	880 (8.1%)	2.32 (2.16-2.5)	541 (5.0%)	3.12 (2.84-3.43)
	No	98,875	4,869 (4.9%)	1	2,168 (2.2%)	1

Appendix D6. Pairwise collinearity between baseline covariates (study 4)

Cramer's V does not show the direction of associations, only the magnitude.
No association is 0, full collinearity is 1.

	Any.disorder	Any.disorder (exl. substance use)	Drug.use	Alcohol.use	Depression	Age	Employed	Prior.self.harm	Income.support	Anxiety	Prior.prison	Prior.violence	ADHD	Prior.crime	Other.dev	Sch.spectrum	Education	Male	Married	Bipolar
Any.disorder (exl. substance use)	0.72																			
Drug.use	0.55	0.31																		
Alcohol.use	0.58	0.28	0.28																	
Depression	0.36	0.51	0.16	0.19																
Age	0.17	0.12	0.15	0.24	0.11															
Employed	0.2	0.15	0.22	0.08	0.04	0.22														
Prior.self.harm	0.26	0.27	0.26	0.22	0.19	0.07	0.08													
Income.support	0.19	0.13	0.25	0.07	0.04	0.18	0.4	0.09												
Anxiety	0.35	0.49	0.13	0.12	0.1	0.08	0.05	0.11	0.07											
Prior.prison	0.18	0.07	0.23	0.16	0.01	0.31	0.16	0.05	0.18	0.04										
Prior.violence	0.15	0.1	0.18	0.11	0	0.11	0.16	0.07	0.18	0.04	0.38									
ADHD	0.24	0.33	0.15	0.06	0.09	0.12	0.1	0.12	0.07	0.07	0	0.07								
Prior.crime	0.15	0.08	0.19	0.1	0.02	0.1	0.16	0.06	0.18	0.04	0.31	0.4	0.05							
Other.dev	0.24	0.34	0.1	0.05	0.07	0.11	0.1	0.11	0.06	0.07	0.01	0.07	0.29	0.04						
Sch.spectrum	0.21	0.29	0.17	0.11	0.06	0.09	0.1	0.09	0.05	0.06	0.04	0.07	0.05	0.04	0.09					
Education	0.05	0.04	0.1	0.04	0.04	0.25	0.27	0.04	0.21	0.01	0.09	0.16	0.1	0.17	0.1	0.02				
Male	0.14	0.15	0.1	0.07	0.11	0.08	0.07	0.13	0.08	0.09	0.11	0.14	0	0.09	0.03	0.03	0.06			
Married	0.05	0.04	0.09	0.01	0	0.32	0.11	0.04	0.1	0	0.02	0.1	0.06	0.09	0.05	0.03	0.21	0.03		
Bipolar	0.13	0.18	0.07	0.08	0.04	0.06	0.02	0.09	0	0.04	0	0	0.08	0.01	0.04	0.02	0.05	0.06	0	
Index.violent	0.03	0.04	0.07	0.02	0.01	0.15	0.03	0.01	0.02	0.01	0.12	0.06	0.04	0.11	0.05	0.05	0.03	0.09	0.04	0.01

Appendix D7. Estimation of the direct effect of individual psychiatric disorders on non-external cause mortality (ICD-10 chapters I-XVIII) relative to measured covariates.

	Adjusted Cox regression models, HR (95% CI)			
Previous psychiatric disorder	***Model 1. Age + sex***	***Model 2. Age + sex + sociodemographic factors***	***Model 3. Age + sex + sociodemographic factors + criminal history***	***Model 4. Age + sex + sociodemographic factors + criminal history + history of self-harm***
Any psychiatric diagnosis	2.03 (1.88-2.19)	1.89 (1.75-2.05)	1.86 (1.72-2.02)	1.79 (1.65-1.94)
Any psychiatric diagnosis (other than substance use)	1.37 (1.27-1.48)	1.27 (1.17-1.37)	1.26 (1.16-1.36)	1.17 (1.08-1.27)
Schizophrenia spectrum	1.43 (1.22-1.67)	1.28 (1.09-1.50)	1.28 (1.09-1.51)	1.23 (1.05-1.44)
Bipolar	1.18 (0.89-1.56)	1.06 (0.80-1.42)	1.07 (0.80-1.43)	0.99 (0.74-1.32)
Depression	1.13 (1.01-1.27)	1.06 (0.94-1.19)	1.06 (0.94-1.19)	0.95 (0.85-1.08)
Anxiety	1.30 (1.15-1.47)	1.27 (1.12-1.43)	1.25 (1.11-1.42)	1.21 (1.07-1.37)
Personality disorder	1.51 (1.33-1.70)	1.34 (1.18-1.51)	1.29 (1.14-1.47)	1.20 (1.06-1.37)
Attention-deficit hyperactivity	1.83 (1.30-2.58)	1.56 (1.09-2.23)	1.54 (1.07-2.20)	1.41 (0.98-2.02)
Other developmental or childhood	1.64 (1.31-2.06)	1.53 (1.21-1.92)	1.52 (1.21-1.91)	1.38 (1.10-1.74)
Substance use	2.20 (2.05-2.37)	2.07 (1.92-2.23)	2.03 (1.88-2.20)	1.96 (1.81-2.12)
Alcohol use	2.12 (1.97-2.27)	2.01 (1.86-2.16)	1.97 (1.82-2.12)	1.90 (1.75-2.05)
Drug use	1.93 (1.78-2.09)	1.74 (1.60-1.90)	1.70 (1.55-1.85)	1.59 (1.45-1.74)

APPENDIX E1. TRIPOD CHECKLIST FOR THE PREDICTION MODEL DEVELOPMENT

Section/ topic item	Item page	Checklist	Page
Title and abstract			
Title	1	Identify the study as developing and/or validating a multivariable prediction model, the target population, and the outcome to be predicted.	121
Abstract	2	Provide a summary of objectives, study design, setting, participants, sample size, predictors, outcome, statistical analysis, results, and conclusions.	121
Introduction			
Background and objectives	3a	Explain the medical context (including whether diagnostic or prognostic) and rationale for developing or validating the multivariable prediction model, including references to existing models.	122
	3b	Specify the objectives, including whether the study describes the development or validation of the model or both.	126
Methods			
Source of data	4a	Describe the study design or source of data (e.g., randomized trial, cohort, or registry data), separately for the development and validation data sets, if applicable.	133
	4b	Specify the key study dates, including start of accrual; end of accrual; and, if applicable, end of follow-up.	133
Participants	5a	Specify key elements of the study setting (e.g., primary care, secondary care, general population) including number and location of centres.	133-134
	5b	Describe eligibility criteria for participants.	133-134
	5c	Give details of treatments received, if relevant.	NA
Outcome	6a	Clearly define the outcome that is predicted by the prediction model, including how and when assessed.	135
	6b	Report any actions to blind assessment of the outcome to be predicted.	NA
Predictors	7a	Clearly define all predictors used in developing or validating the multivariable prediction model, including how and when they were measured.	Apndx. E2
	7b	Report any actions to blind assessment of predictors for the outcome and other predictors.	NA
Sample size	8	Explain how the study size was arrived at.	133-134
Missing data	9	Describe how missing data were handled (e.g., complete-case analysis, single imputation, multiple imputation) with details of any imputation method.	138-139
Statistical analysis methods	10a	Describe how predictors were handled in the analyses.	136-137
	10b	Specify type of model, all model-building procedures (including any predictor selection), and method for internal validation.	136-140
	10d	Specify all measures used to assess model performance and, if relevant, to compare multiple models.	141
Risk groups	11	Provide details on how risk groups were created, if done.	NA
Results			

Participants	13a	Describe the flow of participants through the study, including the number of participants with and without the outcome and, if applicable, a summary of the follow-up time. A diagram may be helpful.	Table 5-4
	13b	Describe the characteristics of the participants (basic demographics, clinical features, available predictors), including the number of participants with missing data for predictors and outcome.	Table 5-4
Model development	14a	Specify the number of participants and outcome events in each analysis.	145-146
	14b	If done, report the unadjusted association between each candidate predictor and outcome.	NA
Model specification	15a	Present the full prediction model to allow predictions for individuals (i.e., all regression coefficients, and model intercept or baseline survival at a given time point).	Apndx. E6
	15b	Explain how to the use the prediction model.	Apndx. E6
Model performance	16	Report performance measures (with CIs) for the prediction model.	151-159
Discussion			
Limitations	18	Discuss any limitations of the study (such as nonrepresentative sample, few events per predictor, missing data).	167-168
Interpretation	19b	Give an overall interpretation of the results, considering objectives, limitations, and results from similar studies, and other relevant evidence.	166-167
Implications	20	Discuss the potential clinical use of the model and implications for future research.	169
Other information			
Supplementary information	21	Provide information about the availability of supplementary resources, such as study protocol, Web calculator, and data sets.	Apndx. E6, 173
Funding	22	Give the source of funding and the role of the funders for the present study.	NA

Appendix E2. Predictive model development protocol

Dynamic prediction of general and violent reoffending in individuals receiving community sentences

1. Study summary

Design. Retrospective cohort study.

Data. Swedish databases linking hospital, crime, family relationships, housing and socio-economic status data.

Participants. The cohort consists of individuals aged 18-65 given community sentences (for any offence) at any point during the study period. We will exclude individuals with pseudoreconvictions.

Study period. We will include people that were given community sentences between 1st January 2001 and 31st December 2013. We will use one sentence per person, chosen at random in case of multiple recorded sentences. Relevant historical data of each individual (including biographical, clinical and criminological data) are available for their lifespan. Outpatient medical history is only available from 2001.

Risk factors. We will include demographic characteristics (including gender, age, educational level, receipt of income support), criminal history, medical history, treatment history (see Table 1 for complete list).

Primary outcomes. The outcomes will be calculated for 37 times points: at the start of a sentence and at each month during the 3-year follow-up period. At each time point, we will access two primary outcomes: (1) Violent reoffending - committing a violent crime within 2 years, (2) general reoffending - committing any crime within 2 years.

Outputs. The main objective is to produce models that estimate the probability of events for the two primary outcomes, with appropriate measures of predictive accuracy.

These models will be used to obtain scoring systems for calculating criminal recidivism risk. The models will be used prospectively to assess an individual's risk at the time of receiving a community sentence and for ***risk monitoring*** during the following supervision period. As part of the study, we will also investigate the effect of desistance from crime on an individual's risk of violent and general reoffending.

2. Statistical Analysis

Statistical analysis will be based on Cox proportional hazards regression with sliding window landmarks (van Houwelingen & Putter, 2011), adjusting for covariates, in each of two scenarios:

(i) violent crime within 2 years from a given landmark,

(ii) any crime within 2 years from a given landmark.

We will use 37 landmarks corresponding to the start of a sentence (landmark 0) and each month of the 3-year follow-up period (landmarks 1-36). Each landmark thus represents the point of risk re-evaluation. The separate datasets will be created for individuals at risk at each landmark time, containing the corresponding values of covariates. Then the landmark datasets will be combined into a single landmark superset. On this superset, we will fit a Cox regression model stratified by landmark. We will use a simple landmark model, which will assume that the effect of the covariates is the same across landmarks. However, the baseline hazard is allowed to vary across landmarks, reflecting the effect of event-free survival (i.e. desistance from crime)

The individuals will be followed up until the occurrence of an outcome event or a censoring event. The censoring events are death, permanent emigration out of Sweden, and, in the case of violent reoffending (scenario i), imprisonment for a non-violent crime.

3. Model development

The development of the current model follows the general approach implemented during the development of the OxRec (Fazel et al., 2016), FoVoX (Wolf et al., 2018), and OxMIV (Fazel, Wolf, Larsson, Mallett, & Fanshawe, 2019) prediction tools.

Risk factors will be considered in four groups. Table 1 below specifies the groups to which each covariate is assigned.

Group 1 consists of covariates measured at the time of a sentence (at baseline). These include sex and prior criminal history. **Group 2** consists of covariates that are measured at baseline but can change value during the follow-up, representing the current status of an individual. They include current age, sociodemographic variables assessed within a year prior to a given landmark, and any history of psychiatric diagnoses prior to a given landmark. For demographic factors, the values for the first three months will be forward-propagated from the baseline. See 'Missing data' for rationale.

Group 1 and Group 2 variables are strongly suspected on the basis of previous research to be associated with the two outcome measures. Group 1 and 2 variables are included in the model by default.

Group 3 consists of covariates measured at baseline that are likely to show an association with the outcomes. These include several sociodemographic measurements and a history of self-harm. **Group 4** consists of covariates measured only during the follow-up period that are likely to show an association with the outcomes. These covariates represent adverse events that can occur during the follow-up period. Group 4 includes triggers for violence (Sariaslan et al., 2016) and psychiatric hospitalisation.

The inclusion of Group 3 and 4 covariates is not required to achieve face validity. Group 3 and 4 covariates will be included if they are significantly associated with the outcomes (see 'Variable selection' for criteria). To ensure a straightforward interpretation of the model's coefficients, interactions between covariates will not be considered.

Table 1. List of covariates

Covariate	Type	Description
GROUP 1. FIXED BASELINE MEASURES (INCLUDED BY DEFAULT)		
Sex	Binary	Male / Female
Any previous offence	Binary	Any prior offence committed before the index offence
Previous violent offence	Binary	Any violent offence committed before the index offence
Violent index offence	Binary	The most recent offence was a homicide, assault, robbery, arson, any sexual offence (rape, sexual coercion, child molestation, indecent exposure, or sexual harassment), illegal threats, or intimidation.
Prior imprisonment	Binary	Any prison sentence served before the index offence
GROUP 2. BASELINE MEASURES DYNAMICALLY UPDATED DURING THE FOLLOW-UP (INCLUDED BY DEFAULT)		
Current age	Continuous	Age in years
Current civil status	Binary	Single/Married. Recorded once a year in November. The record done within a year prior to a given landmark will be used. If the record is missing, then its value will be imputed.
Current years of education	Categorical	< 9 years, 9-11 years, ≥ 12 years. Recorded once a year in November. The record done within a year prior to a given landmark will be used. If the record is missing, then its value will be imputed.
Current employment	Binary	Employment status. Recorded once a year in November. The record is done within a year prior to a given landmark will be used. If the record is missing, then its value will be imputed.
Current receipt of income support	Binary	Receipt of financial support due to low income. Recorded once a year in November. The record done within a year prior to a given landmark will be used. If the record is missing, then its value will be imputed.
Current unstable housing	Binary	More than 3 changes of address during last year. Recorded once a year in November. The record done within a year prior to a given landmark will be used. If the record is missing, then its value will be imputed.
Any prior mental disorder	Binary	Any lifetime psychiatric diagnosis prior to a given landmark. ICD-10 codes: F00-F99.
Prior severe mental disorder	Binary	Any lifetime diagnosis of schizophrenia spectrum or bipolar disorder prior to a given landmark. ICD-10 codes: F20-F29, F30-F31

Prior alcohol use disorder	Binary	Any lifetime diagnosis of alcohol use disorder prior to a given landmark. ICD-10 codes: F10.
Prior drug use disorder	Binary	Any lifetime diagnosis of drug use disorder prior to a given landmark. ICD-10 codes: F11-F19.
GROUP 3. FIXED BASELINE MEASURES (POTENTIALLY INCLUDED)		
Civil status at sentence	Binary	Single/Married. Recorded once a year in November. The record done within a year prior to the index sentence will be used. If the record is missing, then its value will be imputed.
Employment at sentence	Binary	Employment status. Recorded once a year in November. The record done within a year prior to the index sentence will be used. If the record is missing, then its value will be imputed.
Receipt of income support at sentence	Binary	Receipt of financial support due to low income. Recorded once a year in November. The record done within a year prior to the index sentence will be used. If the record is missing, then its value will be imputed.
Unstable housing at sentence	Binary	More than 3 changes of address during last year. Recorded once a year in November. The record done within a year prior to the index sentence will be used. If the record is missing, then its value will be imputed.
History of self-harm or suicide attempts	Binary	Any history of self-harm prior to the index sentence. ICD-10 codes: X60-X84, Y10-Y34.
GROUP 4. EVENTS OCCURING DURING THE FOLLOW-UP PERIOD (POTENTIALLY INCLUDED)		
Being a victim of a violent assault	Binary	Data from medical records. Measured continuously.
Time since the last assault	Continuous	In weeks
Traumatic brain injury (TBI)	Binary	Measured continuously. ICD codes: S01, S02.0-S02.3, S02.7-S02.9, S04, S06.0-S06.9, S07.0-S07.1, S07.8-S07.9, S09.7-S09.9, T01.0, T02.0, T04.0, T06.0, T90.1-T90.2, T90.4-T90.5, T90.8-T90.9
Time since the last TBI	Continuous	In weeks
Self-harm episode	Binary	Self-harm episode or a suicide attempt during a follow-up. Measured continuously. ICD codes: X60-X84, Y10-Y34
Time since the last self-harm episode	Continuous	In weeks

Injuries from other causes (not TBI, not SH)	Binary	Measured continuously. ICD codes: V00-V99, W00-W99, X00-X59
Time since the last injury	Continuous	In weeks
Any psychiatric hospitalisation	Binary	Measured continuously. Inpatient only. ICD-10 codes: F00-F99.
Time since last psychiatric hospitalisation	Continuous	In weeks
Substance intoxication	Binary	Measured continuously. ICD codes: F10.0-F12.0, F14.0-F16.0, F19.0
Time since the last intoxication	Continuous	In weeks

4. Variable selection

All variables from Groups 1 and 2 will be included in the model by default. To further select predictors from Groups 3 and 4, we will implement backwards selection on combined estimates from all imputed datasets (Wood et al., 2008). As the inclusion criteria, we specify the p-value of 0.157 as suggested by (Heinze & Dunkler, 2017). We choose the significance of the coefficient as the criterion, as the criteria based on partial likelihood (such as BIC and AIC) may not be appropriate to use with robust variance estimators.

This strategy recognises that the final model must demonstrate face validity whilst simultaneously allowing the inclusion of additional risk factors if they show an association with outcome variables. The variables are considered in four groups in this way to recognise that a parsimonious model is preferable (i.e. easier to use in practice), provided that it has an acceptable predictive ability.

In addition, the performance of the fitted model will be compared with two other models. First, *the standard baseline model* will be trained using only the information available at a time of a sentence. For this model, the values of covariates will be fixed at their baseline level and will not be updated during the follow-up. Also, this model will only incorporate the baseline hazard estimated at time 0 of the follow-up. Thus, the standard baseline model corresponds to the basic approach to developing a risk prediction model.

Second, *the landmark model with fixed covariates* will be fitted using the information available at all landmark times but without a dynamic update of the covariates' values. For this model, the values of covariates will be fixed at their baseline values and will not change during the follow-up. The occurrence of the trigger will not be considered either. However, the landmark model with fixed covariates will incorporate separate baseline hazard estimations for each landmark time. This model corresponds to the basic approach of developing a dynamic risk prediction model.

No variable selection will be performed for the standard baseline model and the landmark model with fixed covariates. They will consist only of covariates included by default. Comparing the two models will allow the estimation of the effect of the desistance from offending. In addition, the comparison of the two models with the fully dynamic model will be used to assess the prognostic usefulness of the dynamic covariates.

5. Splitting

The data will be divided into a derivation set (~80% of data) and an external validation (~20% of data). The test sample will be derived based on the residential geographical location of the individual at the time of the sentence. Regions are primarily based on the counties of Sweden, derived from the first two digits of the SAMS code. Exceptions are the municipalities of Gothenburg and Malmö, which are separated from their respective counties, and Stockholm municipality, which is separated from its county and sub-divided into northern and southern parts by identifying each SAMS area with the historical province in which it is located. Regions are allocated to four groups, which are proxy measures of urban/rural status: the four urban areas (Group I); the three counties in which the urban areas are located (Group II); four counties with low population (Group III); and all other counties (Group IV). The external validation set will be selected randomly, with equal probability, choosing one region from each of the first three groups and selecting sequentially from the fourth group.

This approach was implemented before during the development of OxMIV to maximise the number of outcomes event in the external validation sample (Fazel, Wolf, Larsson, et al., 2019). In addition, it will increase the representativeness of the external validation sample.

Table 2. The rule for splitting the full dataset into derivation and external validation samples.

Group I	Group II	Group III	Group IV
Major urban centres	**Counties with major urban centres removed**	**Counties with small population**	**Counties with medium population**
1 Stockholm City North	1 Stockholm County Other	7 Kronoberg	3 Uppsala
1 Stockholm City South	12 Skåne Other	9 Gotland	4 Södermanland
12 Malmö	14 Västra Götaland Other	10 Blekinge	5 Östergötland
14 Gothenburg		23 Jämtland	6 Jönköping
			8 Kalmar
			13 Halland
			17 Värmland
			18 Örebro
			19 Västmanland
			20 Dalarna
			21 Gävleborg
			22 Västernorrland
			24 Västerbotten
			25 Norrbotten

Note: Regions will be selected at random, with equal probability, as follows:

- *One region from Group I*
- *One region from Group II (under the constraint that no more than one region in Groups I and II from the same county can be selected)*
- *One region from Group III*
- *Sequentially select from Group IV until the number of individuals in the external validation sample is equal to or exceeds 20% of the entire dataset.*

6. Missing data

Covariates that have more than 30% missing data will be excluded. Missing data will be imputed via multiple imputations (with ten imputations) using a regression model that uses other risk factors and the outcome variables as explanatory variables (Sterne et al., 2009).

We will impute the values of the sociodemographic records at the time of a sentence (at baseline). In addition, we will impute the values of the sociodemographic covariates missing during the follow-up. All sociodemographic records are recorded once a calendar year for all individuals in Sweden. Therefore, the time from the s tart of the follow-up until the subsequent record will vary across the individuals. To compensate for this effect, in case of missing records for sociodemographic factors during the follow-up, we will use values at baseline for the first three months of the follow-up period. Three months is the median time from the start of the follow-up to the subsequent measurement. For education level, we will forward-propagate the last measurement until the next new available measurement without time restrictions. We will impute all other missing sociodemographic records.

Clinical covariates, including triggers, are measured in an ongoing manner. We assume that we have complete information about clinical covariates in our dataset; thus, no missing value imputation is required for them.

7. Assumptions of the model

We will check the ***proportionality assumption*** over the whole observational period and over different time windows by assessing Schoenfeld residuals' plots. Martingale residual plots with LOESS line will be used to check the ***linearity assumption*** and the appropriate ***functional form*** of the continuous variables (Karlsson, 2016). If any assumptions are violated, appropriate adjustments will be made (which may include changing the functional form of the variable or including interaction terms). In general, the landmark model can mediate the effect of the non-proportionality of covariates' effects by accounting for time-effects (van Houwelingen & Putter, 2011).

The functional form of the time-since-trigger variables will be decided after examining their polynomial spine functional form in univariate complete-case analysis on the derivation dataset.

Again, we assume complete information about clinical covariates in our dataset. The database records are likely underreporting the true prevalence of the clinical covariates. This is one of the limitations of our design.

8. Internal validation

The model's predictive accuracy will be summarised by several summary measures, including the concordance index (**c-index**) and the ***Brier score***. To obtain the model performance estimates corrected for optimism, we will implement Harrell's bias correction method (Harrell et al., 1996; Iba et al., 2021). The final accuracy estimates and the estimates of the model's optimism will be obtained by pooling the results from all iterations. We will derive the 95% confidence intervals empirically from the corresponding quantiles.

The proportions of predicted and observed events at different levels of predicted probability and different time windows will be compared using calibration plots.

9. Evaluation of final model

Estimates of coefficients in the final prediction rule will be combined across imputations, using standard methodology (Barnard & Rubin, 1999).

To test the final model, we will use an external validation (holdout) sample derived based on the residential geographical location of the individual at the time of the sentence. The rationale for the selection of the external validation sample is described in 'Splitting'. The predictive accuracy of the model will be summarised by several summary measures, including the concordance index (***c-index***), the ***Brier score***, and time-dependent ***ROC*** curve plots. The proportions of predicted and observed events at different levels of predicted probability will be compared using a calibration plot.

10. Presentation of findings

The main output of the model will be a predictive probability, indicating the probability of occurrence of a recidivism event in 2 years after the time of the assessment. The estimated coefficients of individual risk factors will be examined with a view to (i) simplifying the prediction rule in order to make it easier to use in practice; and (ii) justifying a categorisation of the predicted probability into risk categories (for example 'Low risk' (<5%), 'Medium risk' (5-10%) and 'High risk'

(>10%)). The latter would also benefit from an assessment of the number and characteristics of individuals that fall into the proposed categories in the model development sample and the validation sample.

11. Generalisability

Variables will be defined in such a way as to help the generalisability of the model to other settings. When the final model is used for prediction, it is possible that some variables that are included as covariates may be missing. We will provide guidelines for using the model for prediction in this scenario, for example, by using the mean imputation of the missing variable.

References

Barnard, J., & Rubin, D. B. (1999). Miscellanea. Small-sample degrees of freedom with multiple imputation. Biometrika, 86(4), 948–955.

Fazel, S., Chang, Z., Fanshawe, T., Långström, N., Lichtenstein, P., Larsson, H., & Mallett, S. (2016). Prediction of violent reoffending on release from prison: derivation and external validation of a scalable tool. The Lancet Psychiatry, 3(6), 535–543. https://doi.org/10.1016/S2215-0366(16)00103-6

Fazel, S., Wolf, A., Larsson, H., Mallett, S., & Fanshawe, T. R. (2019). The prediction of suicide in severe mental illness: development and validation of a clinical prediction rule (OxMIS). Translational Psychiatry, 9(1), 1–10.

Harrell, F. E. J., Lee, K. L., & Mark, D. B. (1996). Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine, 15(4), 361–387. https://doi.org/10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4

Heinze, G., & Dunkler, D. (2017). Five myths about variable selection. Transplant International : Official Journal of the European Society for Organ Transplantation, 30(1), 6–10. https://doi.org/10.1111/tri.12895

Iba, K., Shinozaki, T., Maruo, K., & Noma, H. (2021). Re-evaluation of the comparative effectiveness of bootstrap-based optimism correction methods in the development of multivariable clinical prediction models. BMC Medical Research Methodology, 21(1), 9. https://doi.org/10.1186/s12874-020-01201-w

Karlsson, L. (2016). An Evaluation of Methods for Assessing the Functional Form of Covariates in the Cox Model.

Sariaslan, A., Lichtenstein, P., Larsson, H., & Fazel, S. (2016). Triggers for Violent Criminality in Patients With Psychotic Disorders. JAMA Psychiatry, 73(8), 796–803. https://doi.org/10.1001/jamapsychiatry.2016.1349

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., … Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ, 338, b2393. https://doi.org/10.1136/bmj.b2393

van Houwelingen, H., & Putter, H. (2011). Dynamic Prediction in Clinical Survival Analysis.

Wolf, A., Fanshawe, T. R., Sariaslan, A., Cornish, R., Larsson, H., & Fazel, S. (2018). Prediction of violent crime on discharge from secure psychiatric hospitals: A clinical prediction rule (FoVOx). European Psychiatry : The Journal of the Association of European Psychiatrists, 47, 88–93. https://doi.org/10.1016/j.eurpsy.2017.07.011

Wood, A. M., White, I. R., & Royston, P. (2008). How should variable selection be performed with multiply imputed data? Statistics in Medicine, 27(17), 3227–3246. https://doi.org/10.1002/sim.3177

Appendix E3. Missing data in the total cohort

The rates of missing data for sociodemographic covariates by the time after being sentenced.

Months after sentence	N	Employment	Marital status	Education level	Receipt of income support	House changes
			Derivation sample			
0	43,192	0.8%	0.8%	3.2%	0.8%	1.0%
12	33,868	0.9%	0.9%	2.0%	0.9%	1.0%
24	26,395	0.5%	0.5%	1.5%	0.5%	0.5%
36	19,750	0.4%	0.4%	1.2%	0.4%	0.5%
			External validation sample			
0	16484	0.5%	0.5%	3.0%	0.5%	0.6%
12	12884	0.7%	0.7%	1.9%	0.7%	0.7%
24	9992	0.4%	0.4%	1.4%	0.4%	0.4%
36	7472	0.2%	0.2%	1.3%	0.2%	0.3%

APPENDIX E4. VARIABLE SELECTION RESULTS (FIRST ITERATION)

Variable selection results the subset of candidate variables (covariates of Groups 3 and 4 in the protocol). The model was trained on 10 imputations. The estimates and their standard errors were combined using Rubin's rule. The selection was performed by backwards elimination with the exclusion threshold of p = 0.157. The subset of variables included in the mode by default: sex (male), current age, current civil status, current years of education, current employment, current receipt of income support, current unstable housing, any prior mental disorder, prior severe mental disorder, prior alcohol use disorder, prior drug use disorder, any previous offence, previous violent offence, violent index offence, and prior imprisonment. Variables included in the model by default were not subjected to backwards elimination.

	2-year violent reoffending			2-year general reoffending		
Covariates	Included?	Eliminated at step …	p-value at elimination	Included?	Eliminated at step …	p-value at elimination
Triggers						
Being a victim of a violent assault	✓			✓		
- prior week	✕	4	0.788	✓		
- prior month	✓			✕	3	0.626
Traumatic brain injury (TBI)	✕	9	0.489	✕	2	0.908
- prior week	✕	2	0.767	✕	7	0.371
- prior month	✕	3	0.861	✕	8	0.514
Self-harm episode	✕	10	0.480	✕	10	0.174
- prior week	✕	6	0.709	✕	5	0.471
- prior month	✕	8	0.547	✕	1	0.915
Injuries from other causes	✓			✓		
- prior week	✓			✓		
- prior month	✕	7	0.580	✓		
Any psychiatric hospitalisation	✓			✓		
- prior week	✕	5	0.691	✓		
- prior month	✕	1	0.833	✕	6	0.412
Substance intoxication	✕	12	0.393	✓		
- prior week	✕	15	0.212	✕	4	0.561
- prior month	✓			✓		
Demographics at baseline						
Civil status (single)	✕	13	0.217	✓		
Employed	✓			✓		
Receipt of income support	✕	14	0.217	✓		
Unstable housing	✕	16	0.211	✕	9	0.182
Mental health at baseline						
History of self-harm or suicide attempts	✕	11	0.434	✓		

APPENDIX E5. VARIABLE SELECTION RESULTS (SECOND ITERATION)

Variable results for the subset of reoperationalised candidate variables. The variables included covariates of Groups 3 and 4 with the following changes: (a) 'TBI' and 'Injuries from other causes' were combined into one variable 'TBI or injuries from other causes'; (b) 'Self-harm during the follow-up' and 'History of self-harm of suicide attempts' were combined into 'Any prior self-harm or suicide attempt'; (c) 'Marital status at sentence variable was eliminated. The model was trained on 10 imputations. The estimates and their standard errors were combined using Rubin's rule. The selection was performed by backwards elimination with the exclusion threshold of p = 0.157. The subset of variables included in the mode by default: sex (male), current age, current civil status, current years of education, current employment, current receipt of income support, current unstable housing, any prior mental disorder, prior severe mental disorder, prior alcohol use disorder, prior drug use disorder, any previous offence, previous violent offence, violent index offence, and prior imprisonment. Variables included in the model by default were not subjected to backwards elimination.

	2-year violent reoffending			2-year general reoffending		
Covariates	**Included?**	**Eliminated at step …**	**p-value at elimination**	**Included?**	**Eliminated at step …**	**p-value at elimination**
Triggers						
Being a victim of a violent assault	✓			✓		
- prior week	✕	1	0.840	✓		
- prior month	✓			✕	1	0.608
TBI or injuries from other causes	✓			✓		
- prior week	✓			✓		
- prior month	✕	2	0.792	✓		
Any psychiatric hospitalisation	✓			✓		
- prior week	✕	4	0.850	✓		
- prior month	✕	3	0.694	✕	3	0.338
Substance intoxication	✕	6	0.398	✓		
- prior week	✕	9	0.197	✕	2	0.551
- prior month	✓			✓		
Mental health (current)						
Any prior self-harm or suicide attempt	✕	5	0.603	✓		
Demographics at baseline						
Employed	✓			✓		
Receipt of income support	✕	7	0.213	✓		
Unstable housing	✕	8	0.212	✕	4	0.176

APPENDIX E6. FORMULA

The model estimates the probability of an individual to commit a new offence or a new violent offence in the next 2 years after the time of an assessment. The assessment can be conducted at any point within 3 years after an individual received a community sentence. The probability of reoffending is calculated using the formula:

$$P(\text{reoffending}) = 1 - \exp(-BH_{LM})\text{^}\exp(\textstyle\sum \text{beta} \times RF),$$

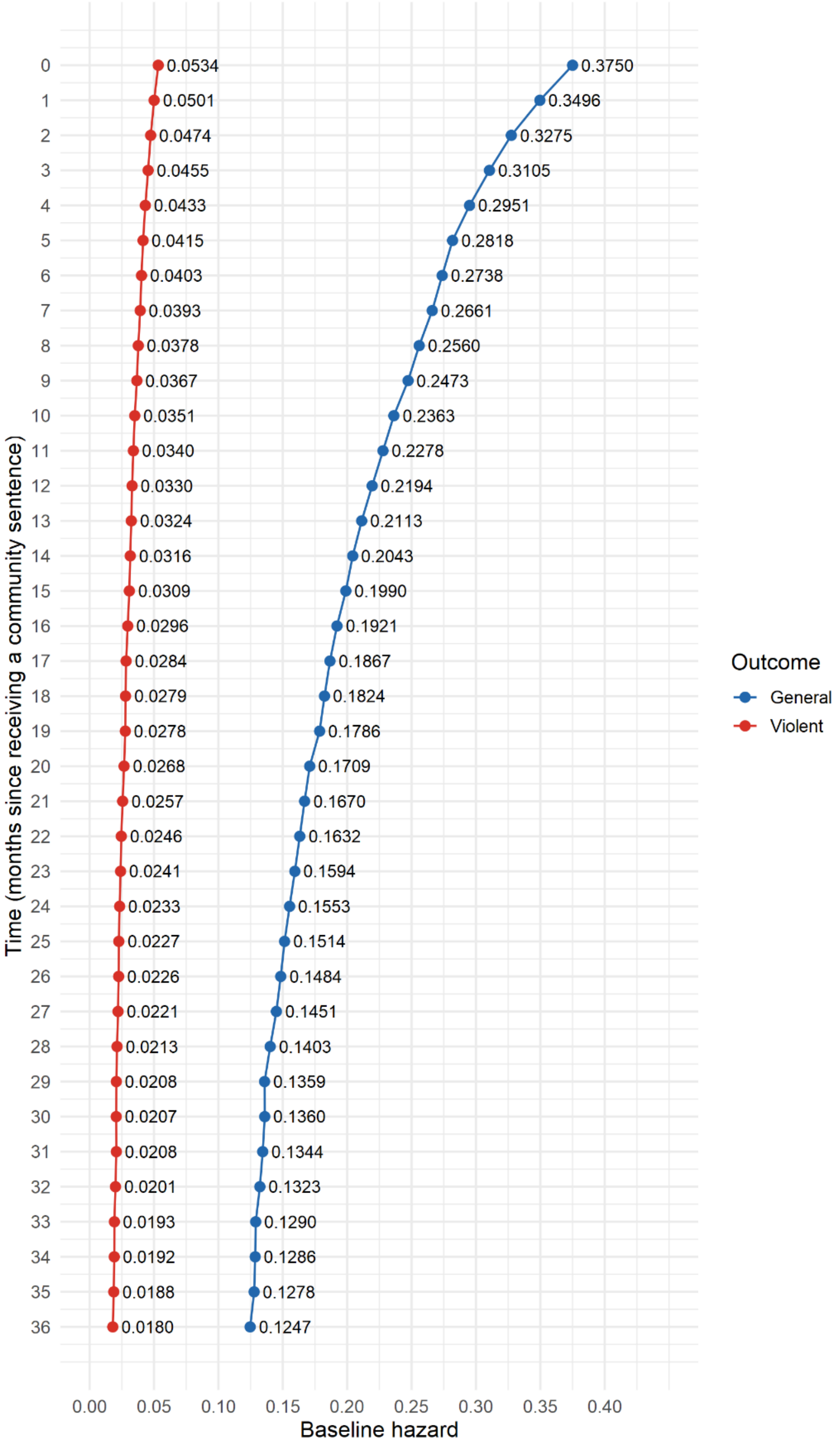
where BH_{LM} is the baseline hazard corresponding to a given landmark (in months since sentence), beta representing coefficients for the corresponding outcome, and RF is a presence or an absence of a given risk factor (coded as 1 or 0 respectively). The list of risk factors and corresponding coefficients are in the table below. The relevant baseline hazard estimations are provided in the plots below the table. General and violent reoffending are estimated using the same formula, but with different coefficients and baseline hazards. Note that some variables are only used for general offending predictions.

Baseline hazards estimates correspond to the time of the assessment. To estimate the risk at the time of receiving a community sentence, use baseline hazard corresponding to landmark 0. To estimate the risk at 6 months after receiving a community sentence, use the landmark 6.

	Coefficients (beta)	
Variables	**2-year violent reoffending**	**2-year general reoffending**
Block 1. Current status		
This block includes (a) items that denote the current status (i.e. current employment, housing, etc.) (b) items that code past events that might have happened at any point before the current assessment (i.e. receiving a psychiatric diagnosis)		
1.1 Demographics		
Sex (male)	0.618	0.408
Age (in years)	-0.034	-0.025
Civil status (single)	-0.135	-0.059
Education level:		
• 9-11 years of educations (finished high school)	-0.197	-0.107
• 12 or more years of education (undergraduate and further)	-0.514	-0.354
Currently employed	-0.215	-0.213
Unstable housing situation	0.080	0.024
Receipt of income support	0.311	0.187
1.2 Mental health history		

History of any psychiatric disorder (excluding drug or alcohol use)	0.178	-0.023
History of severe psychiatric disorder (schizophrenia spectrum or bipolar disorder)	0.140	-0.070
History of alcohol use disorder	0.249	0.066
History of drug use disorder	0.072	0.393
History of self-harm or suicide attempts	NA	-0.055
1.3 Criminal history		
Any prior offence (excluding the index offence)	0.368	0.565
Any prior violent offence (excluding the index offence)	0.602	0.168
Index violent offence	0.579	-0.129
Prior imprisonment	0.380	0.349
Block 2. Status at sentence		
This block includes items that code an individual's status at the time of receiving a community sentence.		
Employed at the time of a sentence	-0.177	-0.211
Receipt of income support at the time of a sentence	NA	0.142
Block 3. Triggers and stressful events		
This block includes items that code adverse events, which might have occurred since the time of receiving a community sentence and before the current assessment. These events could be associated with an ongoing distress and disadaptation.		
3.1 Occurred at any point since receiving a sentence		
Psychiatric hospitalisation	0.326	0.100
Sustaining an injury from any cause (including traumatic brain injury, excluding self-harm)	0.227	0.172
Being a victim of a violent assault	0.301	0.254
Substance intoxication	NA	0.131
3.2 Occurred within last 30 days		
Sustaining an injury from any cause (including traumatic brain injury, excluding self-harm)	NA	0.049
Substance intoxication	0.134	0.086
Being a victim of a violent assault	0.151	0.087
3.3 Occurred within last 7 days		
Psychiatric hospitalisation	NA	0.094
Sustaining an injury from any cause (including traumatic brain injury, excluding self-harm)	-0.180	-0.078

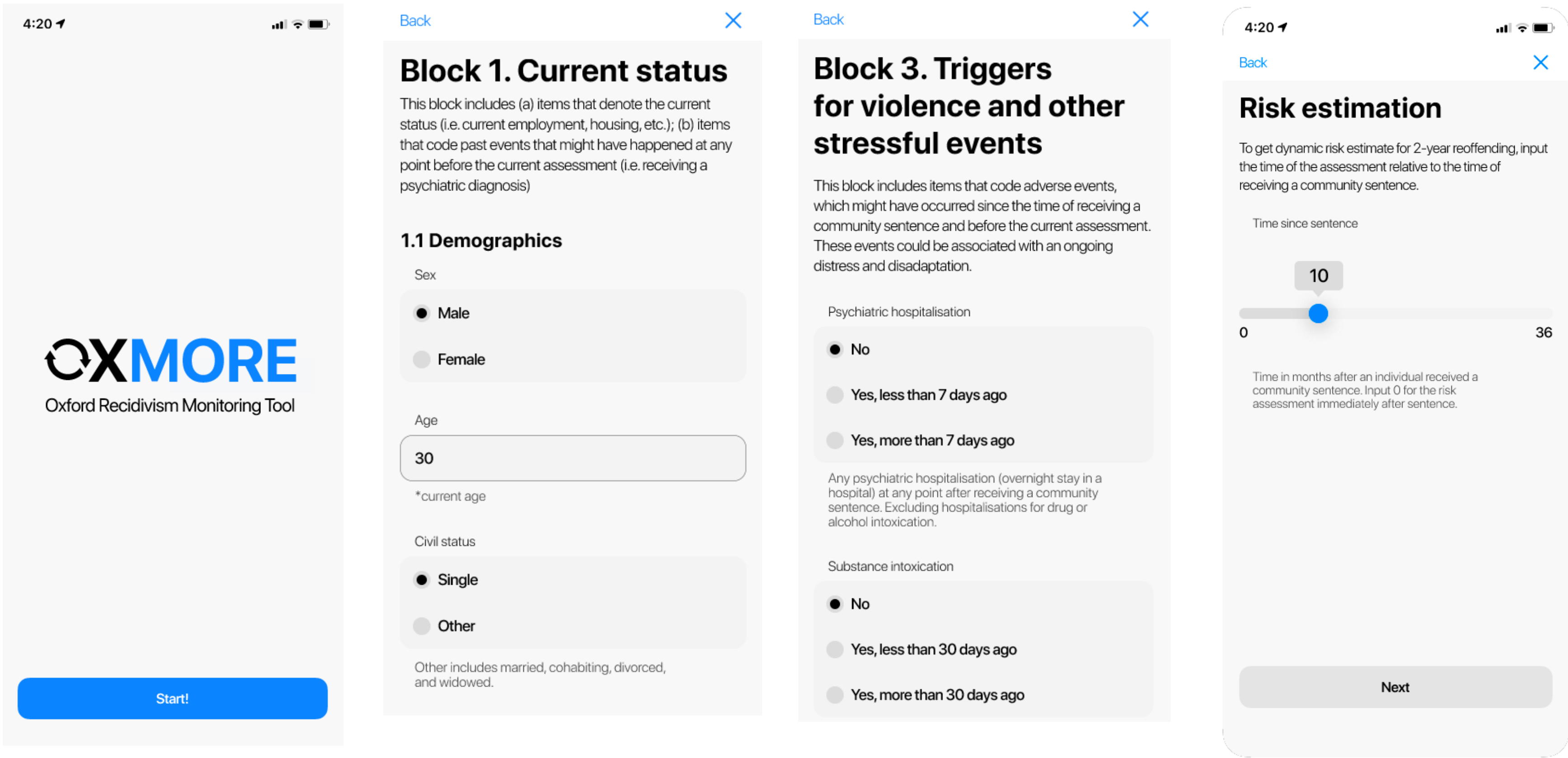
Baseline hazard for 2-year general and violent reoffending (BH_{LM})



APPENDIX E7. OXMORE EXAMPLE INTERFACE FOR MOBILE APPLICATIONS

Designed by Vlad Galanov.

Example startup screen and main tool interface



Appendix E8. R code for prediction model development

This appendix contains R code that was applied to datasets extracted from the Swedish population registers. It does not contain any actual register data, as the data are protected. Some code was adapted from van Houwelingen and Putter (2011), and Hazewinkel (2018). These works were instrumental in guiding my prediction model development process.

The code is provided here to increase transparency of the model development process, and, hopefully, to assist future researchers with the application of landmark modelling to their data. Not all the code written for the model development is provided in this appendix, only the most relevant parts. However, the code is available upon personal request until it becomes fully available online. I apologise in advance to a future code user for some not so elegant coding decisions that I made during the development of the model. They do work, but they are often written not as clearly as I have hoped for. Learning R was a journey for me during my DPhil study. I hope this code will be improved and reused in other scientific projects.

Data extraction and preparation

Loading libraries

```
library(survival)
library(dynpred)
library(Amelia)
library(MASS)
library(haven)
library(survminer)
library(foreign)
library(openxlsx)
library(DescTools)
library(abe)
library(dplyr)
library(data.table)
library(gmodels)
```

Specifying the variables for the model development.

```
#Main outcome + violent reoffending
Outcome <- 'Violent' #General or Violent repffendng
```

```
#Landmark settings
w <- 24 #prediction horizon in months (default is 24 months)
grid <- 1 #setting spacing between landmarks (default is 1 month)
Landmarks <- seq(0,36, by=grid) #setting up landmarks from 0 to year 3  (36 months)


#loading data
#consists of id, start.date, end.date, time difference (days), outcome status, censoring status
data <- read_sas(…)) %>% rename(id = lopnr) %>% filter(YearConviction >= 2007)
```

Selecting the external validation dataset (a holdout dataset) based on the pre-specified criteria.

```
#loading data with with ids and counties codes
full_set <- read_sas(…))
full_set <- full_set %>% rename(id = lopnr) %>% filter(id %in% data$id)

#creating sets with group names
group1 <- c('Malmo', 'Gothebur', 'SH_South', 'SH_North')
group2 <- c('Skane_Ot', 'SH_Other', 'Vastra_G')
group3 <- c('Blekinge', 'Jamtland', 'Kronoberg', 'Gotland')
group4 <- c('Orebro',  'Gavlebo', 'Kalmar',  'Varmlan', 'Norrbot',
            'Halland', 'Ostergo', 'Uppsala', 'Dalarna', 'Jonkopi', 'Vastern', 'Vasterb', 'Soderma', 'Va
stman')

#unique(full_set$GROUP4)


#calculating the size of the external dataset
full_length <- nrow(full_set)
threshold_length <- round(full_length*0.25)
print(paste0('Full length: ', full_length,'. Threshold length: ',threshold_length))

#selecting from groups 1-3
set <- full_set %>%
  filter(GROUP1 == group1[round(runif(1, min = 1, max = length(group1)))]) %>%
  rbind(filter(full_set, GROUP2 == group2[round(runif(1, min = 1, max = length(group2)))])) %>%
  rbind(filter(full_set, GROUP3 == group3[round(runif(1, min = 1, max = length(group3)))]))


#selecting counties from group 4
for (n in 1:length(group4)) {
  if (nrow(set) < threshold_length) {
        random_number <- round(runif(1, min = 1, max = length(group4)))
        set <- set %>% rbind(filter(full_set, GROUP4 == group4[random_number]))
        print(paste0('Step ',n,'. ','County added: ', group4[random_number],'. Sample size: ', nrow(set
)))
  } else {
    print(paste0('Step ',n,'. No new counties added.'))
    break
  }
}

#adding 20% of individuals without SAMS codes
group5 <- full_set %>% filter(GROUP5 == 'NO_DATA')
group5 <- sample_n(group5, round(0.2*length(group5$id)))
set <- set %>% rbind(group5)
```

```
print(paste0('Adding ',round(0.2*nrow(group5)),' individuals without known SAMS code. Final hold-out se
t size: ', nrow(set)))

#creating list of ids that were selected for external validation
holdout_ids <- set$id
saveRDS(holdout_ids, file = paste0(…))
```

Extraction of the follow-up data.

```
holdout_ids <- readRDS(file = paste0(…))


#only retaining those that haven't been selected for hold-out (i.e. external validation dataset)
data <- subset(data, !(data$id %in% holdout_ids)) #selecting ids NOT in hold-out ids

#saving the ids for derivation dataset
derivation_ids <- data$id
saveRDS(derivation_ids, file = …))

#selecting the follow-up data and converting to required format
follow_up  <- data %>%
  select(id, DateRelease, DateOutcome, REC, Censored) %>%
  rename(start.date = DateRelease, end.date = DateOutcome) %>%
  mutate(start.date = as.Date(start.date, origin = "1960-01-01"),  #converting from SAS date
         end.date = as.Date(end.date, origin = "1960-01-01"),
         time.stop = as.numeric(difftime(end.date, start.date, units="days")/30.42)) #follow-up time in
 months
```

Loading fixed covariates.

```
#loading additional fixed covariates and merging with loaded data

#any criminal record
any_prior_crimes <- read_sas(…)) %>%
  rename(id = lopnr)

any_prior_crimes <- filter(any_prior_crimes, id %in% unique(data$id))

data <- merge(data, any_prior_crimes, by = 'id', all.x=TRUE)
data$Prior_crimes[is.na(data$Prior_crimes)] <- 0

#any prior convictions to prison
any_prior_prison <- read_sas(…)) %>%
  rename(id = lopnr)
any_prior_prison <- filter(any_prior_prison, id %in% unique(data$id))

data <- merge(data, any_prior_prison, by = 'id', all.x=TRUE)
data$Prior_prison[is.na(data$Prior_prison)] <- 0

#prior self-harm
self_harm <- read_sas(…) %>%
  rename(id = lopnr)
self_harm <- filter(self_harm, id %in% unique(data$id))


#converting to date-time variable
self_harm$DateSH <- as.Date(self_harm$DATUM, tryFormats = c("%Y%m%d"))
self_harm$DateRelease <- as.Date(self_harm$DateRelease, origin = "1960-01-01")
```

```
#saving past self-harm
self_harm$SH_time<- difftime(self_harm$DateSH, self_harm$DateRelease, units="days")
past_self_harm <- self_harm %>%
  filter(SH_time < 0) %>%
  select(id, SELFHARM) %>%
  distinct()

#merging with self-harm data
data <- merge(data, past_self_harm, by = 'id', all.x=TRUE)
data$SELFHARM[is.na(data$SELFHARM)] <- 0


#selecting and renaming fixed covariates (with some extra fixed covariates for imputation)
fixed_covs <- data %>%
  select(id, Age, GENDER, BYEARR, YearConviction,
         INCOME, INCOMESUPPORT,
         HOUSECHANGES, DAYSUNEMPLOYED, UNEMPLOYMENTBENEFIT,
         EMPLOYED, MARITAL, EDUCATION,
         Immigrant,
         Prior_crimes, Previous_violent_crime, Index_violent,
         Prior_prison,
         SELFHARM,
         SCH, BP, ANX, DEP, PD, OD, OTHER,
         ANY, ANY_noSU, AU, DU, SU) %>%
  rename(Sex = GENDER,
         Birth.year = BYEARR,
         Conviction.year = YearConviction,
         Employed = EMPLOYED,
         Marital.status = MARITAL,
         Education.level = EDUCATION,
         Immigrant.status = Immigrant,
         Income = INCOME,
         Income.support = INCOMESUPPORT,
         Prior.offence = Prior_crimes,
         Prior.violent.offence = Previous_violent_crime,
         Index.violent.offence = Index_violent,
         Prior.prison = Prior_prison,
         House.changes = HOUSECHANGES,
         Days.unemployed = DAYSUNEMPLOYED,
         Unemployment.benefit = UNEMPLOYMENTBENEFIT,
         Prior.self.harm = SELFHARM,
         Any.psych = ANY,
         Any.psych.exlSU = ANY_noSU,
         Alcohol.use = AU,
         Drug.use = DU,
         Substance.use = SU,
         Sch.spectrum = SCH,
         Bipolar = BP,
         Anxiety = ANX,
         Depression = DEP,
         Person.dis = PD,
         Other.develop = OD,
         Uncategorised.dis = OTHER)

#converting some variables into proper format
```

```
fixed_covs$Conviction.year <- as.numeric(fixed_covs$Conviction.year)
```

Loading and pre-processing dynamic covariates

```
#Demographics
dyn_demog <- read_sas(paste0(…)) %>%
  rename(id = lopnr) %>% filter(id %in% unique(data$id))

dyn_demog$Lisa_t <- as.Date(dyn_demog$Lisa_t, tryFormats = c("%Y%m%d"))
#checking N of included people
length(unique(dyn_demog$id))

#selecting within follow-up
dyn_demog <- merge(dyn_demog, follow_up, by.x = 'id', by.y = 'id', all.y = TRUE)
dyn_demog$Lisa_t <-  as.numeric(difftime(dyn_demog$Lisa_t,  dyn_demog$start.date, units="days")/30.42)
dyn_demog <- filter(dyn_demog, (Lisa_t < time.stop) & (Lisa_t > 0))

#checking N of included people
length(unique(dyn_demog$id))
#now we have les people than before, since no for every person there's info available during follow-up


#renaming variables
dyn_demog <- rename(dyn_demog,
                    Dyn.employed = EMPLOYED,
                    Dyn.marital= MARITAL,
                    Dyn.education = EDUCATION,
                    Dyn.housing = HOUSECHANGES,
                    Dyn.days.unemp = DAYSUNEMPLOYED,
                    Dyn.unemp.ben = UNEMPLOYMENTBENEFIT,
                    Dyn.income.sup = INCOMESUPPORT)

glimpse(dyn_demog)
CrossTable(dyn_demog$Dyn.employed, dyn_demog$REC)


#creating a datasets for each demographics
Demog_list <- list() #creating a list to store triggers information

#creating a list of demographics to run the cycle over
demographics <- c('Dyn.employed', 'Dyn.marital', 'Dyn.education', 'Dyn.housing', 'Dyn.days.unemp', 'Dyn
.unemp.ben', 'Dyn.income.sup')

#loop that creates datasets with measuments of demographical variables and their time
for (d in demographics){

  demog <- dyn_demog %>%
    select(id, paste0(d), ends_with('_t')) %>% distinct()

  colnames(demog) <- c('id', paste0(d), paste0(d,'_t'))

  #adding information needed for landmark splitting
  demog <- demog %>%
    merge(follow_up, demog, by.x = 'id', by.y = 'id', all = TRUE) %>%
    select(id, time.stop, REC, paste0(d), paste0(d, '_t')) %>% distinct() %>%
    filter(eval(parse(text=(paste0(d, '_t')))) <= time.stop) %>%
```

```
    filter(eval(parse(text=(paste0(d, '_t')))) >= 0) #filtering the results outside follow-up boundarie
s


  Demog_list[[length(Demog_list)+1]] <- demog #adding demog data to the list
  print(paste0(d,' is done'))

}
```

Loading triggers

```
#Triggers
Trigger_list <- list() #creating a list to store triggers information


#creating a list of new diagnoses and triggers to run the cycle over
triggers <- c('New_psyc_noSU', 'New_SU', 'New_AU','New_DU',
              'New_psyc_sev', 'Psyc_hosp', 'TBI', 'Injury', 'Intoxic', 'Self_harm', 'Victim',
              'InjuryTBI') #injury TBI added after reoprationalisation of the variable

#loop that creates datasets with individual triggers and their time
for (t in triggers) {

  Trigger <- read_sas(paste0(…))
  Trigger <- Trigger %>% filter(lopnr %in% data$id) %>% distinct() %>%
    select(ends_with('lopnr'), ends_with('date'), ends_with('_t')) %>%
    mutate(trigger = 1)

  #renaming columns
  colnames(Trigger) <- c('id', paste0(t, '_date'), paste0(t, '_t'), paste0(t))

  #formatting
  Trigger[[2]] <- as.Date(Trigger[[2]], tryFormats = c("%Y%m%d"))
  Trigger[[3]] <- as.numeric(Trigger[[3]]/30.42) #converting to months

  Trigger <- as.data.frame(Trigger) #converting to data frame format


  #adding information needed for landmark splitting
  Trigger <- Trigger %>%
    merge(follow_up, Trigger, by.x = 'id', by.y = 'id', all = TRUE) %>%
    select(id, time.stop, REC, paste0(t), paste0(t, '_t')) %>% distinct() %>%
    filter(eval(parse(text=(paste0(t, '_t')))) <= time.stop) %>%
    filter(eval(parse(text=(paste0(t, '_t')))) >= 0) #filtering the results outside follow-up boundarie
s


  Trigger_list[[length(Trigger_list)+1]] <- Trigger #adding trigger data to the list
  print(paste0(t,' is done'))

}
```

Creating landmark datasets

```
#cleaning follow-up data for cutting
follow_up_clean <- select(follow_up, -end.date, -start.date, -Censored)
```

```
#iterating over follow-up data, demographics and triggers to create landmark datasets
LM_superset <- data.frame()
for (LM in Landmarks) {

  #FOLLOW-UP
  followupLM <- cutLM(data = follow_up_clean, outcome = list(time='time.stop', status= 'REC'),
                    LM = LM, horizon = LM + w, covs = list(fixed=c('id')),  format = c("wide"), id = 'i
d')

  #reordering columns
  followupLM <- select(followupLM, id, LM, time.stop, REC)
    print(paste0('Follow-up: ',length(unique(followupLM$id))))

  #TRIGGERS
  LM_triggers <- list()
  for (i in 1:length(Trigger_list)) {

    #selecting a trigger dataset from a list
    triggerLM <- Trigger_list[[i]]
    cols <- colnames(triggerLM) #reading the names from the table

    triggerLM <- cutLM(data=triggerLM,
                       outcome=list(time=paste0(cols[2]), status=paste0(cols[3])),
                       LM=LM,
                       horizon = LM + w, #landmark + prediction horizon
                       covs=list(varying=paste0(cols[4])), #varying covariates
                       format='long',id=paste0(cols[1]), rtime=paste0(cols[5]))

    #substituting all NAs for zeros. We have all information about triggers, so if there's NA it means
it didn't happen
    triggerLM[,4][is.na(triggerLM[,4])] <- 0


    LM_triggers[[length(LM_triggers)+1]] <- triggerLM #adding to a landmark list

    print(paste0(i,' of ',length(Trigger_list),' triggers for landmark ', LM,' is done'))

  }


  #DEMOGRAPHICS
  LM_demographics <- list()
  for (i in 1:length(demographics)) {

    #selecting a demographic dataset from a list
    demogLM <- Demog_list[[i]]

    #there is a person with two records with different status. let's remove this person
    #if he is present
    check <- cbind(demogLM[1:2], demogLM[5])
    check <- filter(check, duplicated(check))

    #removing this person
    if (length(check$id) != 0) {demogLM <- demogLM[demogLM$id != check$id, ]}


    cols <- colnames(demogLM) #reading the names from the table
```

```
  demogLM <- cutLM(data=demogLM,
                   outcome=list(time=paste0(cols[2]), status=paste0(cols[3])),
                   LM=LM,
                   horizon = LM + w, #landmark + prediction horizon
                   covs=list(varying=paste0(cols[4])), #varying covariates
                   format='long',id=paste0(cols[1]), rtime=paste0(cols[5]))


  LM_demographics[[length(LM_demographics)+1]] <- demogLM #adding to a landmark list

  print(paste0(i,' of ',length(demographics),' demographics for landmark ', LM,' is done'))

}

#debugging
#aaa <- filter(followupLM, !(followupLM$id %in% demogLM$id) & !(followupLM$id %in% triggerLM$id))


#merging triggers and demographics together
#triggers
dataLM <- merge(LM_triggers[[1]], LM_triggers[[2]], by=c('id','time.stop', 'REC','LM'), all = TRUE)
for (i in 3:length(LM_triggers)) {
  dataLM <- merge(dataLM, LM_triggers[[i]], by=c('id','time.stop', 'REC','LM'), all = TRUE)
}
print(paste0('Trigger: ',length(unique(dataLM$id))))

#adding demographics
for (i in 1:length(LM_demographics)) {
  dataLM <- merge(dataLM, LM_demographics[[i]], by=c('id','time.stop', 'REC','LM'), all = TRUE)
}

#creating time variable for trigger
for (t in triggers) {

  time_trigger <- dataLM %>%
    select(id, time.stop, ends_with(paste0(t, '_t')), LM)

  time_trigger$time_after <- time_trigger[4] - time_trigger[3]

  #converting to weeks
  time_trigger$time_after <- round(time_trigger$time_after/0.25)
  time_trigger[3] <- time_trigger$time_after

  time_trigger <- time_trigger %>%
    select(id, time.stop, ends_with(paste0(t, '_t')), LM)

  dataLM <- dataLM %>% select(-ends_with(paste0(t, '_t'))) %>%
    merge(time_trigger, by=c('id','time.stop', 'LM'), all = TRUE)

}
print(paste0('Demographics: ',length(unique(dataLM$id))))

#creating levels for demographics
for (d in demographics) {
```

```
    time_demog <- dataLM %>%
      select(id, time.stop, ends_with(paste0(d)), LM)

    #time_demog[3][is.na(time_demog[3])] <- 'Unknown'
    #no need to replace with Unknown

    time_demog <- time_demog %>%
      select(id, time.stop, ends_with(paste0(d)), LM)

    dataLM <- dataLM %>% select(-ends_with(paste0(d))) %>%
      merge(time_demog, by=c('id','time.stop', 'LM'), all = TRUE)


  }

  #getting all individuals for each landmark
  dataLM <-  merge(followupLM, dataLM, by=c('id','time.stop', 'LM', 'REC'), all.x = TRUE)
  print(paste0('Adding general follow-up information.'))

  #adding to the landmarkk to the superset
  LM_superset <- rbindlist(list(LM_superset, dataLM))
  print(paste0('Landmark ', LM, ' is completed and added to the list.'))

}

saveRDS(LM_superset, file = paste0('Z:/!Community sentences/Predictive modelling/The boostrapped model/
LM_superset_', Outcome,'_incInjuryTBI.rds'))

LM_superset <- merge(LM_superset, fixed_covs, by='id', all.x = TRUE) # adding fixed covariates
print(paste0('IDs in superset: ',length(unique(LM_superset$id))))

# LM_superset <- readRDS(file = paste0('Z:/!Community sentences/Predictive modelling/The boostrapped mo
del/LM_superset_2108.rds'))
#saveRDS(LM_superset, file = paste0('Z:/!Community sentences/Predictive modelling/The boostrapped model
/superset_w_fixed.rds'))
```

Forward propagation of demographic variables

```
#forward propagating the first measured demographics for the first 3 months
#Last observation Carried Forward
LM_superset <- LM_superset %>%
  mutate(Dyn.employed = case_when(is.na(Dyn.employed) & LM <= 3 ~ LM_superset$Employed,
                                  TRUE ~ as.numeric(Dyn.employed)),
         Dyn.marital = case_when(is.na(Dyn.marital) & LM <= 3 ~ LM_superset$Marital,
                                 TRUE ~ as.numeric(Dyn.marital)),
         Dyn.housing = case_when(is.na(Dyn.housing) & LM <= 3 ~ LM_superset$House.changes,
                                 TRUE ~ as.numeric(Dyn.housing)),
         Dyn.income.sup = case_when(is.na(Dyn.income.sup) & LM <= 3 ~ LM_superset$Income.support,
                                    TRUE ~ as.numeric(Dyn.income.sup)),
         Dyn.days.unemp = case_when(is.na(Dyn.days.unemp) & LM <= 3 ~ LM_superset$Days.unemployed,
                                    TRUE ~ as.numeric(Dyn.days.unemp)),
         Dyn.unemp.ben = case_when(is.na(Dyn.unemp.ben) & LM <= 3 ~ LM_superset$Unemployment.benefit,
                                   TRUE ~ as.numeric(Dyn.unemp.ben)),
         #we use dyans unemployed variable for multiple imputations but not for the model
          Dyn.days.unemp = case_when(is.na(Dyn.days.unemp) & LM <= 3 ~ LM_superset$Days.unemployed,
                                     TRUE ~ as.numeric(Dyn.days.unemp)),
```

```
        #we only take the first landmark for education, as we forwrad propagate education all the way
       Dyn.education = case_when(is.na(Dyn.education) & LM <= 1 ~ LM_superset$Education.level,
                                TRUE ~ as.numeric(Dyn.education))
  )


#Education
#for education - propagating the education status from a prior landmark forward
#creating empty data frame to store the results
LM_superset_edu <- data.frame()
ids <- unique(LM_superset$id) #saving ids for iterations

#it's more convenient to iterate using index
#as you can easily return to the place where you stopped
for (i in 1:length(ids)) {

  #extracting data for one individuals
   set <- filter(LM_superset, id == ids[i])

   #checking that there's more than one landmark
   if (max(set$LM) > 0) {

        #iterating over landmarks and forward propagating the last measure
        for (lm in 1:max(set$LM)) {

          if (is.na(set$Dyn.education[set$LM == lm])) {

            set$Dyn.education[set$LM == lm] <- set$Dyn.education[set$LM == (lm-1)]

            }
          }

      #adding the individual's data to the dataframe sorted by landmark
      LM_superset_edu <- rbindlist(list(LM_superset_edu, set[order(set$LM),]))
      #data.frame::rbindlist is 10 faster than rbind

   } else {

         #if only one landmark (0), then just saving the individual's data without
         LM_superset_edu <- rbindlist(list(LM_superset_edu, set[order(set$LM),]))

   }

  #progress indicator
  print(paste0('id: ', ids[i], '. N: ', i , '. Progress: ', round(i*100/length(ids), 2), '%'))

}

#saving the results
LM_superset <-  LM_superset_edu


#transforming age to the current age
LM_superset <-LM_superset %>%
  mutate(Age = round(Age + LM/12))
```

Multiple imputations

```
#read the data file
LM_superset <- readRDS(file = …)

#Amelia imputations with EM. Markov chains
I <- 10 #number of of imputations to perform

#amelia_list <- list()
#the loop is used to save the memory
for (i in 1:10) {

  #imputation of dynamic covariates
  data_imputed_dynamic <- amelia(LM_superset, m=1, p2s = 2,
                                 idvars=c('time.stop'),
                                 noms=c('REC','Education.level', 'Marital.status',
                                        'New_psyc_noSU',
                                        'New_AU',
                                        'New_SU',
                                        'New_psyc_sev',
                                        'Psyc_hosp',
                                        'TBI',
                                        'Injury',
                                        'Intoxic',
                                        'Self_harm',
                                        'Victim',
                                        'Dyn.employed',
                                        'Dyn.marital',
                                        'Dyn.education',
                                        'Dyn.housing',
                                        'Dyn.unemp.ben',
                                        'Dyn.income.sup'),
                                 ords=c('Dyn.days.unemp',
                                        'Age', 'Sex',
                                        'Income', 'Income.support',
                                        'Employed', 'Unemployment.benefit',
                                        'Immigrant.status',
                                        'Prior.offence', 'Prior.violent.offence',
                                        'Index.violent.offence',
                                        'Prior.self.harm',
                                        'Days.unemployed',
                                        'House.changes',
                                        'Birth.year', 'Conviction.year',
                                        'Any.psych', 'Any.psych.exlSU',
                                        'Sch.spectrum', 'Bipolar', 'Anxiety', 'Depression',
                                        'Alcohol.use', 'Drug.use', 'Substance.use',
                                        'Person.dis', 'Other.develop'),
                                 ts = 'LM', cs = 'id', polytime = 1) #using linear time


   #amelia_list[[length(amelia_list)+1]] <- data_imputed_dynamic

   #saving the results
   saveRDS(data_imputed_dynamic, file = paste0(…))
```

```
  print(paste0('Imputation ',i,' is done'))

  gc() #garbage collection
}
```

Model specification and custom functions

Specifying the models to check

```
#NLM and NBL were added for additional comparison,
#but they performance was not systematically tested for the study

#Models:
# - Dynamic landmark model (main model);
# - Fixed landmark model;
# - Fixed baseline model;
# - Null landmark model (age and sex);
# - Null baseline model (age and sex).


##----- Dynamic landmark model (main model)
#For VIOLENT reoffending
v.dyn.lm.formula <-  Surv(time.stop.rel, REC) ~
  #Group 1
  Sex +
  #Criminal history
  Prior.offence +  Prior.violent.offence + Prior.prison +
  Index.violent.offence +
  #Group 2
  #Current demographics
  Age +
  Dyn.marital + Dyn.edu.group +
  Dyn.employed +
  Dyn.unst.housing + Dyn.income.sup +
  #Current diagnoses
  Cur.psych.exlSU + Cur.psych.sev +
  Cur.alcohol.use + Cur.drug.use +
  #Group 3
  #Baseline demographics
  Employed +
  #Group 4
  #Triggers
  Psyc_hosp +
  InjuryTBI + InjuryTBI_w1 +
  Victim + Victim_m1 +
  Intoxic_m1 +
  #technical variables
  #stratification and clustering
  strata(LM) + cluster(id)


#For GENERAL reoffending
g.dyn.lm.formula <-  Surv(time.stop.rel, REC) ~
  #Group 1
  Sex +
  #Criminal history
  Prior.offence +  Prior.violent.offence + Prior.prison +
```

```
  Index.violent.offence +
  #Group 2
  #Current demographics
  Age +
  Dyn.marital + Dyn.edu.group +
  Dyn.employed +
  Dyn.unst.housing + Dyn.income.sup +
  #History of mental health problems (current)
  Cur.psych.exlSU + Cur.psych.sev +
  Cur.alcohol.use + Cur.drug.use +
  Cur.self.harm +
  #Group 3
  #Baseline demographics
  Employed +
  Income.support +
  #Group 4
  #Triggers
  Psyc_hosp + Psyc_hosp_w1 +
  InjuryTBI + InjuryTBI_m1 + InjuryTBI_w1 +
  Victim + Victim_m1 +
  Intoxic + Intoxic_m1 +
  #technical variables
  #stratification and clustering
  strata(LM) + cluster(id)

##----- Fixed landmark model
fixed.lm.formula <- Surv(time.stop.rel, REC) ~ Sex +
  #Criminal history
  Prior.offence +  Prior.violent.offence + Prior.prison +
  Index.violent.offence +
  Age_baseline + #age at baseline
  Marital.status +  Education.group +
  Employed +
  Income.support +
  Unstable.housing +
  #Current diagnoses
  Any.psych.exlSU + Psyc.severe +
  Alcohol.use + Drug.use +
  strata(LM) + cluster(id)


##----- Null landmark model
null.lm.formula <- Surv(time.stop.rel, REC) ~ Age + Sex + strata(LM) + cluster(id)

##----- Fixed baseline model
#Has to be used with data = filter(data, LM == 0)
fixed.bl.formula <- Surv(time.stop.rel, REC) ~ Sex +
  #Criminal history
  Prior.offence +  Prior.violent.offence + Prior.prison +
  Index.violent.offence +
  Age_baseline + #age at baseline
  Marital.status +  Education.group +
  Employed +
  Income.support +
  Unstable.housing +
  #Current diagnoses
```

```
  Any.psych.exlSU + Psyc.severe +
  Alcohol.use + Drug.use


##----- Null baseline model
#Has to be used with data = filter(data, LM == 0)
null.bl.formula <- Surv(time.stop.rel, REC) ~ Age_baseline + Sex
```

Extracting baseline cumulative hazards

```
# ------ lm_hazards

#function that extracts the cumulative hazard for a given time window from a landmark model

#input:
#1. fit - a coxph() fit stratified by landmarks (Breslow estimator).
#2. w - horizon of prediction (default is 24 months).
#The function works with any w <= max(t).
#However, it's best if the data is administratively censored at t = w.

#Output
#1. lm_bhs - baseline hazards estimated for each landmark

#Extracting baseline hazard

lm_hazards <- function(fit, w = 24) {

  bh <- basehaz(fit, centered=FALSE)

  #saving the hazard for the landmark model
  lm_bhs <- data.frame()
  for(n in unique(bh$strata)) {

    LM <- as.numeric(substr(n, 4, 10))

    bs <- filter(bh, bh$strata == n)
    #finding the closest time point to the horizon
    #if you are using Surv(time.start, time.stop, Outcome), i.e. left-censored observations,
    #then time_pred = LM + w, where LM is the landmark time, and w is the sliding window (prediction ho
rizon).
    #if you are using Surv(time.stop, Outcome), i.e. time relative to landmark (no left-censoring),
    #then time_pred = w
    #the results are the same
    time_pred <- w
    bs$dif <- abs((time_pred) - bs$time)
    t <- filter(bs, bs$dif == min(bs$dif))

    #print(paste0(n, ' ', round(t$hazard, 4)))

    vector <- data.frame(LM = LM,
                         Bhazard = round(t$hazard, 4))

    lm_bhs <- rbind(lm_bhs, vector)
```

```
  }

  return(lm_bhs)

}


# ------ fixed_hazard

#function that extracts the cumulative hazard for a given time window from a regular dataset
#and outputs the same fixed baseline hazard for all provided landmarks

#input:
#1. fit - a coxph() fit without stratification (Breslow estimator).
#2. landmarks - the vector of landmark cut points for formatting the output similarly to lm_bhs. Defaul
t is 0-36.
#3. w - horizon of prediction (default is 24 months).
#The function works with any w <= max(t).
#However, it's best if the data is administratively censored at t = w.

#Output
#1. fixed_bh - cumulative baseline hazard estimated for a given t


fixed_hazard <- function(fit, landmarks = seq(0,36, by = 1), w = 24) {

  bh <- basehaz(fit, centered=FALSE)

  #extracting cumulative hazard at a given horizon
  #or as close to it as possible
  time_pred <- w
  bh$dif <- abs((time_pred) - bh$time)
  t <- filter(bh, bh$dif == min(bh$dif))
  bh_w <- filter(t, time == w)

  #formatting the output for the landmarks
  fixed_bh <- data.frame()
  for (l in landmarks) {
    vector <- data.frame(LM = l,
                         Bhazard = round(bh_w[[1]], 4))

    fixed_bh <- rbind(fixed_bh, vector)
  }

  return(fixed_bh)
}
```

Specifying the function for obtaining prognostic indicies

```
# ------ forsight

#function that obtains prognostic indices for a set of coefficients and baseline hazards
```

```
#input:
#1. coefs - a dataframe with cavariate names (first columns) and corresponding coefficients (second col
umn)
#2. lm_h - the output of lm_hazards custom function, estimated for a horizon w.
#3. w - horizon of prediction (default is 24 months). Used for the conviniece. To mark outcomes.
#4. newdata - data for predictors. Has to have LM columns for landmarks
#5. factors (optional) - a vector of variables that have multiple categories. They will be converted to
 dummy variables (fastDummies)
#6. model.name - used for convinience. Wnen comparing multiple models can be used to distinguish betwee
n them .
#We need to force administrative consoring!
#which contains baseline hazards for different landmarks

#Output
#1. newdata - modified dataset with added predictions
#2. predictions - a selected columns with id and predictions from newdata


forsight <- function(coefs, lm_h, w = 24, newdata, factors = NULL, model.name = 'm') {

  require(fastDummies)
  require(stringr)

  #checking if it's null model or not
  if (!is.null(coefs)) {

      #deleting equal signs from grouped variables
      coefs <- as.data.frame(coefs)
      for (r in 1:nrow(coefs)) {

        #Sometimes this piece of code is required, sometime it's not.
        #A bit unpredictable dummification.

        #changing grouP to grouP_ . For later use as dummy variables
        rownames(coefs)[r] <- str_replace(rownames(coefs)[r], 'group', 'group_')

      }

  } else {

    print('No coefficients specified. Calculating null model')

  }

  # lm_h = bhs_24
  #extracting baseline hazards
  # lm_h <-  lm_h %>%
  #   select(LM, mean)
  colnames(lm_h) <- c('LM', paste0('bh_', w))


  if (is.null(factors) | is.null(coefs)) {
    #adding baseline hazards
    #adding stop time relative to landmark for later use
    newdata <- newdata %>% mutate(time.stop.rel = time.stop - LM) %>%
      merge(lm_h, by = 'LM')
```

```
  print('No factor variables specified')

} else {

  newdata <- newdata %>% mutate(time.stop.rel = time.stop - LM) %>%
    dummy_cols(select_columns = factors,
               remove_first_dummy = TRUE,
               remove_selected_columns = TRUE) %>%
    merge(lm_h, by = 'LM')

 print(paste0('Creating dummy variables for ', factors))

}

#checking if it's a null model or not
if (!is.null(coefs)) {

    #storing all coefficient and then combining them
    lp <- select(newdata, id)
    lp$r <- as.numeric(rownames(lp))
    for (k in 1:nrow(coefs)) {

      cov <- paste0(rownames(coefs)[k])

      #applying coefficients (weighting) to the covariate
      cov_w <- eval(parse(text = paste0('newdata$"', cov,'"','* coefs[k,]')))


      # #debugging line
      # print(paste0('newdata$"', cov,'"','* as.numeric(filter(coefs[2], coefs[1] == "',cov, '"))'))

      cov_w <- as.data.frame(cov_w)
      cov_w$r <- as.numeric(rownames(cov_w))
      colnames(cov_w) <- c(cov, 'r')

      lp <- merge(lp, cov_w, by = 'r', all.x = TRUE)

    }

    #summing across all columns to get prognostic index (linear predictor)
    newdata$lp <- rowSums(lp[ , 3:length(lp)])


    #creating a risk column
    #formulas (they are equivalent)
    # Survival(t) =  exp(-baseline_cumulative_hazard(t)*exp(linear_predictor)).
    # Survival(t) =  exp(-baseline_cumulative_hazard(t))^exp(linear_predictor).
    # Risk(t) =  1 - Survival(t)
    # From Putter's "The landmark approach an introduction and application to risk prediction..."

   #formula
    eval(parse(text = paste0('newdata$risk_', w, '<- 1 - exp(-newdata$bh_', w, ')^exp(newdata$lp)'

    #Adding the model's identifiers
    colnames(newdata)[which(names(newdata) == 'lp')] <- paste0(model.name, '.lp')
    colnames(newdata)[which(names(newdata) == paste0('bh_', w))] <- paste0(model.name, '.bh_', w)
```

```
      colnames(newdata)[which(names(newdata) == paste0('risk_', w))] <-  paste0(model.name, '.risk_', w
)

      print(paste0('Prediction is complete for model ', model.name))

      return(select(newdata, 'id', 'LM',
                    'time.stop', 'time.stop.rel',
                    'REC',
                    paste0(model.name, '.lp'),
                    paste0(model.name, '.bh_', w),
                    paste0(model.name, '.risk_', w)))

    } else {

      #if it's null model the making null predictions

      #then prognostic indices will be zeros
      newdata$lp <- 0

      #creating a risk column
      #here is the correct one
      # Survival(t) =  exp(-baseline_cumulative_hazard(t)*exp(linear_predictor)).
      # Survival(t) =  exp(-baseline_cumulative_hazard(t))^exp(linear_predictor).
      # Risk(t) =  1 - Survival(t)
      # From Putter "THe landmark approach an introduction and application to risk prediction..."

      #eval(parse(text = paste0('newdata$risk_', w, '<-  1 - exp(-newdata$bh_', w, '*exp(newdata$lp))')
)) #right
      eval(parse(text = paste0('newdata$risk_', w, '<- 1 - exp(-newdata$bh_', w, ')^exp(newdata$lp)')))
 #right

      #Adding the model's identifiers
      colnames(newdata)[which(names(newdata) == 'lp')] <- paste0(model.name, '.lp')
      colnames(newdata)[which(names(newdata) == paste0('bh_', w))] <- paste0(model.name, '.bh_', w)
      colnames(newdata)[which(names(newdata) == paste0('risk_', w))] <-  paste0(model.name, '.risk_', w
)

      print(paste0('Prediction is complete for null model ', model.name))

      return(select(newdata, 'id', 'LM',
                    'time.stop', 'time.stop.rel',
                    'REC',
                    paste0(model.name, '.lp'),
                    paste0(model.name, '.bh_', w),
                    paste0(model.name, '.risk_', w)))

    }
  }
```

Variable selection

Landmark settings

```
Outcome <- 'General' #or ‘Violent’
I <- 10 #imputations
threshold_p <- 0.157
```

Specification of the model

```
#variable selection performed on non-dummified model
formula <- Surv(time.start, time.stop, REC) ~
  #Group 1
    #Baseline demographics
    Sex +
    #Criminal history
    Prior.offence +  Prior.violent.offence + Prior.prison +
    Index.violent.offence +
  #Group 2
    #Current demographics
    Age +
    Dyn.marital +  Dyn.education +
    Dyn.employed +
    Dyn.unst.housing + Dyn.income.sup +
    #Current diagnoses
    Cur.psych.exlSU + Cur.psych.sev +
    Cur.alcohol.use + Cur.drug.use +
    Cur.self.harm +
  #Group 3
    #Baseline demographics
    #Marital.status +
    Employed +
    Unstable.housing + Income.support +
    #Baseline psychiatric variables
    #Prior.self.harm +
  #Group 4
    #Triggers
    Psyc_hosp + Psyc_hosp_w1 + Psyc_hosp_m1 +
    #TBI + TBI_w1 + TBI_m1 +
    #Injury + Injury_w1 + Injury_m1 +
    InjuryTBI + InjuryTBI_w1 + InjuryTBI_m1 +
    #Self_harm + Self_harm_w1 + Self_harm_m1 +
    Victim + Victim_w1 + Victim_m1 +
    Intoxic +  Intoxic_w1 + Intoxic_m1 +
  #Technical variables
    #Stratification and clustering
    strata(LM) + cluster(id)
```

Variable selection candidates

```
vars_test <- list(#Group 1 and 2 (included by default, so we don't test them)
  #Group 3
 c('Unstable.housing'),
 c('Marital.status'),
 # c('Prior.self.harm'),
   #added later by combining prior self-harm and trigger
   c('Cur.self.harm'),
 c('Employed'),
 c('Income.support'),
  #Group 4
  c('Psyc_hosp'), c('Psyc_hosp_w1'), c('Psyc_hosp_m1'),
  c('Victim'), c('Victim_w1'), c('Victim_m1'),
```

```
  c('Intoxic'), c('Intoxic_w1'), c('Intoxic_m1'),
  # c('Injury'), c('Injury_w1'), c('Injury_m1'),
  # c('TBI'), c('TBI_w1'), c('TBI_m1'),
     #added later by combining Injury and TBI
      c('InjuryTBI'), c('InjuryTBI_w1'), c('InjuryTBI_m1'))
  # c('Self_harm'), c('Self_harm_w1'), c('Self_harm_m1'))
```

Initiating variable selection

```
results <- data.frame() #storage for the results
cox_formula <- formula #initial formula
for (s in 1:length(vars_test)) {


      print('The starting model: ')
      print(paste0(cox_formula[3]))

      #converting to text and cleaning
      f <- as.character(cox_formula[3])
      f <- gsub(' ', '', paste0(cox_formula[3]), fixed = TRUE)
      f <- gsub("[ \t\n\r\v\f]", "", f, fixed = TRUE)
      f <- gsub("\n", "", f, fixed = TRUE)

      #parsing the formula to get the variables
      vars <- strsplit(f, '+' , fixed = TRUE)[[1]]

  #I. Fitting the models on imputations
      models  <- list()
      for (i in 1:I){

        imputation <- data.table(readRDS(file = …))

        print(paste0('Imputation ', i,' loaded'))

                      # #sampling for the debugging
                      # #Loop for creating balaced samples across landmarks
                      # n_lms <- unique(imputation$LM)
                      # data_boot <- data.frame()
                      # for (l in n_lms) {
                      #
                      #   #sample with replacement from a given landmark
                      #   lm_boot <- filter(imputation, LM == l)[sample(nrow(filter(imputation, LM == l
)),
                      #   size = nrow(filter(imputation, LM == l))*0.2, replace = TRUE), ]
                      #   data_boot <- rbind(data_boot, lm_boot) #getting a bootstrapped dataset
                      #
                      # }
                      # print('Resampling complete')
                      #
                      # imputation <- data_boot

                #fitting the model on the imputed dataset
                    #Timing of the model fit. Starting the clock!
                    ptm <- proc.time()
```

```
                    #fitting the model
                    eval(parse(text = paste0('cox <- with(imputation, coxph(', cox_formula[2], ' ~ ', c
ox_formula[3],
                                           ', method = "breslow", x = TRUE, y = TRUE))')))
                    #Stopping the clock
                    t_lapsed <- proc.time() - ptm


                    #cleaning the memory
                    remove(imputation)

                print(paste0('Model fitted. The fit took ', round(t_lapsed[[3]]), ' s.'))
                #print(summary(cox))

                #saving the models
                 models[[length(models) + 1]] <- cox
      }

      print(paste0('All ', I, ' imputations were processed for the step ', s,'. Pooling the estimates')
)

  #II. Pooling the estimates. Applying Rubin's rule
        #some nice summary here https://bookdown.org/mwheymans/bookmi/rubins-rules.html
        #Heymans, M.W., & Eekhout, I. (2019). Applied Missing Data Analysis (first draft). Amsterdam

        #1. Point estimates for coefficients
          #pooling coefficients
          betas <- as.data.frame(coef(models[[1]]))
          colnames(betas) <- 1
          for (i in 2:I) {

            betas[paste0(i)] <- as.data.frame(coef(models[[i]]))

          }
          #estimating the average
          betas$mean <- rowMeans(betas[1:I])

        #2. Within imputation variance
          #pooling robust Standard Errors and squaring them
          rSE2 <- as.data.frame((summary(models[[1]])$coefficients[,'robust se'])^2)
          colnames(rSE2) <- 1
          for (i in 2:I) {

            #saving squred robust estimators
            rSE2[paste0(i)]  <- data.frame((summary(models[[i]])$coefficients[,'robust se'])^2)

          }

            #estimating the mean
          rSE2$wi.var <- rowMeans(rSE2[1:I]) #it is an estimate of within imputation variance
          #adding to the coefficient estimates
          betas$wi.var <- rSE2$wi.var


        #3. Between imputation variance
        bi.vars <- data.frame()
        for (r in 1:nrow(betas)) {
```

```
          #extracting the estimates for one covariate and transposing it for convenience
          covariate <- as.data.frame(t(betas[r,1:I]))

          #subtracting the pooled value of the estimate from the value of each imputation
          covariate$error <- covariate[,1] - betas$mean[r]

          #estimating between imputation variance for a given covariate
          bi.var <- sum(covariate$error^2)/(I-1)

          #saving the result
          bi.vars <- rbind(bi.vars, bi.var)
        }

        #adding to the coefficient estimates
        betas$bi.var <- bi.vars[,1]

        #4. Total variance
        #any TW: Shogun fans?
        betas <- betas %>% mutate(total.var = wi.var + bi.var + bi.var/I)

        #5. Pooled standard errors
        betas <- betas %>% mutate(pooled.se = sqrt(total.var))


      #Calculating z-statistics and estimating p-value
        betas <- betas %>% mutate(z = mean/pooled.se) %>%
          #returning p-values from two tailed hypothesis test
          mutate('Pr(>|z|)' = 2*pnorm(abs(z), lower.tail = FALSE))


      #Creating a column with covariate names and cleaning up
        betas <- round(betas, 9) #Rounding, so the table looks cleaner
        betas <- betas %>%
          mutate(var = rownames(betas)) %>%
          select(var, mean, pooled.se, z, 'Pr(>|z|)')

        rownames(betas) <- 1:nrow(betas) #re-indexing rows

    print(paste0('Pooling complete. ', nrow(betas), ' estimates were pooled. Starting backward eliminat
ion'))

 #III. Backward elimination
      #selecting only covariates that we need to check
        betas <- filter(betas, var %in% unique(unlist(vars_test)))

        #finding the highest p-value
        var_max_p <- filter(betas, betas$'Pr(>|z|)' == max(betas$'Pr(>|z|)') )


        print(paste0(var_max_p$var,' has the p-value of ', round(var_max_p$'Pr(>|z|)', 3)))

        if (var_max_p$'Pr(>|z|)' > threshold_p) {

          print(paste0('The p-value is higher than a threshold of ', threshold_p, '. Eliminating the va
riable from the model'))
```

```
          #deleting the variable from the full list of variables
          vars <- vars[vars != paste0(var_max_p$var)]

          #updating the cox_formula
          cox_formula[3] <- paste(vars, collapse = ' + ')

          outcome <- data.frame('Step' = s,
                                'Var' = var_max_p$var,
                                'Var_p' = round(var_max_p$'Pr(>|z|)', 3),
                                'Threshold_p' = threshold_p,
                                'Result' = 'Eliminated')

          results <- rbind(results, outcome)

          print(results)

        } else {

          print(paste0('The p-value is lower than a threshold of ',
                                threshold_p, '. The model is finalised'))
          print(results)
          final_vars <- vars
          print(paste0('The final model is  ', print(paste0(cox_formula[3]))))

          #stopping the algorithm
          break

        }
}
```

Internal validation. Harrel’s bias correction method

As described in Iba, K., Shinozaki, T., Maruo, K., & Noma, H. (2021). Re-evaluation of the comparative effectiveness of bootstrap-based optimism correction methods in the development of multivariable clinical prediction models. BMC Medical Research Methodology, 21(1), 1-14.

```
library(rms)
library(pec)
library(survcomp)

Outcome <- 'Violent'
N_boot <- 3
rt1 <- 0.1 #risk threshold for sensitivity and specificity statistics
rt2 <- 0.2
rt3 <- 0.5
#default is 0.2 (20%) for violent reoffending and 0.5 (50%) for general
```

Creating lists of formulas

```
#Landmark models
lm_model_names <- c('DLM','FLM','NLM')
lm_models <- list(v.dyn.lm.formula, fixed.lm.formula, null.lm.formula)
```

```
#Baseline models
bl_model_names <- c('FBL','NBL')
bl_models <-  list(fixed.bl.formula, null.bl.formula)
```

Starting the loop over imputations

```
for (i in 1:I) {

#loading data
#######
data <- data.table(readRDS(file = …))  %>% mutate(time.stop.rel = time.stop - LM)
#adding age at baseline for baseline models
data$Age_baseline <- round(data$Age - data$LM/12)
#######

#storing the results per landmark
results_app <- data.frame()
results_boot <- data.frame()
results_orig <- data.frame()
#storing average results
lm_results_app <- list()
lm_results_boot <- list()
lm_results_orig <- list()
for (m in 1:1) {

  model_name <- lm_model_names[[m]]
  formula <- lm_models[[m]]

  #printing the model
  print(paste0('Model: ', model_name))

#APPARENT PERFORMANCE (THETA APP)
#estimating the apparent performance on the original population

  #fitting the model into the full dataset
  cox_app <- coxph(formula, method = 'breslow', data = data)
  hazards <- lm_hazards(cox_app, w = 24) #extracting cumulative hazard for each landmark

  predictions <- forsight(coefs = coef(cox_app),
                          lm_h = hazards,
                          w = 24,
                          newdata = data,
                          factors = c('Dyn.edu.group','Education.group'),
                          model.name = model_name)


        #calculating c-index for each landmark

        for (l in Landmarks){

          #loading a landmark set
          lm_set <- filter(predictions, LM == l)
                   #calculating it's proportion to all entries
          weight <- nrow(lm_set)/nrow(predictions)
```

```
        #calculating c-index for a given ladndmark
        cindex <- concordance.index(x = eval(parse(text=(paste0('lm_set$', model_name, '.risk_', w)))
),
                                    surv.time = lm_set$time.stop.rel,
                                    surv.event = lm_set$REC, method = c("noether"))


        #loading a landmark set without censored individuals
        #for calculating Brier score, TPR, FPR, NPV, and PPV.
        lm_set_full_cases <- filter(lm_set, !(REC == 0 & time.stop.rel < 24))
        lm_set_full_cases$brier <- round(eval(parse(text=(paste0('lm_set_full_cases$',
                                                   model_name,
                                                   '.risk_', w,' - ',
                                                   'lm_set_full_cases$REC'))))^2, 5)
        weight_censored <- nrow(lm_set_full_cases)/nrow(predictions)

        brier <- sum(eval(parse(text=(paste0('lm_set_full_cases$',
                                      model_name,
                                      '.risk_', w,' - ',
                                      'lm_set_full_cases$REC'))))^2)/nrow(lm_set_full_cases)

        #printing the progress
        print(paste0('The apparent metrics for landmark ', l,' are calculated'))

        #saving the results
        vector <- data.frame(model = model_name,
                             imputation = i,
                             performance = 'Apparent',
                             LM = l,
                             weight = weight,
                             weight_censored = weight_censored,
                             cindex = cindex$c.index,
                             brier = brier)


        metrics_apparent <- rbind(metrics_apparent, vector)


      }

  lm_results_app[[length(lm_results_app) + 1]] <- metrics_apparent

  #calculating weighted average of c-index across landmarks
  cindex_average <- sum(metrics_apparent$weight*metrics_apparent$cindex)
  brier_average <- sum(metrics_apparent$weight_censored*metrics_apparent$brier)

  average_apparent <- data.frame(Model = model_name,
                                 cindex = cindex_average,
                                 brier = brier_average)

  results_app <- rbind(results_app, average_apparent)


  #######################
  #ESTIMATING BOOTSTRAPPED PERFORMANCE (THETA BOOT)
  #estimating the apparent performance on the original population
```

```
  for (b in 1:N_boot) {

    #bootstrap resampling
    n_lms <- unique(data$LM)
    data_boot <- data.frame()
    for (l in n_lms) {

      #sample with replacement from a given landmark
      lm_boot <- filter(data, LM == l)[sample(nrow(filter(data, LM == l)), size = nrow(filter(data, LM
== l))*0.25, replace = TRUE), ]
      data_boot <- rbind(data_boot, lm_boot) #getting a bootstrapped dataset

    }

    cox_boot <- coxph(formula, method = 'breslow', data = data_boot)
    #summary(cox_boot)
    hazards_boot <- lm_hazards(cox_boot, w = 24) #extracting cumulative hazard for each landmark


    #BOOTSRAP PERFORMANCE ON A BOOTSTRAP SAMPLE
    predictions <- forsight(coefs = coef(cox_boot),
                            lm_h = hazards_boot,
                            w = 24,
                            newdata = data_boot, #using the bootstrapped sample
                            factors = c('Dyn.edu.group','Education.group'),
                            model.name = model_name)


    #calculating c-index for each landmark
    metrics_boot <- data.frame()
    for (l in Landmarks){

      #loading a landmark set
      lm_set <- filter(predictions, LM == l)
      #calculating it's proportion to all entries
      weight <- nrow(lm_set)/nrow(predictions)


      #calculating c-index for a given ladndmark
      cindex <- concordance.index(x = eval(parse(text=(paste0('lm_set$', model_name, '.risk_', w)))),
                                  surv.time = lm_set$time.stop.rel,
                                  surv.event = lm_set$REC, method = c("noether"))


      #loading a landmark set without censored individuals
      #for calculating Brier score, TPR, FPR, NPV, and PPV.
      lm_set_full_cases <- filter(lm_set, !(REC == 0 & time.stop.rel < 24))
      lm_set_full_cases$brier <- round(eval(parse(text=(paste0('lm_set_full_cases$',
                                                               model_name,
                                                               '.risk_', w,' - ',
                                                               'lm_set_full_cases$REC'))))^2, 5)
      weight_censored <- nrow(lm_set_full_cases)/nrow(predictions)


      brier <- sum(eval(parse(text=(paste0('lm_set_full_cases$',
```

```
                                           model_name,
                                           '.risk_', w,' - ',
                                           'lm_set_full_cases$REC'))))^2)/nrow(lm_set_full_cases)


    #printing the progress
    print(paste0('The bootstrapped metrics for landmark ', l,' are calculated'))

    #saving the results
    vector <- data.frame(model = model_name,
                         imputation = i,
                         performance = 'Boot',
                         iteration = b,
                         LM = l,
                         weight = weight,
                         weight_censored = weight_censored,
                         cindex = cindex$c.index,
                         brier = brier)


    metrics_boot <- rbind(metrics_boot, vector)


  }


  lm_results_boot[[length(lm_results_boot) + 1]] <- metrics_boot

  #calculating weighted average of c-index across landmarks
  cindex_average <- sum(metrics_boot$weight*metrics_boot$cindex)
  brier_average <- sum(metrics_boot$weight_censored*metrics_boot$brier)

  average_boot <- data.frame(Model = model_name,
                             iteration = b,
                                 cindex = cindex_average,
                                 brier = brier_average)

  results_boot <- rbind(results_boot, average_boot)



  #####################
  #BOOTSRAP PERFORMANCE ON ORIGINAL DATA
  predictions <- forsight(coefs = coef(cox_boot),
                          lm_h = hazards_boot,
                          w = 24,
                          newdata = data, #using the original data
                          factors = c('Dyn.edu.group','Education.group'),
                          model.name = model_name)


  #calculating c-index for each landmark
  metrics_orig <- data.frame()
  for (l in Landmarks){

    #loading a landmark set
```

```
  lm_set <- filter(predictions, LM == l)
  #calculating it's proportion to all entries
  weight <- nrow(lm_set)/nrow(predictions)


  #calculating c-index for a given ladndmark
  cindex <- concordance.index(x = eval(parse(text=(paste0('lm_set$', model_name, '.risk_', w)))),
                              surv.time = lm_set$time.stop.rel,
                              surv.event = lm_set$REC, method = c("noether"))


  #loading a landmark set without censored individuals
  #for calculating Brier score, TPR, FPR, NPV, and PPV.
  lm_set_full_cases <- filter(lm_set, !(REC == 0 & time.stop.rel < 24))
  lm_set_full_cases$brier <- round(eval(parse(text=(paste0('lm_set_full_cases$',
                                                             model_name,
                                                             '.risk_', w,' - ',
                                                             'lm_set_full_cases$REC'))))^2, 5)
  weight_censored <- nrow(lm_set_full_cases)/nrow(predictions)


  brier <- sum(eval(parse(text=(paste0('lm_set_full_cases$',
                                       model_name,
                                       '.risk_', w,' - ',
                                       'lm_set_full_cases$REC'))))^2)/nrow(lm_set_full_cases)


  #printing the progress
  print(paste0('The bootstrapped metrics for landmark ', l,' are calculated'))

  #saving the results
  vector <- data.frame(model = model_name,
                       imputation = i,
                       performance = 'Orig',
                       iteration = b,
                       LM = l,
                       weight = weight,
                       weight_censored = weight_censored,
                       cindex = cindex$c.index,
                       brier = brier)


  metrics_orig <- rbind(metrics_orig, vector)


}

lm_results_orig[[length(lm_results_orig) + 1]] <- metrics_orig

#calculating weighted average of c-index across landmarks
cindex_average <- sum(metrics_orig$weight*metrics_orig$cindex)
brier_average <- sum(metrics_orig$weight_censored*metrics_orig$brier)

average_orig <- data.frame(Model = model_name,
                           iteration = b,
                           cindex = cindex_average,
```

```
                               brier = brier_average)

    results_orig <- rbind(results_orig, average_orig)

  }

}
```

Training final models

Obtaining predictors for all imputations

```
#Main outcome + violent reoffending
Outcome <- 'General' #General or Violent repffendng
w <- 24 #prediction horizon in months (default is 24 months)
I <- 10

#cretaing lists of formulas
#Landmark models
lm_model_names <- c('DLM','FLM','NLM')
lm_models <- list(g.dyn.lm.formula, fixed.lm.formula, null.lm.formula)
#Baseline models
bl_model_names <- c('FBL','NBL')
bl_models <-  list(fixed.bl.formula, null.bl.formula)

#all model names
all_model_names <- c('DLM','FLM','NLM','FBL','NBL')
all_models <- c(g.dyn.lm.formula, fixed.lm.formula, null.lm.formula,
                     fixed.bl.formula, null.bl.formula)

#training for landmark models
for (m in 4:length(lm_model_names)){

        #choosing a model
        model_name <- lm_model_names[[m]]
        formula <- lm_models[[m]]

        models <- list()
        for (i in 1:I){

          data <- data.table(readRDS(file = …)) %>% mutate(time.stop.rel = time.stop - LM)
          #adding age at baseline for baseline models
          data <- data %>% mutate(Age_baseline = round(Age - LM/12),
                                  #adding a marker for censoring
                                  Censored = case_when(REC == 0 & time.stop.rel < 24 ~ 1,
                                                       TRUE ~ 0))
          data <- data[sample(nrow(data), 100000), ]

          eval(parse(text = paste0('cox <- with(data, coxph(', formula[2], ' ~ ', formula[3],
                                   ', method = "breslow"))')))
          print(paste0('The model ', model_name,' fitted. Imputation: ', i))
          #cleaning the memory
          remove(data)
```

```
          #saving the models
          models[[length(models) + 1]] <- cox

        }

        saveRDS(models,
                file = …)
}
```

Training for baseline models

```
for (m in 1:length(bl_model_names)){

  #choosing a model
  model_name <- bl_model_names[[m]]
  formula <- bl_models[[m]]

  models <- list()
  for (i in 1:I){

    data <- data.table(readRDS(file = …) %>% mutate(time.stop.rel = time.stop - LM)
    #adding age at baseline for baseline models
    data <- data %>% mutate(Age_baseline = round(Age - LM/12),
                            #adding a marker for censoring
                            Censored = case_when(REC == 0 & time.stop.rel < 24 ~ 1,
                                                 TRUE ~ 0)) %>% filter(LM == 0)


    eval(parse(text = paste0('cox <- with(data, coxph(', formula[2], ' ~ ', formula[3],
                             ', method = "breslow", x = TRUE, y = TRUE))')))

    print(paste0('The model ', model_name,' fitted. Imputation: ', i))
    #cleaning the memory
    remove(data)

    #saving the models
    models[[length(models) + 1]] <- cox

  }

  saveRDS(models,file =…)

}

#Averaging coefficients
#LANDMARK MODELS

for (m in 1:length(lm_model_names)) {

  model_name <- lm_model_names[[m]]
  models_list <- readRDS(file =…)

  coefficients <- data.frame()
  baseline_hazards <- data.frame()
  for (n in 1:length(models_list)) {
```

```
    model <- models_list[[n]]
    summary(model)

    #extracting coefficients
    coefs <- as.data.frame(t(coef(model)))
    coefs$imputation <- n
    coefs$model <- model_name

    coefficients <- rbind(coefficients, coefs)

    #extracting cumulative hazard for each landmark
    hazards <- lm_hazards(model, w = 24)
    hazards <- as.data.frame(t(hazards$Bhazard))
    colnames(hazards) <- seq(0, 36, by = 1)
    hazards$imputation <- n
    hazards$model <- model_name


    baseline_hazards <- rbind(baseline_hazards, hazards)

  }

  remove(models_list)

  #saving the coefficients and hazards for the model
  trained_model <- list(colMeans(coefficients[1:(length(coefficients)-2)]),
                        colMeans(baseline_hazards[1:(length(baseline_hazards)-2)]))
  saveRDS(trained_model,
          file =…)
}

#BASELINE MODELS

for (m in 1:length(bl_model_names)) {

  model_name <- bl_model_names[[m]]
  models_list <- readRDS(file = …)

  coefficients <- data.frame()
  baseline_hazards <- data.frame()
  for (n in 1:length(models_list)) {

    model <- models_list[[n]]
    summary(model)

    #extracting coefficients
    coefs <- as.data.frame(t(coef(model)))
    coefs$imputation <- n
    coefs$model <- model_name

    coefficients <- rbind(coefficients, coefs)

    #extracting cumulative hazard for each landmark
    hazards <- fixed_hazard(model, w = 24)
    hazards <- as.data.frame(t(hazards$Bhazard))
    colnames(hazards) <- seq(0, 36, by = 1)
```

```
    hazards$imputation <- n
    hazards$model <- model_name


    baseline_hazards <- rbind(baseline_hazards, hazards)

  }

  #saving the coefficients and hazards for the model
  trained_model <- list(colMeans(coefficients[1:(length(coefficients)-2)]),
                        colMeans(baseline_hazards[1:(length(baseline_hazards)-2)]))
  saveRDS(trained_model, file =…)


}

#Loading the coefficients
#dynamic landmark model (violent reoffending)
coefs_vdlm <- readRDS(file =…[[1]]
bhs_vdlm <- readRDS(file = …)[[2]]

cv <- as.data.frame(coefs_vdlm)
cv$Vars <- rownames(cv)
rownames(cv) <- 1:nrow(cv)
write.xlsx(cv, file = …)

bv <- as.data.frame(bhs_vdlm)
write.xlsx(bv, file = …)

#dynamic landmark model (general reoffending)
coefs_gdlm <- readRDS(file = …)[[1]]
bhs_gdlm <- readRDS(file = …)[[2]]

cg <- as.data.frame(coefs_gdlm)
cg$Vars <- rownames(cg)
rownames(cg) <- 1:nrow(cg)
write.xlsx(c,
           file =
bg <- as.data.frame

bg <- as.data.frame(bhs_gdlm)
write.xlsx(bg, file = …)
```