POL-2313

Eve Fellows

Wolfson College

Supervisor: Dr Sara Valdebenito

UNIVERSITY OF CAMBRIDGE

Utilising Machine Learning to forecast Sexual Recidivism on the Rail Network

Word Count: 18000

Submitted in part fulfilment of the requirements for the

Master's Degree in Applied Criminology and Police Management

January 2025

Abstract

Study Aim

This study aimed to use existing police data and machine learning algorithms, to enhance the accuracy of sexual recidivism predictions on the UK rail network. Using these forecasts, this research hoped to identify repeat sex offenders and target interventions more effectively.

Research Questions

RQ1: What is the nature and prevalence of sexual offences on the UK's rail network, and what are the rates of sexual recidivism among offenders?

RQ2: What are the characteristics and risk factors associated with sexual recidivism on the rail network?

RQ3: Do characteristics of sex offenders differ, based on type of offending?

RQ4: What is the predictive validity of a machine learning algorithm in forecasting sexual recidivism on the rail network?

Research Design

This research was quantitative in nature. The initial phase of the study was descriptive, whilst the second phase, focused on evaluating the predictive validity of the machine learning model, was predictive.

Data and methodology

The study utilised secondary arrest data from the British Transport Police. A report was created to identify all individuals arrested for at least one sexual offence on the rail network between 1st April 2016 and 31st March 2024. Once these individuals were identified, the report returned details of all

their arrests (for any offences) that occurred within the same time frame. The final dataset included 3445 individuals, arrested for 11, 789 offences.

This study opted to use a 'Random Forest' algorithm to forecast recidivism. To prepare the data, the time point of prediction (at which the forecast begins) was selected as 1st April 2023. The predictor variables included several demographic features alongside previous arrest counts and crime harm score. The primary outcome variable was defined as a further arrest for a sexual offence within the 12 month follow up period and this was binary in nature. The Random Forest model was constructed in R-Studio (Version 4.4.1) using two data subsets; 70% (n= 2411) for training, and 30% (n= 1034) for testing.

Analytic methods

Descriptive statistics, such as mean and standard deviation, were used to analyse the data for RQ1, RQ2 and RQ3. Predictive modelling was employed (using Random Forest) to address the final research question. The Random Forest model was trained using the 'training' subset, and forecasted predictions were compared with actual outcomes in the 'testing' subset. The model's performance was evaluated using the number of true and false predictions to calculate accuracy, sensitivity, and specificity scores. The Gini Index was evaluated to determine variable importance and five-fold cross validation was conducted to assess the reliability of the model.

Key Findings

RQ1:

Sexual recidivism on the rail network is generally low – 3.6% within the 12 month follow up, rising to 10.4% over an 8-year period. The most common offence types include both contact, and non-contact offences – sexual assault on a female, outraging public decency and exposure.

RQ2:

Most offenders in the sample were male, White-North European, with an average age of 35 years old (SD= 12.6). Unemployed, previous number of total arrests, and previous sexual arrests were all associated with an increased likelihood of recidivism, as well as a more diverse range of sexual offending.

RQ3:

Minor differences were observed between subgroups of sexual offenders with those who arrested for digitally enabled sexual offences seemingly the most distinct. These offenders were generally older (Mean age = 37.9, SD= 14.4), had the greatest proportion of employment and highest number of sexual arrests. Furthermore, a pattern of like-for-like prediction was observed for contact, non-contact and digitally enabled offences, however, only the correlation for contact offences was statistically significant ($p = 0.03$).

RQ4:

The initial Random Forest model achieved an overall accuracy of 0.96, however, the substantial decrease in the specificity score between the training and testing data indicates possible overfitting. Given the potential consequences of false predictions, oversampling and under-sampling methods were utilised, alongside fine-tuning of the parameters. The final model maintained an accuracy score of 0.96, correctly identifying non-recidivists 97% of the time (specificity), but only identifying recidivists 55% of the time (sensitivity). The most important variables to the model's accuracy were previous arrests for contact offences and non-contact offences alongside the cumulative crime harm score. The least important variables were gender, drugs-related arrests and ethnicity.

Implications

The study proposes the benefits of using machine learning algorithms to improve recidivism predictions, however it is mindful of the ethical and moral concerns regarding implementation. Given the complexity of machine learning, it advocates further collaboration between practitioners and academics to help police develop reliable predictive models to forecast recidivism.

Acknowledgements

Firstly, I would like to thank my Cambridge Supervisors, Olivia Pinkney and Dr Sara Valdebenito, for their unwavering belief in me, their encouragement and for pushing me beyond my comfort zone. I also want to acknowledge my course mates who shared this journey with me. Their support and the friendships I have gained along the way have made this experience truly unforgettable.

I am also grateful to BTP, without whom this incredible opportunity would not have been possible. To my colleagues within the ViSOR team, thank you for graciously allowing me the time to focus on my studies. To DCI Sam Painter, your steadfast presence has kept me grounded throughout this journey – thank you for helping me stay on course. To Greg and the wider Data Science team, your patience and guidance has been a cornerstone to this work – I am truly grateful. To all other BTP colleagues I met along the way, thank you for helping make this thesis a reality.

Finally, to my friends and family who have supported me unconditionally, I am so grateful to each of you! Most especially Dan, your endless support, encouragement and patience has been at the heart of this accomplishment, and I could not have done it without you.

Contents

List of Tables

List of Figures

Chapter 1: Introduction

Violence Against Women and Girls ('VAWG') is a **national threat** in England and Wales (National Police Chiefs Council (NPCC) & College of Policing (COP), 2024). It is an umbrella term, encompassing offences that disproportionately effect women and girls, such as Sexual Offences, Domestic Abuse, Stalking and Harassment (COP & NPCC, 2024). Every year in the United Kingdom (UK), an estimated 1 in 12 women (two million) will be victim of VAWG offences (COP & NPCC, 2024). Globally, this rises to 1 in 3; however, due to issues of underreporting, the true scale of the problem is difficult to comprehend (Mannell, Lowe, Brown, Mukeriji, Devakumar, Gram, Jansen, Minckas, Osrin, Prost, Shannon & Vyas, 2022; COP & NPCC, 2024). The potential consequences of VAWG offences are devastating; 75% of sexual assault victims meet the criteria for Post Traumatic Stress Disorder, whilst instances of suicide following domestic abuse increase year on year (Dworkin, Jaffe, Bedard-Gilligan & Fitzpatrick, 2021; VKPP & NPCC, 2024). Concerningly, the number of police recorded VAWG crimes has risen by 37% in the last four years (COP & NPCC, 2024). The policing of such offences also received increased scrutiny following several high-profile incidents (Garg, 2023; Hohl & Stanko, 2022; VKPP, 2024). This heightened attention adversely affected trust and confidence in the policing of VAWG perpetrators and highlighted the need for improvements (Garg, 2023; Hohl & Stanko, 2022; VKPP, 2024; COP & NPCC, 2024).

Sexual Offences

In 2024, UK police recorded 195,315 victims of sexual offences, however the true figure was estimated at over one million, revealing the extent of potential underreporting (COP & NPCC, 2024). Despite experiences of crime generally decreasing, the Crime Survey for England and Wales (CSEW) report an increase in sexual offences (Office for National Statistics, 2024). Though they may not represent the highest *crime counts*, sexual offences are considered *high harm* and generate significant public concern (Hanson & Morton-Bourgon, 2005; Sherman, Neyroud & Neyroud, 2016). Individuals subject to unwanted sexual behaviour (USB) experience a range of negative effects, including physical, emotional, financial and psychological distress (Ariel, Langton, Peters, Webster & Assaraf, 2024a; Ariel, Ceccato, McDonnell, & Webster, 2024b).

Furthermore, sexual offences encompass a variety of behaviours, from contact offending (e.g. Rape), to non-contact offences (e.g. Exposure), as well as digitally enabled offences (e.g. Voyeurism) (Ceccato & Loukaitou-Sideris, 2022). As society advances, so does the nature and prevalence of sexual offences. For example, the accelerated growth and affordability of technology provides new opportunities for individuals to perpetrate sexual offences online (McAlister, 2014). In response, the legal system must adapt, continually amending existing legislation or introducing new offences such as 'Cyberflashing' (College of Policing, 2024). In recent years, the generally accepted definition of USB has broadened significantly, to include behaviours such as sexualised comments and suggestive gestures. Society's growing awareness of the diverse forms of USB, reflects a shift towards addressing these harmful behaviours and recognising potential patterns of escalation. This evolving understanding subsequently contributed to the implementation of new legislation regarding Sex-based Harassment (Protection from Sex-based Harassment in Public Act, 2023).

Sexual Offences on Public Transport

Transport networks provide a vital role in connecting communities and enabling individuals to travel to work, education and daily activities (Ariel et al, 2024b). Alarmingly, research indicates high rates of USB occur within public transport environments (Ariel et al, 2024a; Gekoski, Gray, Adley & Horvath, 2017; VKPP, 2024). Compared to other countries, the UK was initially reported to have one of the lowest prevalence rates of USB on public transport, however recent findings reveal almost three quarters (74.3%) of rail users have experienced USB on trains or in train stations (Ariel et al, 2024a; Gekoski et al, 2017). The motivation-facilitation model purports whilst most sexual offences are primarily motivated by sexuality factors (such as paraphilias) and facilitated by personality traits (such as impulsivity), they are also influenced by situational factors (such as access to potential victims) (Seto, Augustyn, Roche & Hilkes, 2023). Inherently, transport hubs present several situational factors that facilitate sexual offences (Ariel et al, 2024a). Firstly, females are more likely to rely on public transport, making it a particularly vulnerable setting, given the gendered nature of these offences (Forsdike, Ison, Hooker, Henry, & Taft, 2024; Williams, Malik & McTarnaghan, 2020). Busy, crowded services enable offenders to easily conceal their behaviour and provide a sense of anonymity, whilst quiet late-night services or unmanned platforms isolate vulnerable victims (Ariel et al, 2024a). Furthermore, due to the transient nature of public transport, offenders can avoid detection by regularly switching services, whilst victims will likely remain on the service until their intended destination (Ariel et al, 2024a; Gekoski et al, 2017). This high prevalence of USB within public transport environments can have significant economic and social consequences, resulting in females feeling more fearful and likely to adopt avoidance strategies (e.g. find another method of travel, or decide not to travel at all) (Ceccato & Loukaitou-Sideris, 2022).

Police response to VAWG

In response to this growing epidemic, the NPCC advocated for a '**whole-system approach'** and more recently, the government pledged to utilise all available tools to target offenders (Labour Party, 2024; COP & NPCC, 2024, p5). VAWG was subsequently introduced to the Strategic Policing Requirement, ensuring the police response is prioritised alongside counter terrorism (Home Office, 2023; COP & NPCC, 2024). UK policing also adopted the '4P' approach, setting out key commitments to '*prepare', 'pursue', 'protect*' and consequently '*prevent*' offences (COP & NPCC, 2024, p4). With an estimated 2.3 million perpetrators of VAWG offences annually, police need to target resources to disrupt the "*power few"* - the small percentage of offenders responsible for the largest proportion of harm (COP & NPCC, 2024; Sherman, 2007, p299). To successfully '*pursue'* high-harm sexual offenders, the NPCC and COP propose using *"big technology"* (COP & NPCC, 2024, p. 24). The Operation Soteria National Operating Model (NOM), developed to improve Rape and Serious Sexual Offence (RASSO) investigations, also outlines the importance of using existing police data strategically, to disrupt repeat offenders (Hohl & Stanko, 2022; College of Policing, 2023). One working example, is the

Metropolitan Police's VAWG100, which uses data analytics to identify the top 100 VAWG offenders, and optimise resource allocation (COP & NPCC, 2024, p.25). Early evaluation shows the VAWG100 has successfully brought perpetrators to justice, protected victims and championed a whole-system approach (COP & NPCC, 2024).

Existing Risk Assessments

To identify these high-harm individuals, police forces previously relied on clinicians' professional judgements for recidivism predictions; however, research revealed these predictions were no more accurate than random chance (Bengston & Långström, 2007; Bonta & Andrews, 2007; Craig & Rettenberger, 2016; Howard, 1998). In current practice, most police forces utilise one of many existing risk assessment tools to predict an individual's level of risk (such as the Risk Matrix 2000, Offender Assessment System, Sexual reoffending predictor or Active Risk Management System) (College of Policing, 2017; Howard & Wakeling, 2021). These tools transformed risk assessments, by reducing dependency on expert opinion and incorporating static and dynamic risk factors, empirically linked to recidivism (Bonta & Andrews, 2007; Hanson, 1998). This shift enabled a more evidence-based approach to predicting further offence and empowered police to prioritise offenders, allocate resources and target interventions more effectively. Individuals identified as high-risk were subject to increased scrutiny, and more prohibitive conditions, compared to those assessed as lower-risk. To justify subjecting individuals to increased scrutiny and imposing such restrictive conditions, police require a robust and defensible framework (Hanson, 1998). Whilst available evidence largely supports the moderate predictive validity of existing risk assessment tools, they are time-consuming and lack external validity across the evolving nature of sexual offences (Fellows, 2024a; Kewley, Osman & McGuiness, 2020; Långström, 2004; McNaughton Nicholls & Webster, 2014; Parent, Guay & Knight, 2012; Tully, Chou & Browne, 2013). Additionally, these risk assessment tools are only utilised *following conviction* for a sexual offence. This differs significantly to other VAWG offences, such as Domestic Abuse, Stalking and Harassment, where risk assessments are a key priority, completed at the earliest opportunity (College of Policing, 2016; Fellows, 2024a). Conducting these risk assessments following conviction limits opportunities for early interventions, particularly given the low charge rates for sexual offences in the UK (9%) (Prime Ministers Office, 2024). This process risks the misallocation of police resources and potential missed opportunities for prevention strategies (Fellows, 2024a; Whitten, 2024). For example, Whitten (2024) revealed most Sexual Risk Orders were only applied for following an average of 2.8 failed prosecutions. Alarmingly, research indicates that the misallocation of such resources can result in harmful consequences, potentially increasing recidivism rates of would-be low-risk offenders (Andrews & Bonta, 2014; Bonta & Andrews, 2007).

Artificial Intelligence and Machine Learning

More recently, emerging research provided support for the use of artificial intelligence (AI) and machine learning (ML) to effectively predict criminal recidivism (Travaini, Pacchioni, Bellumore, Bosia & De Micco, 2022). AI is a broad, multidisciplinary field that enables digital technology to undertake tasks typically associated with human intelligence (Bland, 2020; O'Connell, 2024; NPCC, 2024). By reasoning, analysing, and interacting with data, AI mimics human cognition and delivers valuable insights (Bland, 2020; O'Connell, 2024). ML is a subset of AI, whereby digital systems identify patterns in data (*input*) to make predictions (*outputs*) (Bland, 2020). Using iterative processes, ML algorithms learn from the data, continually refining their performance without having to be explicitly reprogrammed (Bland, 2020; O'Connell, 2024; NPCC, 2024). Early research indicates that ML methods achieve 'good performance' in predicting criminal recidivism, with an average accuracy score of 0.81 (where 0 reflects no ability to predict, and 1 demonstrates perfect predictions) (Travaini et al, 2022). Additionally, ML algorithms can process large datasets and identify evolving patterns much faster than current practices (Farayola, Tal, Connolly, Saber & Bendechache, 2023). Given current pressures on UK policing to proactively manage increasing demand with diminishing resources, ML algorithms offer a promising alternative to effectively identifying high-risk individuals

(Bland, 2020). It also aligns with Operation Soteria NOM and the NPCC's commitment to 'pursue' offenders by using big technology and existing police data.

The current study

The current study used existing police data and ML algorithms, to enhance the accuracy of sexual recidivism predictions on the rail network. Using these data-driven forecasts, this study hopes to identify high-harm, repeat sex offenders and target interventions, more effectively. This will add to the emerging literature, and support policy decisions regarding sex-offender risk assessments and interventions. The following research questions have been devised to deliver targeted results:

> **RQ1:** What is the nature and prevalence of sexual offences on the UK's rail network, and what are the rates of sexual recidivism among offenders?
>
> **RQ2:** What are the characteristics and risk factors associated with sexual recidivism on the rail network?
>
> **RQ3:** Do characteristics of sex offenders differ, based on type of offending (e.g. contact, non-contact, digitally enabled or attempted)?
>
> **RQ4:** What is the predictive validity of a machine learning algorithm in forecasting sexual recidivism on the rail network?

The next section will review relevant literature to identify gaps that this study seeks to address. This is followed by a detailed explanation of the methodology and data, a presentation of the study's findings and a discussion of the practical and policy implications.

Chapter 2: Literature Review

Introduction

To understand the societal impact of repeat sex offenders, this literature review begins by exploring the prevalence of sexual recidivism, particularly within transport environments. The characteristics of sex offenders are considered alongside risk factors for repeat offending, with a focus on the similarities and differences between various offence types. The research regarding existing risk assessment tools is then reviewed, considering the strengths, limitations and effectiveness in identifying risk. Finally, the literature review considers emerging research examining the predictive validity of ML algorithms to forecast criminal recidivism and addresses critical concerns regarding the ethical and moral implications of algorithmic policing.

Sexual recidivism

In England and Wales, criminal recidivism costs the economy approximately £18.1 billion, with far-reaching consequences for victims and communities (Travaini et al, 2022). Despite the significant harm caused by repeat VAWG offenders, a lack of robust data exists regarding these individuals (COP & NPCC, 2024). Whilst identifying VAWG recidivists is challenging (due to underreporting) by failing to recognise the prevalence of repeat offending, the impact may be underestimated (Furby, Weinrott & Blackshaw, 1989; Hohl & Stanko, 2022).

Generally, reconviction rates for sex offenders are ‘low’, though there are variations across different studies (Falshaw, Bates, Patel, Corbett & Friendship, 2003; Loukaitou-Sideris & Ceccato, 2021). From a global perspective, Hanson and Morton-Bourgon (2009) conducted a meta-analysis using a strict inclusion and exclusion criteria. The analysis incorporated 118 studies from various countries examining rates of sexual, violent and general recidivism amongst convicted sex offenders (Hanson & Morton-Bourgon, 2009). Using an average follow-up period of 5 years 10 months, the sexual recidivism rate was 11.5% (Hanson & Morton-Bourgon, 2009). These findings reflect an earlier meta-analysis from the same authors, which reports a sexual recidivism rate of 13.7% over a 5-year follow-up (Hanson & Morton-Bourgon, 2005). Notably, the reoffending rate was higher for violent (including sexual) offences (19.5%) and increased further for *general recidivism* (33.2%) which supports the notion that sex offenders are not specialist offenders, but often commit a range of crimes (Hanson & Morton-Bourgon, 2009; Hohl & Stanko, 2022).

To understand the '*base rate'* of sexual recidivism in the UK, Craig, Browne, Stringer and Hogue (2008) conducted a systematic review of 'sex offender reconviction' literature. Sixteen UK studies were identified using computer searches and reference lists, including published and non-published papers from 1991-2005 (Craig et al, 2008). During a 2-year follow-up period, the average sexual recidivism rate was 5.8%, rising to 6.9% over 4 years and reaching 17.4% for follow-up periods of 6 years or more (Craig et al, 2008). The substantial increase observed in the 6-year follow-up could be partly attributed to the absence of a defined end date (Craig et al, 2008). These findings appear to suggest sexual reconviction rates increase over time; however, this pattern may be influenced by data aggregation and the cumulative nature of reconviction rates. In contrast, a longitudinal study that observed a cohort of sex-offenders over a 20-year period, reported sexual recidivism risk was highest immediately after release from custody, but decreased significantly over time (Hanson, Harris, Helmus & Thornton, 2014; Helmus, 2018). The risk of recidivism reportedly halved for every five years the individual remained in the community without reoffending (Hanson et al, 2014; Helmus, 2018).

Transit systems are reportedly the second most frequent setting for sexual harassment (Williams, Malik & McTarnaghan, 2020). Despite this, there is little research regarding the nature and prevalence of sexual recidivism within public transport environments. Existing literature primarily focuses on victimisation, leaving a knowledge gap regarding the rates of reoffending and making it challenging to establish reliable recidivism rates (Ariel et al, 2024a). A rapid evidence assessment conducted by Gekoski et al (2017) reported that females experiences of sexual offences or harassment on public transport ranged between 15% to 95% globally. These findings were supported by a subsequent exploratory study, whereby college students in multiple continents were surveyed to assess the prevalence of sexual harassment in transit environments (Loukaitou-Sideris & Ceccato, 2021). These survey responses also reflected a vast range of experiences of sexual harassment, varying from 11% to 89% (Loukaitou-Sideris & Ceccato, 2021). The wide variations reported in both studies may be attributed to different geographical contexts, legislation, cultural definitions of sexual harassment and the research designs (Loukaitou-Sideris & Ceccato, 2021). More recently, Ariel et al (2024b) utilised surveys to examine experiences of USB on the UK's rail network. The findings revealed 73.4% of participants had personally experienced USB whilst on trains, or in train stations across Great Britain (Ariel et al, 2024b). The most common forms of USB were '*verbal comments or jokes of a sexual nature'* (53%), '*unwanted touching or groping'* (51.8%), '*sexual gestures/mimes'* (30.5%) and '*indecent exposure'* (20.5%) (Ariel et al, 2024b, p.16). These findings align the VAWG Strategic Threat Risk Assessment (STRA), whereby sexual assaults (touching/groping) accounted for approximately half of all RASSO (48%) (VKPP, 2024). However, it also indicates levels of indecent exposure may be higher on the rail network compared to the national average (8%) (VKPP, 2024).

In recidivism studies, terms such as 'reoffending', 'rearrest' 'reconviction' and 'recidivism' are often used interchangeably despite having distinct meanings (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2009). 'Recidivism' generally has the broadest definition, encompassing any offence-related behaviours (Falshaw et al, 2003). Multiple measures can be utilised for 'recidivism', such as named suspects, arrest or reconviction data. However, the use of different measures hinders comparisons across studies, as each may adopt varying definitions of further offending (Craig et al, 2008; Hanson & Morton-Bourgon, 2009). For example, the COP and NPCC reported 45.6% of *named suspects* were named in more than one sexual offence (COP & NPCC, 2024). This is much higher than the sexual *reconviction* rates reported earlier and demonstrates how reported rates vary considerably depending on measures used. Each measure presents distinct advantages, and limitations. Arrest data likely captures more incidents, however it may over-estimate recidivism as not all arrests result in conviction (Whitten, 2024). Conversely, reconviction data may be distorted by low charge rates and potential plea bargaining at court (Craig et al, 2008).

Characteristics and Risk Factors

By analysing existing data, researchers can identify common characteristics, and risk factors associated with sexual recidivism. This has clear practical implications, particularly in development of effective risk assessment tools (Emeagi, Sullivan, Landsiedel, Craik & Howard, 2024; Hanson & Morton-Bourgon, 2005; Seto et al, 2023). A '*characteristic*' is considered a general attribute, whilst a '*risk factor*' refers to specific characteristics associated with increased likelihood of reoffending (Seto et al, 2023). These factors are often categorised as either 'static' or 'dynamic'. *Static characteristics* refer to historical attributes that do not change or can only change in one direction (e.g. age at time of first offence or number of previous convictions) (Bonta & Andrews, 2007; Seto et al, 2023). *Dynamic factors*, on the other hand, are enduring factors that can change bi-laterally over time (e.g. employment or marital status) (Bonta & Andrews, 2007, Craig, Browne, Stringer, & Beech, 2005; Seto et al, 2023; Tollenaar & M van der Heijden, 2013). Static characteristics help to identify individuals who pose a long-term risk of sexual recidivism and may require intensive supervision (Bonta & Andrews, 2007; Craig et al, 2005; Seto et al, 2023). Whilst dynamic risk factors can inform specific interventions to reduce risk, as described by the Risk-Factor Prevention Paradigm (Farrington, 1995; Farrington, Tofi & Piquero, 2015; Seto et al, 2023).

Regarding *characteristics*, it is widely recognised that VAWG crimes are gendered in nature, with a large proportion perpetrated by males (Fellows, 2023; Garg, 2023; Hohl & Stanko, 2022; VKPP, 2024). Across all VAWG offences, males account for 75% of *suspects*, however for sexual offences specifically, males represent over 90% of *prosecutions* (Ministry of Justice (MOJ), Home Office (HO) & the Office for National Statistics (ONS), 2013; VKPP, 2024). Though this percentage is very high, female sex offenders do exist. A meta-analysis conducted by Cortoni, Babchishin and Rat (2017) reported female sex offender prevalence rates ranged from 0.4%-6.8% within police records.

Interestingly, when using victimisation surveys, the range of female sex offenders' prevalence increased to 3.1%-24.4%, suggesting female perpetrators of sex offences may be less likely to be reported (Cortoni, Babchishin & Rat, 2017). Age is another characteristic commonly utilised within risk assessment tools and is considered a "robust predictor" (Rice & Harris, 2014, p. 151). The most common age group of males with convictions for sexual offences was reported between 30-39 years old (Emeagi et al, 2024). This is supported by the VAWG National Statement which states the average age of RASSO suspects is 37-years old (COP & NPCC, 2024). Beyond this, the literature generally supports the notion that sexual offending decreases as age increases, however this varies slightly across different subgroups of sexual offences (Craig et al, 2008; Hanson, 2002). With regards to ethnicity, the available research presents mixed findings. MOJ, HO and ONS (2013) report the largest proportion of sexual offenders are White (78%), followed by Black (9.9%) and Asian (9.7%). However, research on *multiple perpetrator* sexual assault, revealed the opposite, with almost half sexual offenders were Black (46.8%), followed by White (39.6%) (Morgan, Brittain and Welch, 2012). These findings may reflect differences in offender demographics between single and multiple perpetrator offences.

Regarding *risk factors*, an early meta-analysis by Hanson and Bussière (1998) identified the strongest predictors of sexual recidivism were measures of sexual deviancy (*such as a sexual interest in children*), followed by criminal lifestyle (*including the number of previous offences*). Individuals previously convicted for sexual offences were more likely to reoffend, particularly those who had targeted strangers or male victims, those who began sexual offending at an early age, and those who committed a diverse range of sexual crimes (Hanson & Bussière, 1998). Prior non-sexual offences were moderately positively correlated with sexual recidivism, however prior violent offences showed little association, though both measures had large variability across studies (Hanson & Bussière, 1998). These findings were supported in a further meta-analysis by Hanson & Morton-Bourgon (2005, p.1154) who also identified deviant sexual preferences and anti-social orientation (*such as unemployment and substance abuse*) as "major predictors" of sexual reoffending. Interestingly, the meta-analysis revealed many of the issues typically targeted during sex offender treatment programmes (*such as adverse childhood experiences*) had little or no relationship with sexual recidivism (Hanson & Morton-Bourgon, 2005). More recently, a systematic review by Seto et al (2023) identified atypical sexuality (*including paraphilic sexual interests*), self-regulation problems (*such as unstable employment/housing*), anti-social cognitions (*such as non-compliance with interventions*) and relationship problems were linked to increased sexual recidivism. Conversely, stable relationships and employment were recognised as potential protective factors, reducing the likelihood of further sexual offending (Seto et al, 2023). The association between unstable housing/employment and sexual recidivism is particularly concerning, given that many convicted sex offenders may face difficulties in obtaining work or housing due to their criminal history (Cann, 2017). Seto et al (2023) also identified '*hostility towards women*' as a predictor of sexual recidivism, which may be unsurprising, given the gendered nature of these offences. These risk factors reflect the "*Central Eight*" major predictors recognised in Bonta and Andrew's Risk-Need-Responsivity (RNR) Model for predicting criminal behaviour (Bonta & Andrews, 2007). These eight factors, identified through empirical research, summarize the criminogenic risk and/or need factors associated with reoffending, including antisocial behaviour, personality, cognition and associates, alongside family, education, leisure and substance misuse (Andrews, Bonta & Wormith, 2006). Critically, the impact of intercorrelations between these factors are largely unknown so careful consideration is required when weighting each factor within risk assessment tools (Hanson & Bussière, 1998).

Unfortunately, the predictive validity of risk factors is not consistent and their effectiveness in forecasting recidivism may vary depending on the type of offence (Craig et al, 2005; Emeagi et al, 2024). Given the breadth of offences encompassed by 'sexual offending,' it is crucial to understand how the characteristics and risk factors differ across various offence type (Emeagi et al, 2024). In a meta-analysis, Babchishin, Hanson and Hermann (2011) concluded that online offenders were more likely to identify as White compared to offline offenders and were also generally younger (38.6 years) compared to offline offenders (43.6 years). A further meta-analysis by Babchishin, Hanson and

VanZuylen (2013) reported online offenders had greater academic achievements and greater levels of anti-sociality compared to offline offenders, including a higher number of prior offences. Despite these differences, Jung, Ennis, Stein, Choy and Hook (2012) concluded contact, non-contact and child pornography offenders share more similarities than differences. More recently, Emeagi et al (2024, p.29) observed that sexual offending behaviours are becoming increasingly distinct from one another and a "pattern of like for like prediction" is emerging. This suggests that prior *contact* sexual offences predict future *contact* offences, and so on. As such, a previous conviction for one type of sexual offence may decrease the likelihood of a further arrest for a different sexual offence type (Emeagi et al, 2024).

Existing Risk Assessment Tools

Within the criminal justice system (CJS), various risk assessment tools are employed to identify repeat sex offenders (Tully, Chou & Browne, 2013). Over the years, these risk assessment methods have evolved from professional judgements, to utilising evidence-based instruments incorporating both static and dynamic risk factors (Bonta & Andrews, 2007). These risk assessments are integral to the management of sex offenders, serving as the foundation for decisions regarding resource allocation and prioritisation of treatments (Craik, Han, Sullivan, Landsiedel, Travers Spaull & Howard, 2024; Emeagi et al, 2024; Tully, Chou & Browne, 2013). To effectively use these tools to identify high-harm offenders and reduce recidivism, the risk principle of the Risk-Need-Responsivity (RNR) Model emphasizes the importance of aligning interventions with the level of assessed risk (Andrews & Bonta, 2010).

Many risk assessment instruments used to predict sexual recidivism are classified into two main methodological approaches: clinical and actuarial (Craig et al, 2008). Clinical risk assessments, rely on clinicians' judgement of risk, drawing on their experience and knowledge of offending behaviour (Craig et al, 2008). These can be structured, or unstructured, and enable a person-centred, ideographic approach. Clinical risk assessments benefit from the inclusion of dynamic risk factors such as the individual's engagement in available treatment programmes (Tully, Chou & Browne, 2013). Actuarial risk assessments, on the other hand, employ a nomothetic approach, scoring pre-determined risk factors, previously associated with sexual recidivism (Bengston & Långström, 2007; Tully, Chou & Browne, 2013). Offenders are scored based on the presence or absence of such factors and these scores summed to calculate the assessed risk level (Craig et al, 2008; Tully, Chou & Browne, 2013). The scoring of risk factors in actuarial approaches seeks to reduce bias and enhance consistency when utilised by different practitioners (Craik et al, 2024).

Within the literature, there is large debate surrounding the benefits and disadvantages of each approach, however it is widely accepted that actuarial methods deliver greater overall predictive validity (Bonta & Andrews, 2007; Craig et al, 2005; Parent, Guay & Knight, 2012). In three different meta-analyses by the same authors, actuarial risk assessment tools predicted sexual reconviction more accurately than non-actuarial or clinical risk assessments (Hanson & Morton-Bourgon, 2004, 2007 & 2009). Within the most recent meta-analysis, actuarial measures specifically designed for sexual recidivism demonstrated the strongest predictive accuracy (d. = 0.67), followed by actuarial measures intended for general recidivism (d.= 0.62) (Hanson & Morton-Bourgon, 2009). Despite the moderate predictive validity of actuarial risk assessments, they are not without flaw and are criticized for reliance on static risk factors (Craig et al, 2005). Research indicates actuarial risk assessments lack external validity, as the findings are not generalizable across all sex-offender sub-groups (Långström, 2004). For example, Långström (2004) revealed actuarial risk assessments could not differentiate between recidivists from non-recidivists in African Asian offenders (Långström, 2004). Additionally, Parent, Guay and Knight (2012) reported the predictive accuracy of actuarial risk assessments varies depending on the type of offender to which they are applied. Furthermore, studies indicate actuarial risk assessment tools may also lack internal validity as the same offender can be assessed differently, depending on which instrument is used, in some cases being risk-assessed as the lowest risk category on one measure, but almost the highest rank on another (Barbaree, Langton & Peacock, 2006). The lack of a consistent method for identifying high-harm sex offenders, coupled with the increasing variety of risk assessment tools and emerging distinctions between offender types, risks undermining effective offender management.

Unfortunately, few studies exist that directly compare the use of clinical versus actuarial approaches, using the same population of sex offenders (Sjöstedt & Grann, 2002). In one such study, Bengston and Långström (2007) compared the predictive validity of unstructured clinical judgement to an actuarial approach on a sample of 121 adult male sex offenders. factors. The study reported when using an open-ended time frame for recidivism *"the predictive accuracy of unguided clinical judgement did not exceed chance"* for neither sexual nor violent reconviction (Bengston & Långström, 2007, p.135). Conversely, the actuarial measures predicted sexual, severe sexual and violent reconviction more accurately than chance (Bengston & Långström, 2007). Interestingly, when the recidivism time frame was reduced to a fixed 2-year period, none of the prediction methods yielded significant results, however this may be attributed to the low reconviction rates during this timeframe (Bengston & Långström, 2007). Notably, where rates of reoffending are low, the likelihood of false positives increased significantly (Craig et al, 2005; Tully, Chou & Browne, 2013). The potential consequences of false predictions are significant, with potentially harmful offenders being subject to low levels of scrutiny (Boer & Hart, 2008; Fellows, 2023). Generally, existing risk assessment tools use binary or categorical outcomes such as low, medium, high or very high risk (Emeagi et al, 2024). However, what does this really tell us? To effectively help police to target resources, these tools could benefit from outcomes that incorporate the potential imminence, severity, and frequency of further offending (Sjöstedt & Grann, 2002).

Machine Learning (ML) Algorithms

Given the clear policy implications of accurate recidivism predictions and the limitations of existing tools as outlined above, researchers and practitioners are constantly seeking improvements (Etzler, Schönbrodt, Pargent, Eher & Rettenberger, 2024). The use of statistical modelling to forecast criminal recidivism is not novel; however recent advancements reveal ML algorithms could provide more efficient and effective alternative risk assessment methods (Travaini et al, 2022). Though the literature in this area is growing, the application of ML algorithms to forecast risk within the CJS remains limited (Travaini et al, 2022). In a systematic review by Travaini et al (2022), a strict criterion was used to identify 12 studies that applied various ML methodologies to predict *general* recidivism. To effectively compare the different ML models, the review evaluated 'accuracy' (ACC) (*how often the algorithm correctly classified data points*), and the 'area under curve' (AUC) (*the model's ability to distinguish between recidivists and non-recidivists*) (Travaini et al, 2022). Across the 12 studies, the average ACC was 0.81, and the mid-score for AUC was 0.74 (*0 meaning the model predicted poorly, 1 meaning the model predicted perfectly).* This indicates ML models have moderate to high predictive validity for *general* recidivism, and perform better than actuarial risk assessment tools; however, there was a wide range in scores across different ML models.

The literature regarding the use of ML algorithms to forecast *sexual reoffending* specifically is limited, as this remains an emerging area of research. In the aforementioned systematic review, three of the twelve studies examined *sexual recidivism*. Among these three studies, the highest ACC score was 0.96, whilst AUC ranged from 0.71 to 0.77, which suggests that ML models may outperform existing tools in identifying repeat sexual offenders (Travaini et al, 2022). However, in one of these studies, Tollenaar and van der Heijden (2011) compared several ML, data mining and classical statistical methods in predicting sexual reoffending. The study revealed the best model at forecasting sexual recidivism was linear discriminant analysis (LDA), a classical statistical method (AUC = 0.725, ACC = 0.955) (Tollenaar and van der Heijden, 2011). Tollenaar and van der Heijden (2011) subsequently concluded that ML models did not provide a significant advantage over conventional statistical methods, highlighting issues with calibration, interpretability, and consistency. More recently, Etzler et al (2024) used actuarial risk assessment scores of a small sample of male sex offenders (N=511) as predictor variables in a Random Forest (RF) ML algorithm. The study reported the RF model did not demonstrate “superior predictive performance” when compared with conventional methods, also citing issues with transparency and interpretability (Etzler et al, 2024, p2).

The application of ML algorithms to support police decision making is contentious due to concerns regarding ethics, farness and potential bias (Farayola et al, 2023; Travaini et al, 2022). Since ML algorithms learn from patterns within data, they are vulnerable to biases within the original datasets such as racial, selection or labelling bias. Additionally, due to the complex nature of ML models, it is difficult to articulate how they convert raw data into predicted outcomes or classifications (Travaini et al, 2022). To address these concerns, several frameworks have been developed to encourage the ethical and responsible use of AI, including the European Commission’s “Seven Requirements for trustworthy AI” and the NPCC’s Principles of AI in Policing (Farayola et al, 2023, p.4; NPCC, 2024b). These frameworks highlight the need for transparency, accountability, technical robustness, fairness and lawfulness (Farayola et al, 2023; NPCC, 2024b). The EU Requirements also emphasize the importance of human oversight in AI activities *(human in command*), particularly during the design cycle *(human-on-the-loop)* and decision-making process *(human-in-the-loop*) (Farayola et al, 2023). In a systematic literature review, Farayola et al (2023) examined 69 papers to explore the ethics and trustworthiness of AI in predicting criminal recidivism. The authors proposed four additional requirements – consistency, reliability, explainability and interpretability - to complement the seven outlined by the European Commission (Farayola et al, 2023). These requirements are essential for developing robust, trustworthy and ethical AI models for forecasting recidivism in the CJS (Farayola et al, 2023). Amongst the 69 papers included in the literature review, most addressed issues of fairness, interpretability, and transparency, however none explored privacy, data governance or accountability (Farayola et al, 2023). Within the review, several challenges hindering the implementation of ML risk assessment models were also identified, including dataset limitations, lack of standardisation and inconsistent metrics (Farayola et al, 2023). To minimize the potential risk of introducing bias within ML algorithms, it is crucial to consider the quality and representativeness of the initial dataset, the importance of predictor variables, and the definition of outcome variable. Addressing potential sources of bias and ensuring fairness at each step is essential for developing an ethical and reliable model.

Chapter Summary

This chapter reviewed the literature on four key topics relevant to the research questions. Whilst sexual recidivism rates are generally considered to be low, comparisons across different studies are hindered by varying definitions of repeat offending. Notably, there is a gap in the literature regarding sexual recidivism within transport environments, as most existing research focuses on victimisation. The available research indicates most sex offences are perpetrated by young males, with sexual deviance and antisocial traits consistently identified as the strongest predictors of recidivism. However, emerging research suggests characteristics and risk factors may vary across different offence types. Regarding existing risk assessment tools, the literature largely supports their moderate predictive accuracy. Although, concerns remain regarding external and internal validity, as the findings are not generalisable, and individual scores differ depending on the risk assessment tool used. ML algorithms initially show promise, particularly in predicting general recidivism. However, limited research exploring their application for forecasting sexual recidivism suggests these methods may not provide significant predictive advantages. Lastly, the principles of trustworthy AI were explored to align with the objectives of the current study.

Chapter 3: Methods

Introduction

In alignment with the requirements of trustworthy AI, this chapter provides a comprehensive explanation of the methodology employed in the current research. It details the research design and process of data collection, including steps taken to prepare and organise the data prior to analysis. Following this, the chapter elaborates on the analytic methods used to address each research question and explains the construction, testing and analysis of the ML model.

Research Design

The initial phase of this study was descriptive, utilising a large dataset (N=3445) to systematically describe the characteristics of individuals arrested for sexual offences on British Transport Police (BTP) jurisdiction. It also described the nature of sexual offences on the rail network and calculated the rate of recidivism. The second phase of the study was predictive, employing prediction models to analyse patterns in the data to determine the likelihood of a further arrest for a subsequent sexual offence. The study was reliant on quantitative data, including both continuous and categorical variables. Given the sensitive nature of the research, ethical approval was obtained from University of Cambridge Institute of Criminology (Appendix 1), permission was granted by the BTP Research Board, and a Data Protection Impact Assessment was authorised prior to data collection.

Data Collection

To address the research questions, this study utilised secondary data from BTP. BTP are a specialized police force, responsible for policing the rail network across England, Scotland and Wales (British Transport Police Authority, n.d). Given their national remit, BTP's jurisdiction is structured into three geographic regions: 'B Division' which polices London and the Southeast of England, 'C Division' which polices the Pennines, Midlands, Southwest and Wales, and 'D Division' which is responsible for policing Scotland.

To ensure the data was appropriate for the research objectives, a new report was created within 'Business Objects', also known as BOXI. This is the reporting tool of choice within BTP and produces reports from information held on the BTP Occurrence Management System (Niche). The report was designed to identify **all** individuals arrested for *at least one sexual offence* in B-Division or C-Division between 1st April 2016 and 31st March 2024. These were detected using crime codes listed in BTP Force Crime Group for sexual offences (02A). Once these individuals were identified, the report returned details of all their arrests (for any offences) that occurred within the same time frame. The eight-year period was selected to maximise sample size, reflecting BTP's transition onto Niche in 2016 and encompassing data up to the most recent financial year. Arrests for offences in Scotland (D-Division) were excluded due to legislative differences.

Arrest data was used, instead of charge or conviction data, as it is frequently updated and readily accessible through automated reports. It excludes instances without an identified suspect, which would not meaningfully contribute to the study. Furthermore, using arrest data supports the identification of *early* predictors and promotes early risk assessments for sex offenders, offering potential practice and policy benefits.

Creating Variables

Person Data

Once the relevant individuals were identified, a range of variables were incorporated into the report from data held on Niche. To facilitate the analysis of characteristics, demographic variables such as age, gender, ethnicity, and occupation were included. The report also contained the individual's residential City/Town and County but omitted street names and postcodes to prevent identification beyond the research's purpose. When incorporating the occupation variable, the report initially returned 127 distinct job types, however many of these were very similar (such as ‘cleaner’ and ‘contract cleaner’). To simplify analysis, these were consolidated into 29 broader occupational categories (e.g. cleaner). Additionally, a new column of data was created to indicate if the individual was ‘*employed*’, ‘*unemployed’,* ‘*retired*,’ ‘*student*,’ or classified as ‘*other’.* This information was collated into one worksheet in BOXI named ‘Person data’ which initially contained 5646 rows of data.

Arrest Data

The new report also contained operational details of all arrests for individuals listed in the ‘*Person data’* worksheet, including arrest dates, offence dates, offence categories and offence locations. To support practical implications, the report captured the precise railway location where the offence occurred (e.g. train or platform) and distinguished the type of station (e.g railway station or underground station). To enable cross-referencing with the Cambridge Crime Harm Index (CCHI), the report returned the arrest reason and corresponding Home Office crime code. Additional variables that were not essential to the analysis such as ‘disposal reason,’ were also included to facilitate the identification of duplicate offences and reduce the risk of bias. This information was collated into a second worksheet in BOXI, named ‘*Arrest Data’* and initially contained 11, 809 rows of data.

Once these variables had been generated in the BOXI, the report was downloaded into Microsoft Excel, and a unique ID number was created for all individuals to ensure privacy and data governance. The dataset was stored within the author’s work-issued One Drive and access limited to the project team. Further variables of interest were then calculated in Microsoft Excel. Given the high-harm nature of sexual offences, and to ensure crime counts were not considered in isolation, the CCHI score for each offence was calculated using the VLOOKUP function. A pivot table was used to determine the cumulative CCHI score for each individual in the ‘person data’ worksheet. To assist with answering RQ3, a new column of data was created to classify the **subgroups** of sexual offences. All arrests categorised as ‘Sexual’ were assigned into one of the following subgroups (Appendix 2):

- Contact (*sexual offences involving physical touch*),

- Non-contact (*sexual offences without physical touch*),
- Digitally Enabled (*sexual offences facilitated by technology or internet access*),
- Attempted (*where the intent was evident, but the act was not completed*)
- Other (*for sexual offences that did not clearly fit into the previous categories*).

Finally, due to the unique nature of BTP jurisdiction – as a **national** force with a specialist focus on **transport**, an additional column was created to identify if the offence occurred in the same location as the offender's residence.

Data Preparation and Organisation

After creating all variables, the data was checked for duplicates and anomalies. Unfortunately, the inclusion of the 'occupation' variable resulted in multiple duplicates within the *'person data'*. This occurred due to numerous occupations being recorded on Niche for each individual, particularly when their occupation changed over several years of reoffending. To overcome this, all duplicates were manually reviewed by the author. The duplicated individuals were cross-referenced with Niche, and only the occupation recorded at the time of the arrest remained in the dataset. A similar issue was also identified within 'self-defined ethnicity', where multiple individuals recorded both a blank answer and a positive response. Again, these were manually reviewed by the author and the most complete row of data (e.g. where the self-defined ethnicity had been provided) remained in the dataset. Overall, 1180 duplicates were removed from the '*person data'* worksheet, reducing the number of rows to 4466. Some individuals recorded multiples rows of data, due to offending at different ages, however the decision was made not to condense these, to accurately reflect the age at time of arrest.

Duplicates were also located in the '*arrest data*' worksheet, primarily due to input errors on Niche, such as officers incorrectly creating too many offences. These were identified using the '*disposal reason*' column, as this highlighted comments entered by Officers stating 'incorrect,' 'added in error' or similar. These offences were cross-referenced with Niche to ensure they had been recorded accurately and any duplicate offences were removed. Identifying duplicate offences was essential, as their inclusion could have skewed results and inflated offenders' CCHI scores. Overall, 20 offences were removed from the '*arrest data'* worksheet. The remaining data included 3608 individuals (4466 rows of data), responsible for 11, 789 offences.

Once all duplicates had been removed, each variable was analysed for missing data. Within 'person data', gaps were identified in gender (89, *2%*), age (8, 0.*2%*), ethnicity (468, *10.6%*) and self-defined ethnicity (794, *17.8%*). Where possible, this data was filled through cross-referencing with Niche, however where the data was not available, the decision was made to categorise these as 'Unknown', rather than remove the entire row of data. Furthermore, several offences (49) within the arrest data worksheet lacked corresponding CCHI scores. Many of these were low-level Railway Byelaw offences but some were immigration and sexual offences. While the author considered calculating CCHI scores, due to time constraints it was decided to record missing scores as 0, rather than remove these offences from the study entirely.

Figure 9*. Illustration of the time point of prediction within the current dataset (Fellows, 2024b).*

In preparation for the algorithm, the arrest data was divided into two; arrests made prior to April 2023 and those made after, as illustrated in Figure 1. To create the predictor variables, the total number of arrests prior to April 2023 was summed for each individual and categorised by offence type. To create the outcome variable, the number of sexual arrests that occurred post April 2023 were calculated for each individual. Any individual who had **only** been arrested for a sexual offence within the follow up period were removed from the dataset, as they did not meet the criteria of sexual ‘**recidivism.’** All information was collated into an Excel Spreadsheet named ‘*Master,’* where individuals were presented as rows, and variables were presented as columns. Finally, all variables were coded numerically to ensure compatibility with the algorithm (e.g. male = 1, female = 2). Where appropriate, continuous variables (such as age), were grouped into categories to facilitate analysis (e.g. aged 15-20 = 1, aged 21-25 = 2). Any rows that became duplicated as a result, were subsequently condensed. This reduced the final data set to 3445 rows of data within the Master Worksheet.

Analysis – Descriptives

The study utilised descriptive statistics, such as mean, mode, range and standard deviation (SD) to describe the characteristics of offenders in the dataset such as the average age, gender, ethnicity and occupation. Descriptive statistics were also used to identify possible risk factors, and to summarize the nature of the sexual offences committed on the rail network, such as the most common offence type.

Constructing the Algorithm

For the predictive modelling, this study opted to use a ‘Random Forest’ (RF) model. RF are inductive ML algorithms that generate predictions by combining results from numerous classification trees (Berk et al, 2009; Fellows, 2024a). The classification trees continually divide the data into smaller subsets, using different predictor variables at each step, to increase uniformity in relation to the predicted outcome variable (Berk et al, 2009; Fellows, 2024a). The process usually begins by identifying the strongest predictor linked to the outcome variable (typically using the Gini Index) and continues until there is no further splits that could meaningfully enhance the classification (Berk et al, 2009, p.196; Berk, 2012; Fellows, 2024a). Each classification tree can be considered a different expert, giving its own ‘vote’ on the outcome. The RF takes the answers from all the trees and uses the majority ‘vote’ as the final classification. E.g. If most trees have voted ‘high risk’, then the model will classify the individual as ‘high risk’ (as seen in Figure 2). This method is especially good at handling complex data with many different variables and is more accurate than using a single decision tree (Etzler et al, 2024). RF algorithms recognise patterns of intercorrelation and allow the cost of false predictions to be built directly into the model, which is critical when forecasting criminal recidivism (Berk et al, 2009).

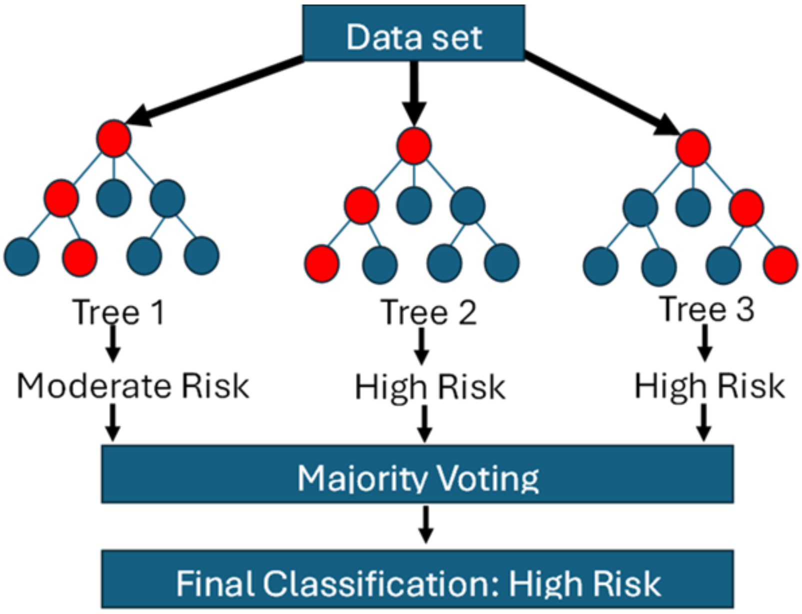

Figure 10. *Illustration of a Random Forest Model (Fellows, 2024b).*

When constructing the algorithm, the time point of prediction (at which the forecast begins), was selected as 1st April 2023. This enabled a 12-month follow up period and ensured consistency across all offenders within the data.

Drawing on the insights from the literature review, the available data and potential practical/policy implications, the following predictor variables were built into the initial RF model:

- Age (5-year intervals).
- Age (10-year intervals).
- Gender.
- Ethnicity.
- Self-Defined Ethnicity.
- Occupation (broad occupation types).
- Occupation (Employed, unemployed, student, retired or unknown).
- Cumulative CCHI score.
- Total number of previous arrests.
- Number of previous arrests for contact sexual offences.
- Number of previous arrests for non-contact sexual offences.
- Number of previous arrests for digitally enabled sexual offences.
- Number of previous arrests for attempted sexual offences.
- Number of previous arrests for other sexual offences.
- Total number of previous arrests for all sexual offences.
- Number of previous arrests for violent offences.
- Number of previous arrests for robbery offences.
- Number of previous arrests for public order offences.
- Number of previous arrests for drugs-related offences.
- Number of offences that occurred in the same city/town as offender's residential address ('City/Town Match').
- Number of offences that did **not** take place in the same city/town as offender's residential address ('City/Town No Match').
- Number of offences that occurred in the same County as offender's residential address ('County Match').
- Number of offences that did **not** take place in the same County as offender's residential address ('County No Match').

For the current study, sexual recidivism was defined as a subsequent arrest for a further sexual offence, following an earlier sexual offence arrest. As such, the primary outcome variable was considered a further arrest for a sexual offence within the 1-year follow up period. The outcome variable was categorised into two classifications; '0' where there was **no further arrest** for a sexual offence, or '1', meaning **a further arrest** for a sexual offence during the follow up period.

Once prepared, the data was uploaded to R-Studio (Version 4.4.1) and the model constructed using the 'randomForest' package. The data was cleaned using the 'janitor' function to ensure consistent formatting (all lower case, with no lingering spaces). The proposed predictor variables were assessed for collinearity, using Variance Inflation scores (VIF) and a correlation matrix. To prevent issues of multicollinearity, such as biased variable importance or overestimation of model performance, any highly correlated features were removed from the model. The data was randomly divided into a training subset (70%, N=2411) and a testing subset (30%, N=1034). Randomly allocating the data into these subsets minimised the risk of selection bias, and enhanced external validity by ensuring the model's performance generalised well to previously unseen data (Ratcliffe, 2022). The 'training' subset enabled the model to learn patterns from the data, whilst testing various combinations of predictors in relation to the outcome variable (Berk, 2012). The 'testing' subset was used to evaluate the model's predictions by comparing these to actual outcomes from data it had not previously encountered. The testing subset remained separate until this stage to enable a thorough evaluation of the model's performance once new cases were introduced (Berk, 2012; Fellows, 2024a). Given the potential consequences of false predictions, the RF model was fine-tuned, using a plot of error rates versus number of trees to determine the optimal number of classification trees and splits (mtry). Oversampling and under-sampling methods were also utilised to minimise the risk of overfitting, whereby the model 'overlearns' and underfitting, where the model does not learn enough from the data.

Analysing the Algorithm

The model's overall accuracy was calculated using the number of true and false predictions. A *true positive* indicates the model correctly forecast a further arrest for a sexual offence, whilst a *true negative* indicates the model correctly predicted no further arrests for a sexual offence (Fellows, 2024a). On the other hand, a *false positive* occurs when the model incorrectly predicted a further arrest for a sexual offence, and a *false negative* indicates that the model incorrectly predicted no further arrest for sexual offences (Fellows, 2024a). The number of true and false predictions were presented in a Confusion Matrix (illustrated in Table 1) and used to calculate the model's specificity and sensitivity (Berk et al, 2009; Fellows, 2024a). A high sensitivity score indicates the model identifies positive cases effectively, whilst high specificity reflects the algorithm's ability to accurately identify negative cases (Fellows, 2024a; Kovalchuk et al, 2023).

Table 1. *Example confusion matrix to demonstrate true and false predictions*.

	Actual outcome – No further arrest for a sexual offence	**Actual outcome – Further arrest for a sexual offence**
Predicted Outcome – No further arrest for a sexual offence	*True Negative*	*False Negative*
Predicted Outcome – Further arrest for a sexual offence	*False Positive*	*True Positive*

To determine the **predictive importance** of each variable, the Gini index was used to analyse any decrease to the model's accuracy when each feature is removed in turn (Berk et al, 2009; Berk, 2012; Berk, 2013; Brieman, 2001). Finally, to assess the reliability of the model, five-fold cross validation was applied, whereby the original dataset was randomly divided into five 'folds'. Each fold was used to create different training and testing subsets and the RF model was assessed using each of these.

Limitations

Poorly recorded variables is a common limitation in forecasting data, and this study was no exception. Missing data was located in several of the *'person data'* columns, most notably self-defined ethnicity. To address this, unknown values were assigned a category, or the affected rows were removed during the missing data analysis (Berk, 2012). The absence of CCHI scores was also a notable limitation, particularly for sexual offences, although this only impacted a small proportion of the sample (3.3% of all arrests , 0.5% for sexual arrests).

There are also limitations in using 'arrest' data, as it is unlikely that all individuals arrested will ultimately be convicted, however this was selected with consideration for its practical implications. Using arrest data also excludes sexual offences where the suspect has been invited for interview or reported for consideration, rather than arrested as per the traditional process. However, this was anticipated to only exclude a small number of cases, as the College of Policing (n.d.) encourages that sexual offenders should be promptly arrested to aid effective investigation.

Furthermore, it is worth noting, some offences included in Force Crime Group 2A may not be inherently sexual, such as Outraging Public Decency. This offence encompasses a variety of behaviours ranging from public urination (not inherently sexual) to masturbation under clothing (clearly sexualised behaviour). Unfortunately, due to recording processes on Niche, there is no way to distinguish these, so rather than exclude a large portion of data, these offences remained within the final data set.

Chapter Summary

This chapter sought to promote transparency, explainability and interpretability in alignment with the frameworks of ethical and trustworthy AI. It described the research design, data and methodology used to address the research questions. The chapter also explained the analytical methods employed to examine the data in relation to the research questions, before addressing the potential limitations.

Chapter 4: Findings

Introduction

This chapter presents the findings for each research question in turn. It employs descriptive statistics to examine the nature and prevalence of sexual offences on the rail network, detailing frequency of offence types and patterns of recidivism. The chapter then analyses the characteristics of individuals arrested for sexual offences and considers risk factors associated with reoffending. For RQ3, these characteristics and risk factors are scrutinised across different sexual offence subgroups. Lastly, the predictive validity of the ML algorithm is assessed, including an analysis of variable importance and measures implemented to prevent over or under-fitting.

Nature of sexual offences, and rates of recidivism

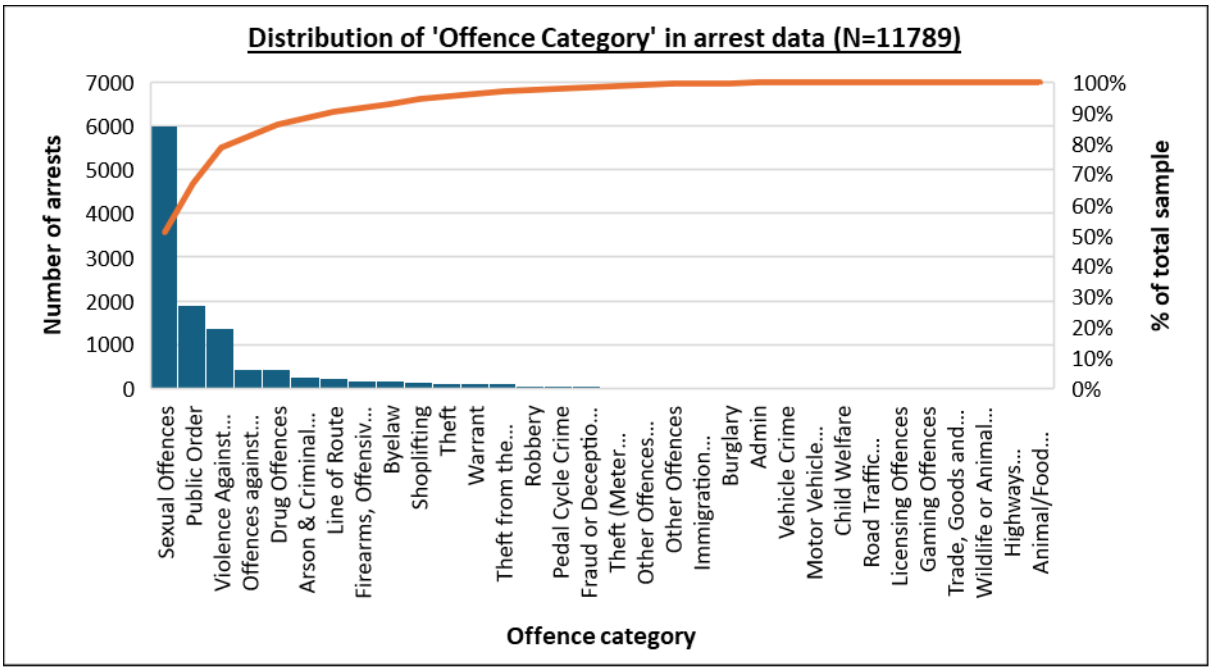

***Figure 11**. Distribution of 'Offence Category' in arrest data.*

To understand the nature and prevalence of sexual offences on the rail network, an initial analysis of the '*arrest data'* was conducted. Figure 3 reveals over 50% of arrests within the dataset related to sexual offences (N=6013, 51%). However, because the report was designed to capture *all arrests* for individuals arrested for at least one previous sexual offence, the final dataset also included 32 other offence categories. Among these, Public Order (N=1901, 16%) and Violence Against the Person (N=1377, 12%) were the next most common offence categories. Of the 32 offence categories included in the dataset, 22 accounted for less than 1% of the overall sample.

Table 2. *The most common sexual offences on the rail network*.

Offence Wording	Number of arrests	% of total sample	Sex Offence Subgroup
Sexual assault on a female	2520	41.9%	Contact
Act of Outraging Public Decency – Common Law	1213	20.2%	Non-contact
Exposure – SOA 2003	638	10.6%	Non-contact
Sexual assault on a male	194	3.2%	Contact
Public Nuisance – Indecent Exposure	144	2.4%	Non-contact

The most common sexual offences on the rail network are outlined in Table 2, comprising of two contact offences and three non-contact offences. These were ‘Sexual Assault on a Female’, ‘Outraging Public Decency’, ‘Exposure’, ‘Sexual Assault on a Male’ and ‘Indecent Exposure.’ Collectively, these offences contributed to almost 80% of all arrests for sexual offences within the dataset. Figure 4 illustrates the proportion of arrests across the different sexual offence subgroups. This analysis reveals more than half of the sexual arrests were for ‘contact’ offences, followed by ‘non-contact’, which accounted for the second largest proportion. ‘Attempted’, ‘digitally enabled’ and ‘other’ offences all contributed to less than 10% each of the total number of sexual arrests.

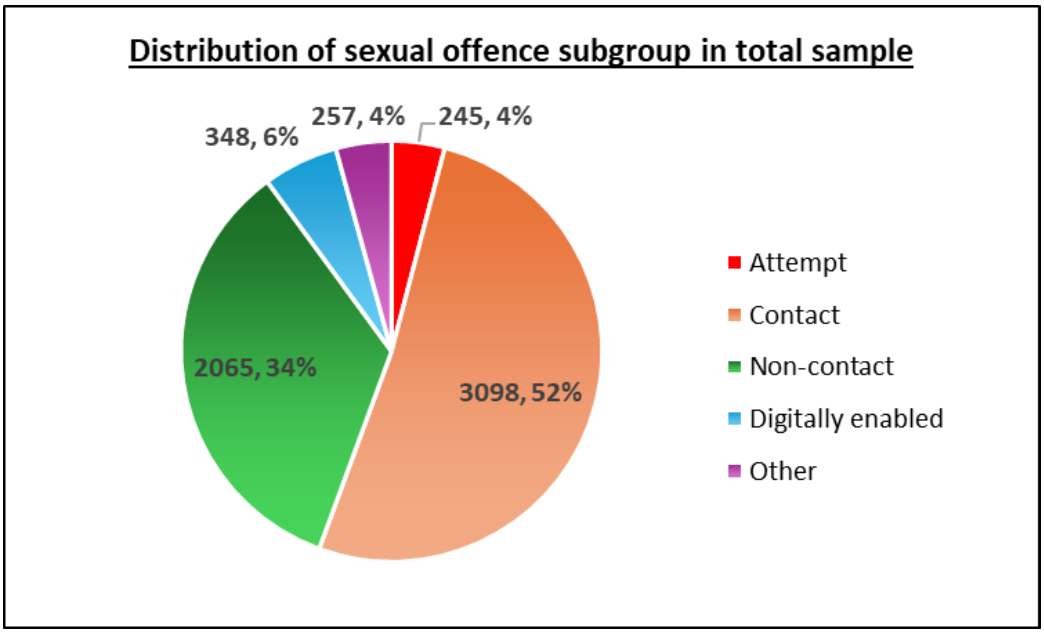

Figure 4. *Distribution of sexual offence subgroups in total sample.*

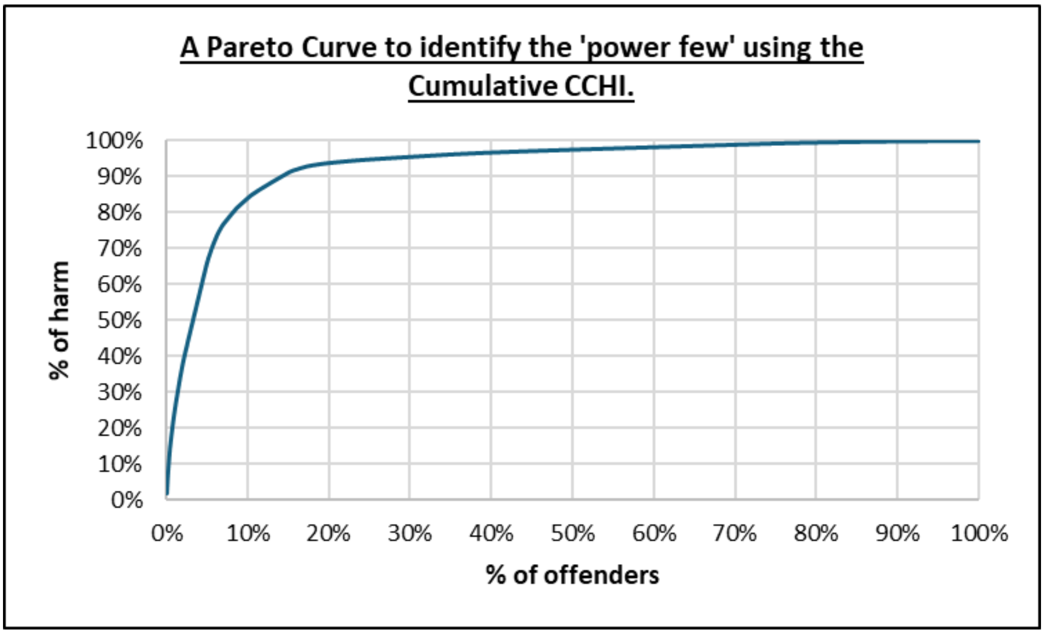

Figure 5. *Pareto Curve to identify the 'power few' using the Cambridge Crime Harm Score.*

The total number of arrests per individual ranged from 1 to 73 (M = 3.3, SD = 4.3) with 70 individuals (2%) identified as outliers based on z-scores. For sexual offences, the number of arrests per individual ranged from 1 to 21 (M = 1.7, SD = 1.6), with a similar number of outliers (76). The total harm score across all offences amounted to 859,156 and the Cumulative CCHI score per individual ranged from 0 to 14655 (M = 272.7, SD = 923.1). The harm scores for sexual offences totalled

694,762 and ranged from 0 to 14643 per individual (M = 225.0, SD = 830.4). Figure 5 presents a Pareto Curve illustrating the percentage of total harm, caused by the percentage of offenders. It reveals over 80% of the total harm was attributed to less than 10% of offenders, and 20% of offenders were responsible for over 90% of the total harm.

Almost two-thirds of sexual arrests occurred on B Division (N=3925, 65.3%), whilst C Division accounted for just over one-third (N=2088, 35%). Most sexual offences occurred whilst on the train (N=3140, 52.2%), followed by in the station (N=1566, 26%) and on the platform (N=903, 15%). Railway stations were the most common offence location (N=4401, 73%), followed by London Underground stations (N=1298, 22%), collectively accounting for 95% of incidents. The number of sexual offences on the rail network remained relatively stable throughout the week, with the lowest rates on Sundays (N=760, 12.6%) and Mondays (N=753, 12.5%) and peaks on Fridays (N=922, 15.3%) and Saturdays (N=1070, 17.8%). Of the sexual offences, 29% (N=1777) occurred in the same City or Town as the offender's residential address.

Regarding recidivism, 125 individuals were arrested for a subsequent sexual offence during the 12-month follow up period, indicating a *sexual* recidivism rate of 3.6%. However, when the 12-month time frame is removed, this number increased to 360 individuals, resulting in a *sexual* recidivism rate of 10.4%. The average time between a first sexual offence, and a subsequent arrest for a further sexual offence was calculated as 545 days. The rate of *general* recidivism (a subsequent arrest for **any** further offence), was 30% (N=1031).

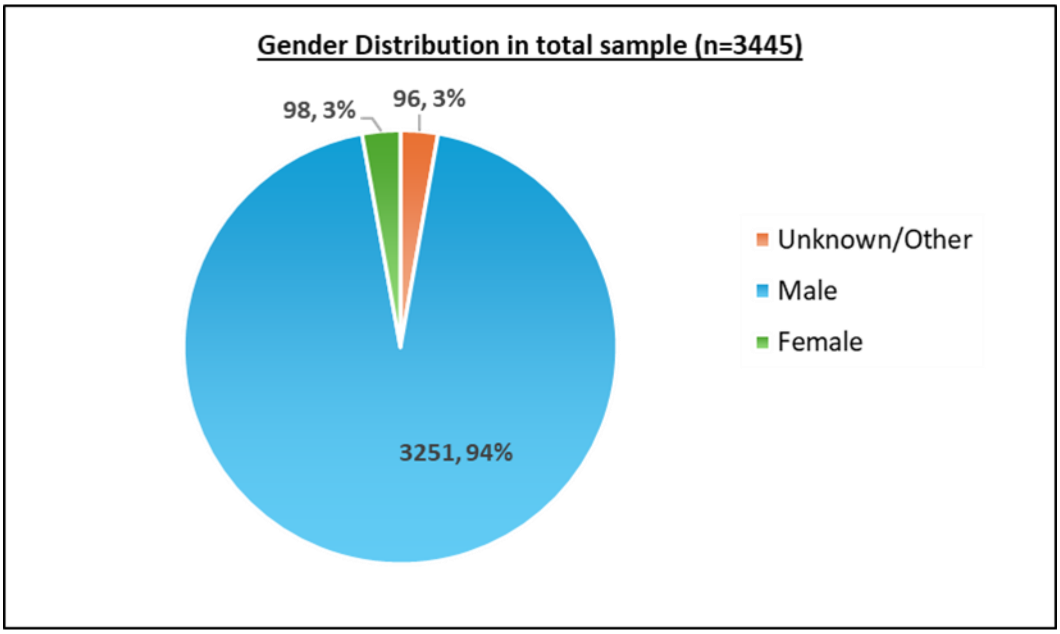

Figure 6. *Gender distribution in total sample.*

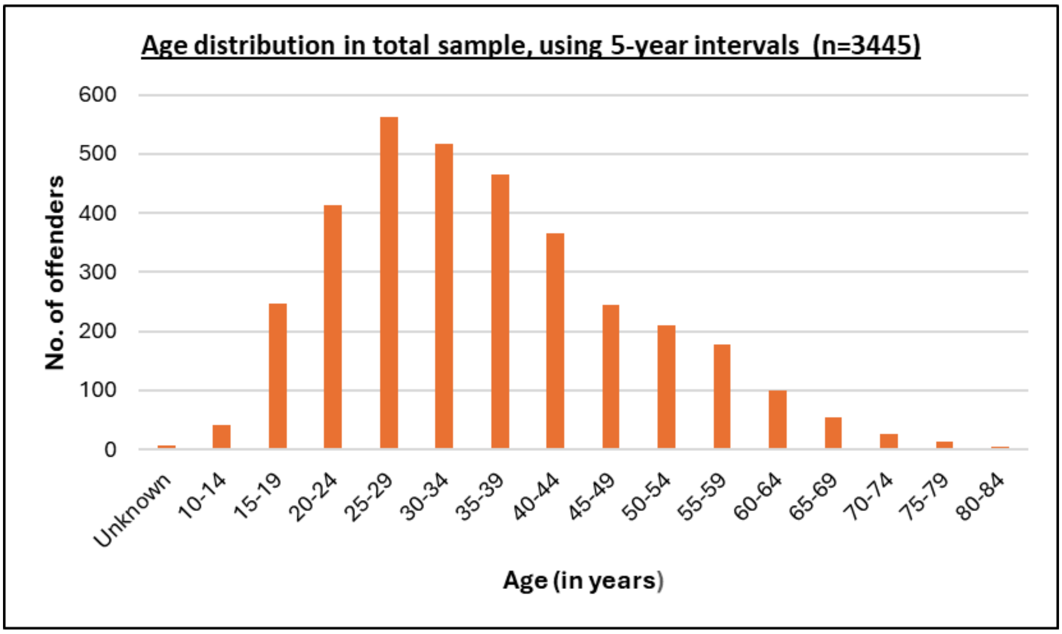

Figure 7. *Age distribution in total sample, using 5-year intervals.*

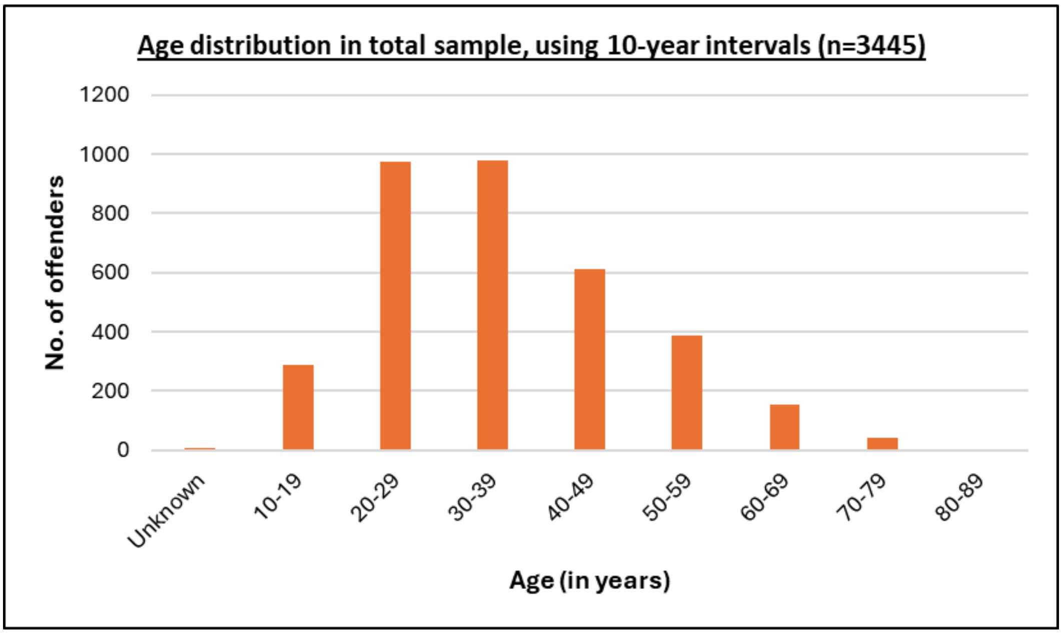

Figure 12. *Age distribution in total sample, using 10-year intervals.*

Figure 6 revealed most offenders within the sample were male (94%), whilst females and individuals of unknown or other genders accounted for only 3% each. Ages within the sample ranged from 12 to 84 years old (M = 35, SD = 12.6) and the most common age group was 25 to 29 years old (*when using 5-year intervals*). The number of offenders per age group increased with age, peaking at 25 to 29 years, after which, it steadily declined. Notably, when ages are grouped using 10-year intervals, the 20 to 29 and 30 to 39 age groups account for nearly equal proportions of the total sample, as shown in Figure 8.

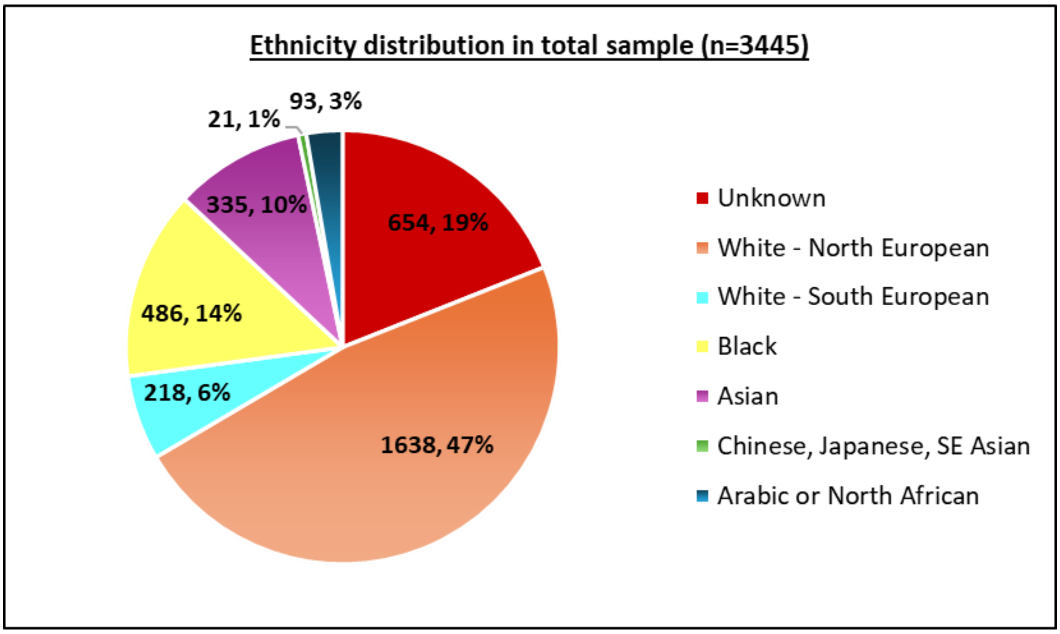

Figure 9. *Ethnicity distribution in total sample.*

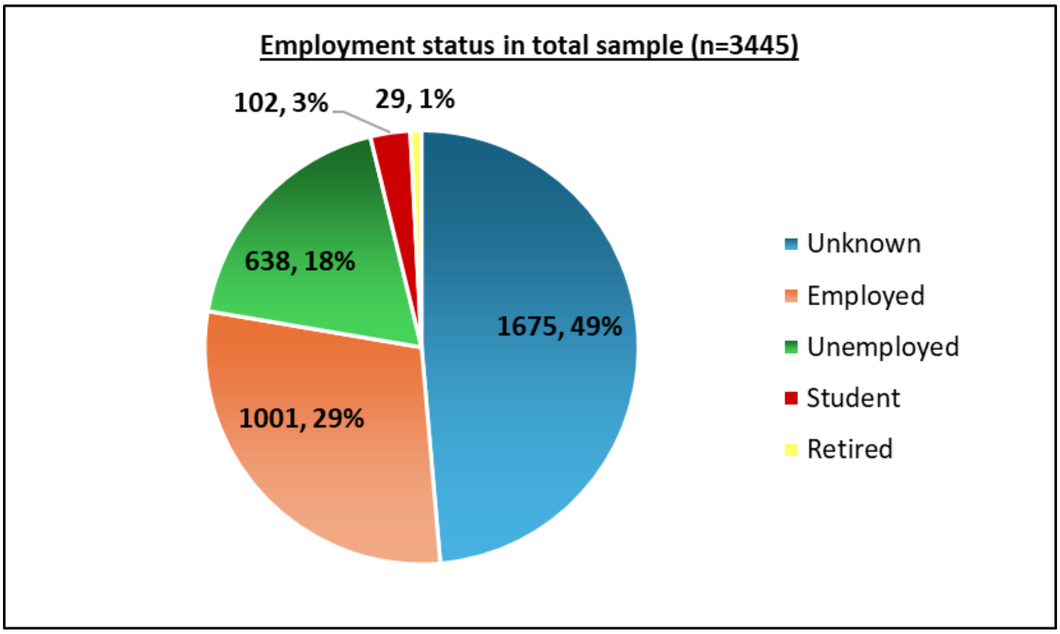

Figure 10*. Employment status in total sample.*

Regarding ethnicity, just under half the sample were White – North European (47%), followed by Black (14%), Asian (10%), White – South European (6%), Arabic and North African (3%) and Chinese, Japanese and SE Asian (1%). Unfortunately, 19% did not have any recorded ethnicity on Niche. Figure 10 reveals 49% of the sample had no employment information recorded at the time of arrest. Among the remaining half, 29% were employed, 18% unemployed, 3% students and 1% were retired. Of the total sample, only 4% were recorded as homeless at the time of arrest.

Risk factors for recidivism.

To identify potential risk factors for recidivism, data was compared between individuals who were not arrested a further sexual offence during the 12-month follow up (*non-recidivists*) and those who were (*recidivists*). Consistent with the overall sample, *males* represented the highest percentage of offenders amongst both non-recidivists (94.3%) and recidivists (96.8%). Notably, no individuals of unknown or other genders reoffended during the follow up period. The average age for offenders was 35 years for both samples; however, the age range was slightly narrower for recidivists (14-70) compared to non-recidivists (12-84). When analysed in 5-year intervals, the most common age range for both groups was 25 to 29 years old. However, in 10-year intervals, recidivists were most commonly aged 20 to 29 years, whilst non-recidivists were older (most commonly aged 30 to 39 years). In terms of ethnicity, White – North Europeans constituted the largest proportion in both samples. However, individuals with unknown ethnicity and those recorded as Black, were more prevalent in the recidivist sample. Interestingly, unemployment was more than double amongst those who were rearrested (40%), compared to those who were not (17.7%). The mean cumulative CCHI score was higher for the non-recidivist sample (M= 308.3, SD = 1034.9), compared to the recidivist sample (M=129.3, SD=373.3) however the large SD indicates significant variation in the data.

Given the small sample size of 12-month recidivists, a further analysis was conducted to compare those arrested for an additional sexual offence at any time during the 8-year period. This largely reflected the same patterns within the 12-month recidivism sample, however, the mean age of offenders was slightly older at 37 years old and the mean cumulative CCHI was substantially higher.

Table 3. *Characteristics of recidivists compared to non-recidivists.*

	Non-recidivists (N=3320)	Sexual Recidivists (N=125) *(12 month fixed follow up period)*	Sexual Recidivists (N=360) *(8-year follow up period)*
Gender			
Male	94.3% (3130)	96.8% (121)	98.3% (354)
Female	2.8% (94)	3.2% (4)	0.8% (3)
Unknown/Other	2.9% (96)	0% (0)	0.8% (3)
Age			
Mean	35.9 *(SD 13.1)*	35.5 *(SD 12.6)*	37.0 *(SD 13.7)*
Range	12-84	14-70	13-78
Mode (5-year)	25-29 years	25-29 years	25-29 years
Mode (10-year)	30-39 years	20-29 years	20-29 years
Ethnicity			
White – North European	47.5% (1578)	48% (60)	41.1% (148)
Unknown	18.9% (627)	21.6% (27)	21.4% (77)
Black	13.8% (458)	22.4% (28)	20.6% (74)
Asian	9.9% (330)	4% (5)	9.2% (33)
White – South European	6.5% (216)	1.6% (2)	3.9% (14)
Arabic or North African	2.7% (90)	2.4% (3)	3.1% (11)
Chinese, Japanese, SE Asian	0.6% (21)	0% (0)	0.8% (3)
Employment Status			
Employed	29.3% (974)	21.6% (27)	30.3% (109)
Unemployed	17.7% (588)	40% (50)	35.0% (126)
Student	2.9% (96)	4.8% (6)	5.6% (20)
Retired	0.8% (28)	0.8% (1)	1.1% (4)
Unknown	49.2% (1634)	32.8% (41)	28.1% (101)
Mean Cumulative CCHI Score	308.3 (SD = 1034.9)	129.3 (SD = 327.3)	364.2 (SD = 974.9)

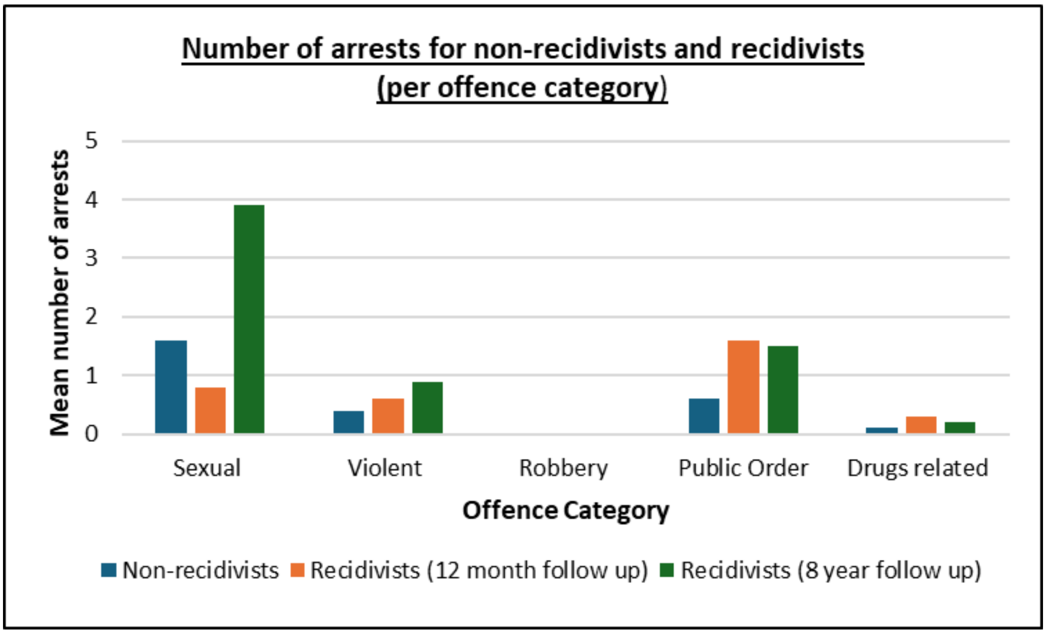

Figure 14. *Number of arrests for non-recidivists and recidivists (per offence category).*

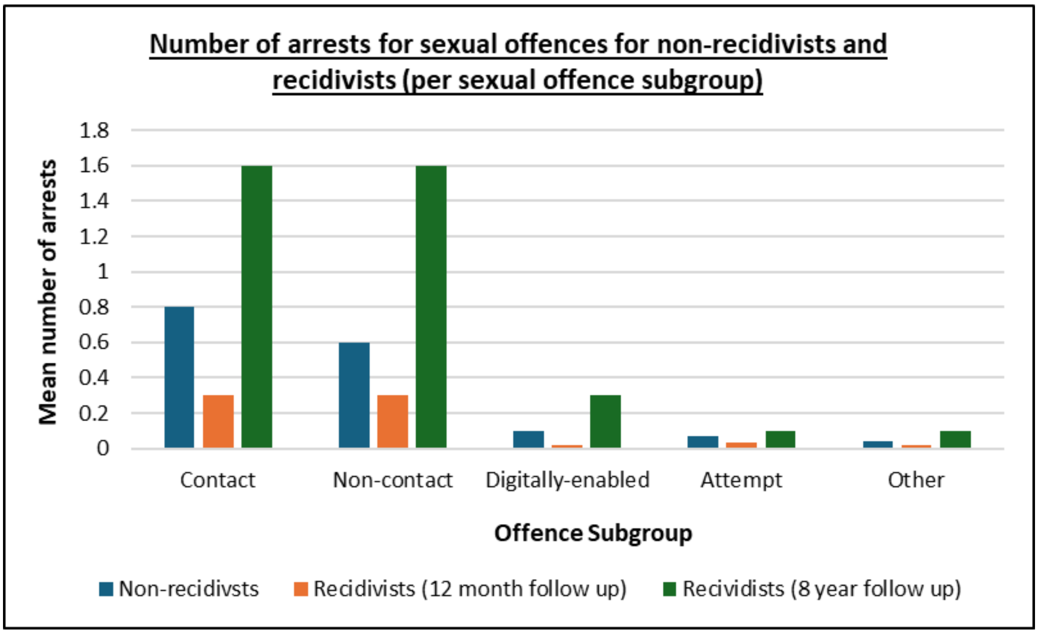

Figure 13. *Number of arrests for sexual offences for non-recidivists and recidivists (per sexual offence subgroup).*

These samples were then analysed regarding the total number of arrests per offence category. Overall, both recidivist samples had a higher number of total arrests (*12 month* - M=5.3, SD=6.1; *8 year* - M=8.5, SD=8.9) compared to the non-recidivist sample (M=3.4, SD=4.5), though all groups showed large SD. Recidivists in the 8-year time frame had the highest number of arrests for sexual offences (M=3.9, SD = 3.0), followed by non-recidivists (M=1.6, SD=1.6) and the 12-month recidivist sample (M=0.8, SD=1.3). For violent, public order and drugs-related offences, both recidivist samples had a greater number of arrests than non-recidivists. Regarding location, offences occurring outside of the offender's residential city, or county were more common than local offences, and this was consistent across all samples. Figure 12 further illustrates the number of arrests for the different subgroups of sexual offences for recidivists and non-recidivists. These findings reveal the 8-year recidivist sample had the highest number of previous arrests across all subgroups of sexual offences, followed by non-recidivists, and the 12-month recidivist sample.

Table 4. *Proportion of offenders arrested across multiple sexual offence subgroups.*

	Total Sample (N= 3445)	Non-recidivists (N = 3320)	Recidivists (12 month follow up) (N=125)	Recidivists (8-year follow up) (N=360)
One Subgroup	92.1% (3173)	92% (3055)	94.4% (118)	78.9% (284)
Two Subgroups	7.3% (251)	7.3% (244)	5.6% (7)	19.7% (71)
Three Subgroups	0.5% (18)	0.5% (18)	0% (0)	1.4% (5)
Four Subgroups	0.1% (3)	0.1% (3)	0% (0)	0% (0)

Table 4 presents the distribution of individuals arrested for offences in multiple different sexual offence subgroups. Within the total sample, most offenders were arrested for offences within only one subgroup (92%) – e.g. only contact offences. This pattern was consistent across non-recidivist and both recidivist groups. However, a larger proportion of offenders were arrested for offences across two and three subgroups within the 8-year recidivism sample.

Differences between sexual offence subgroups

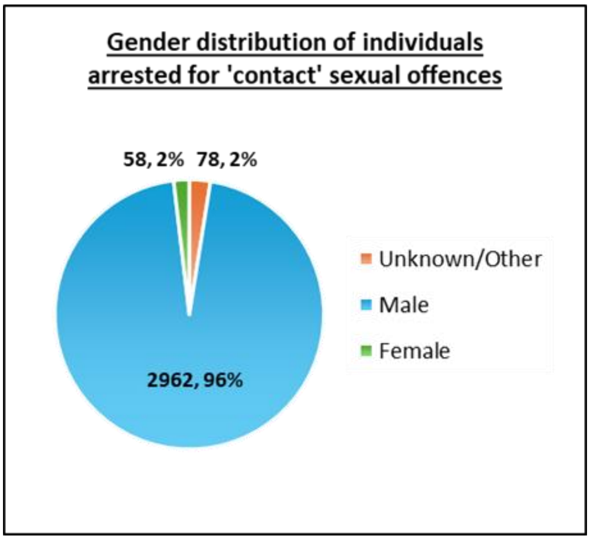

Figure 13. *Gender distribution of individuals arrested for a 'contact' sexual offence.*

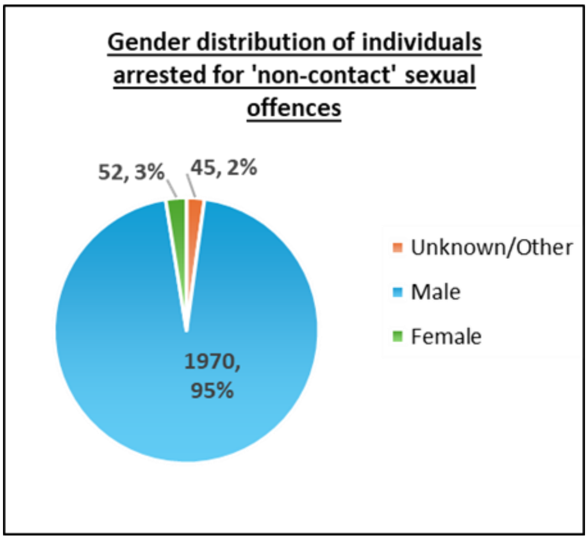

Figure 14. *Gender distribution of individuals arrested for 'non-contact' sexual offences.*

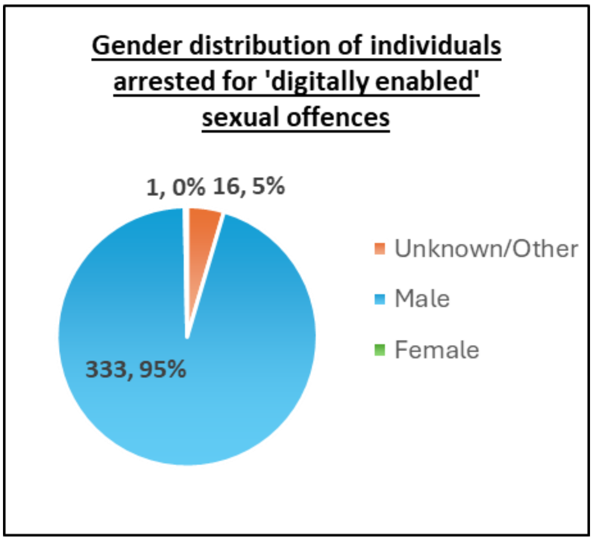

Figure 15. *Gender distribution of individuals arrested for 'digitally enabled' sexual offences.*

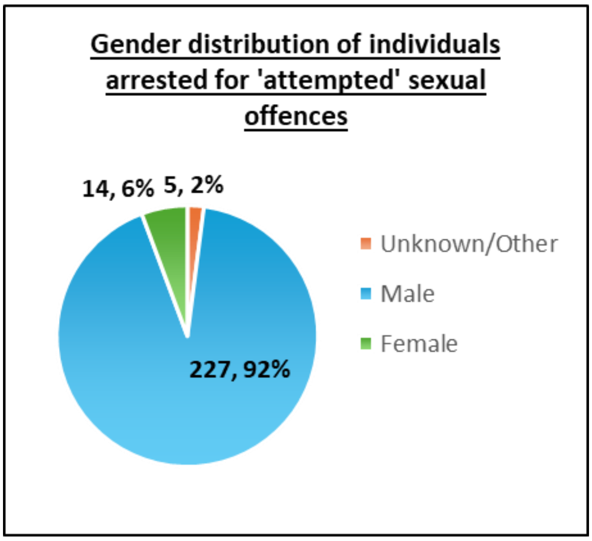

Figure 16. *Gender distribution of individuals arrested for 'attempted' sexual offences.*

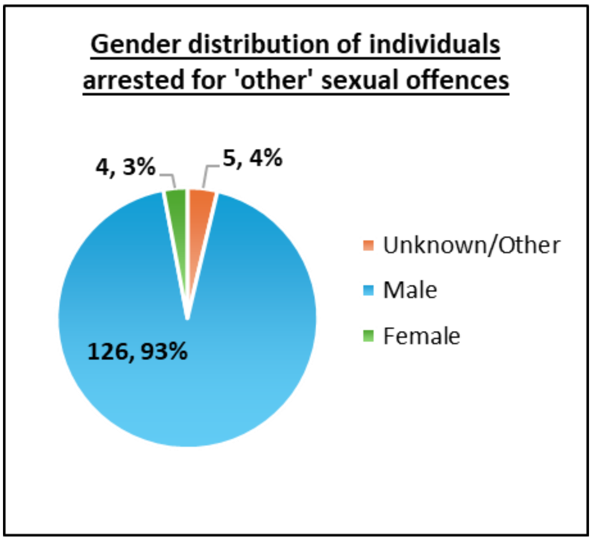

Figure 17. *Gender distribution of individuals arrested for 'other' sexual offences.*

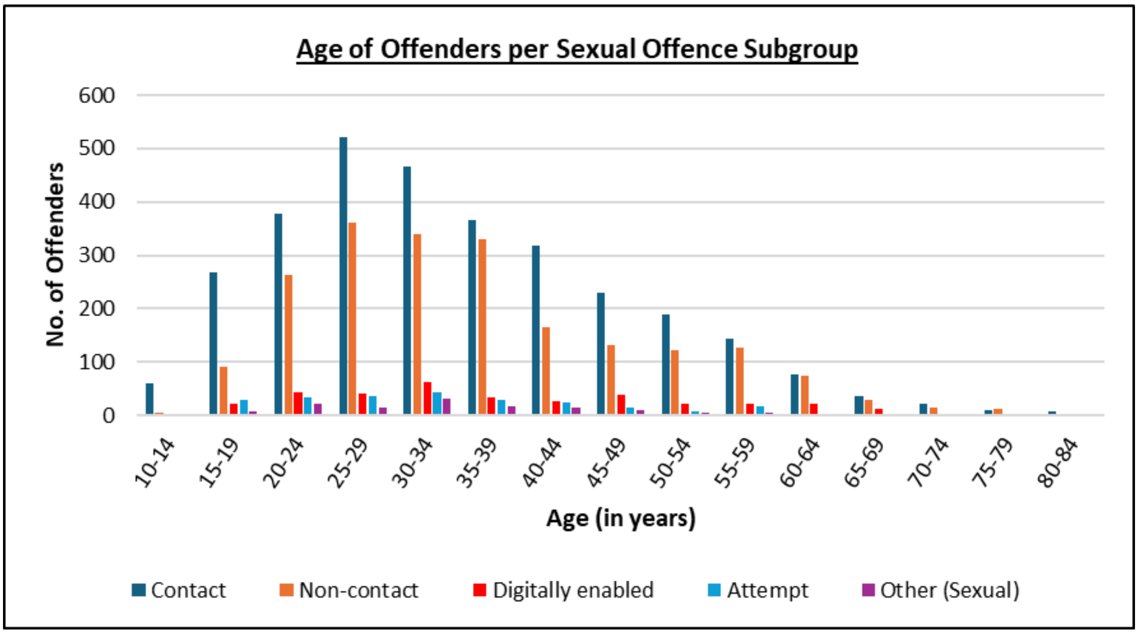

Figure 15. *Distribution of age, per sexual offence subgroup.*

To address RQ3, the analysis compared the characteristics of individuals within each distinct sexual offence subgroup. Figures 13 to 17 reveal males accounted for over 90% of offenders in all samples. Notably, female offenders were absent in cases of 'digitally enabled' sexual offences, however they constituted 6% of 'attempted' sexual offences, slightly higher than other subgroups. Regarding age, 'contact' offenders generally mirrored the overall sample; however, 'non-contact' offenders did not steadily decrease with age as expected, instead remaining relatively stable between 25 to 39 and 40 to 59 years old. Additionally, 'digitally enabled' offenders seemingly peaked slightly later (Mean age = 37.9, SD= 14.4) than the total sample. Unfortunately, meaningful patterns within the 'attempted' and 'other' sexual offences, were unclear due to small sample sizes.

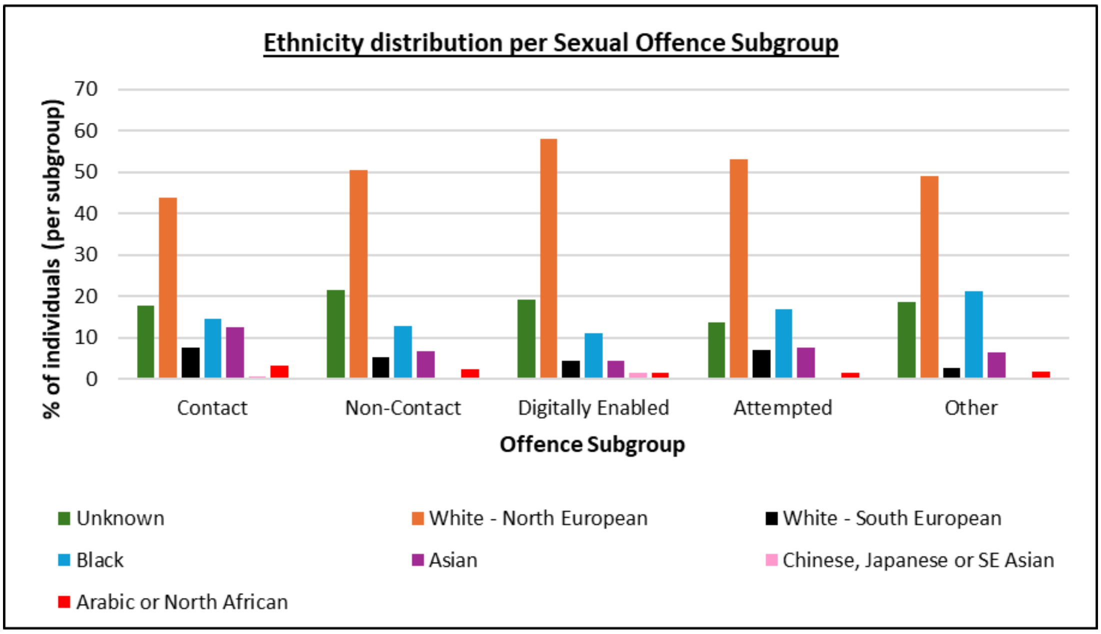

Figure 19. *Ethnicity distribution per sexual offence subgroup.*

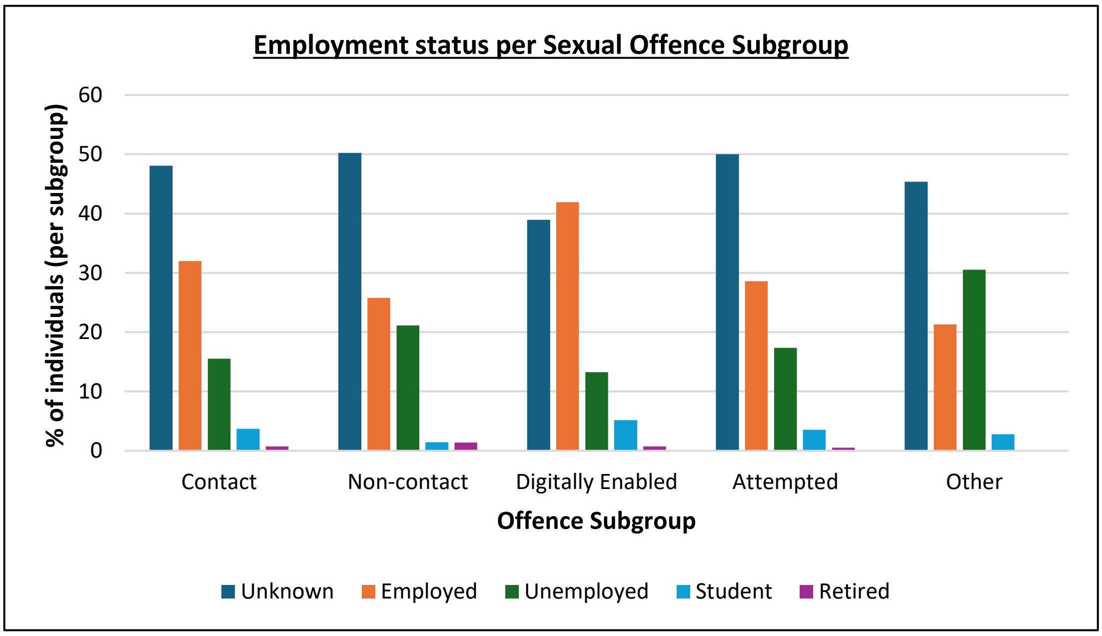

Figure 20. *Employment status per sexual offence subgroup.*

Figure 19 shows White – North European accounted for the largest proportion of offenders across all subgroups, contributing to over 50% for non-contact, digitally enabled and attempted sexual offences. Interestingly, Figure 20 shows individuals who had been arrested for digitally enabled offences had the highest employment rates compared to other subgroups. The cumulative CCHI score was not analysed at this stage, as any differences could be due to the offence subgroups themselves, e.g. contact offences likely carry a higher CCHI score than non-contact offences.

Table 5. *Number of previous arrests, per sexual offence subgroup*

		Offence Subgroup				
		Contact	Non-contact	Digitally Enabled	Attempt	Other
Previous arrests	Total	M = 3.1 SD = 5.4	M = 4.2 SD = 5.1	M = 5.2 SD = 5.3	M = 5.2 SD = 8.0	M = 6.1 SD = 7.9
	Sexual	M = 1.7 SD = 1.6	M = 2.0 SD = 2.0	M = 3.3 SD = 3.4	M = 2.2 SD = 2.2	M = 3.1 SD = 3.4
	Violent	M = 0.4 SD = 1.1	M = 0.4 SD = 1.2	M = 0.3 SD = 0.9	M = 0.6 SD = 1.6	M = 0.8 SD = 1.8
	Robbery	M = 0.0 SD = 0.3	M = 0.0 SD = 0.1	M = 0.0 SD = 0.2	M = 0.0 SD = 0.1	M = 0.0 SD = 0.0
	Public Order	M = 0.5 SD = 1.5	M = 0.9 SD = 2.1	M = 0.5 SD = 1.5	M = 0.7 SD = 1.4	M = 0.7 SD = 1.5
	Drugs	M = 0.1 SD = 0.5	M = 0.1 SD = 0.5	M = 0.3 SD = 1.2	M = 0.1 SD = 0.4	M = 0.1 SD = 0.5

The total number of arrests was analysed for each sexual offence subgroup, as presented in Table 5. The analysis revealed that individuals arrested for 'other' sexual offences had the highest number of **total** arrests (M = 6.1, SD = 7.9), followed by those arrested for 'attempted' offences (M = 5.2, SD =8.0) and 'digitally enabled' offences (M = 5.2, SD = 5.3). When scrutinized by offence type, individuals arrested for 'digitally enabled' offences had the highest number of total **sexual** arrests (M = 3.3, SD = 3.4), followed by those arrested for 'other' offences (M = 3.1, SD = 3.4), and 'attempted' offences (M = 2.2, SD = 2.2). 'Other' offences also recorded the highest number of violent arrests (M = 0.8, SD = 1.8). Regarding Public Order offences, the individuals who had committed 'non-contact' l offences had the highest number of arrests (M = 0.9, SD = 2.1). Notably, the mean number of robbery-related arrests was 0 across all subgroups of sexual offences. Arrests related to drug-related offences were also generally low, with a slightly higher average observed among those arrested for 'digitally enabled' offences (M = 0.3, SD = 1.2). Across all subgroups, offences outside of the offender's residential address were more common than local offences.

Table 6. *Correlation Matrix to show the risk of recidivism between subgroup of sexual offences.*

	n	1	2	3	4	5
1. Contact Offences	47	.32*	.05	-.03	0	-.03
2. Non-Contact Offences	45	-.06	.25	.07	0	0
3. Digitally Enabled Offences	3	-.03	-.02	.57	0	-.01
4. Attempted Sexual Offences	5	.25	-.03	-.01	0	-.01
5. Other Sexual Offences	1	.11	-.01	-.01	0	1

Note: **p*<0.05, two-tailed.

Table 6 shows a correlation matrix of the relationship between arrests pre-2023, and post 2023 for each of the different subgroups of sexual offences. Arrests for contact, non-contact and digitally enabled offences all showed small to moderate positive correlations with a further arrest for the same subgroup offence within the follow up period. Though, only the correlation between contact sexual offences was found to be statistically significant (p=0.03). Additionally, previous attempted sexual offences showed a small positive correlation with a further arrest for a contact sexual offence. Whilst 'other' sexual offences appeared to show a perfect correlation, this may be attributed to the small sample size. Interestingly, previous arrests for many of the subgroups were negatively correlated with a further arrest for an offence within a different subgroup.

ML Algorithm

Lastly, to address RQ4, Table 7 presents the five variables most highly correlated with each other, as identified whilst constructing the RF model. To prevent issues arising from collinearity, the following variables were excluded from the algorithm: Age (10-year intervals), County No Match, City/Town No Match and total number of previous arrests. Table 8 then reveals the five variables with the highest VIF scores after removing these highly correlated variables. Since 'previous sexual arrests' had the highest VIF score, this was also removed from the model to address multicollinearity concerns. The VIF scores were rechecked, and all deemed under an acceptable level (<2). For completeness, the correlation matrix was rechecked after removing high correlations. The only remaining variables that appeared highly correlated with one another was the two occupation variables (0.64) (as these had been separated into broader and specific categories). Therefore, the decision was made to remove the specific categories variable, leaving only the broad categories in the data set.

Table 7. *The five most highly correlated variables in the RF model.*

Variable1	Variable2	Correlation
Age (10-year intervals)	Age (5-year intervals)	0.98
County No Match	Total Previous Arrests	0.93
County No Match	City/Town No Match	0.86
City/Town No Match	Total Previous Arrests	0.85
Public Order Offences	Total Previous Arrests	0.70

Table 8. *VIF scores – a measure of multicollinearity.*

Variable	VIF Score
Previous Sexual Arrests	38.27
Previous Non-Contact Sexual Offences	21.61
Contact Sexual Offences	19.43
Digitally Enabled Sexual Offences	6.59
Attempted Sexual Offences	2.66

Table 9 shows the accuracy, sensitivity, and specificity scores for the training and testing data across various iterations of the RF model developed during the parameter fine-tuning process. RF1 was the original model, created using **all features** to obtain a base level of performance. The split of the outcome variable in the training data for RF1 was 2322 (0) and 89 (1). RF2 to R4 presents the scores when the collinear and aliased features were removed, whilst still using the original sample.

Due to concerns regarding the imbalanced nature of the outcome variable, this was artificially inflated using the 'ROSE' function in R Studio. When the minority class in the training data was oversampled to 20%, this increased the number of positive cases to 585 (1), whilst negative cases remained at 2322 (0) *(as seen in RF5 to RF7).* When oversampled to 10%, the positive cases increased to 280 (1), and negative remained at 2322 (0) *(as seen in RF8).* For completeness, the model was also tested with under-sampled training data. When under-sampled to 20%, the number of positive cases was 89 (1), and the number of negative cases was reduced to 357 (0) *(as seen in RF9 to RF11).* When under-sampled to 10%, the number of positive cases remained at 89 (1), and the negative cases was 790 (0) *(as seen in RF12).* All models achieved consistently high accuracy scores and sensitivity scores in both the training and testing data; however, the specificity score varied substantially, ranging from 0.08 to 0.53. Specificity also showed a significant decline between the training and testing data, indicative of overfitting. All models using either the oversampled or under-sampled data, achieved higher specificity scores than those using the original unbalanced data.

Table 9. *Accuracy, Sensitivity and Specificity scores for all RF models (when evaluated against testing data).*

Model	Features	No. of trees	No. of splits (mtry)	Sample	Accuracy		Sensitivity		Specificity	
					Train	*Test*	*Train*	*Test*	*Train*	*Test*
RF1	All	500	5	Original	1.00	0.96	1.00	0.99	0.91	0.19
RF2	Collinear features removed	500	5	Original	1.00	0.96	1.00	0.99	0.89	0.08
RF3	Collinear features removed	100	5	Original	0.99	0.96	1.00	0.99	0.85	0.08
RF4	Collinear features removed	100	4	Original	0.99	0.96	1.00	0.99	0.80	0.06
RF5	Collinear features removed	500	5	Oversampled (20%)	1.00	0.96	1.00	0.98	0.99	0.28
RF6	Collinear features removed	200	5	Oversampled (20%)	1.00	0.96	1.00	0.99	0.99	0.31
RF7	Collinear features removed	200	4	Oversampled (20%)	0.99	0.96	1.00	0.99	0.96	0.31
RF8	Collinear features removed	500	5	Oversampled (10%)	0.99	0.96	1.00	0.99	0.95	0.22
RF9	Collinear features removed	500	5	Under-sampled (20%)	0.99	0.94	1.00	0.96	0.96	0.53
RF10	Collinear features removed	300	5	Under-sampled (20%)	0.99	0.94	1.00	0.96	0.97	0.53
RF11	Collinear features removed	300	4	Under-sampled (20%)	0.99	0.95	1.00	0.96	0.96	0.47
RF12	Collinear features removed	500	5	Under-sampled (10%)	0.99	0.96	1.00	0.98	0.94	0.42

RF10 was selected for further evaluation as it produced the highest specificity score without compromising the sensitivity or overall accuracy. To validate the model's performance, a five-fold cross validation was conducted. Table 10 shows the overall accuracy of the RF10 model when tested against each of the five folds. Furthermore, Table 11 presents a confusion matrix of the average number of true and false predictions made against the actual outcomes using RF10. Using these predictions, the overall specificity of the model was calculated as 0.97 and the sensitivity was 0.56. The overall accuracy was 0.96.

Table 10. *RF10 Accuracy scores achieved using five-fold cross validation.*

	Accuracy Score
Fold 1	0.94
Fold 2	0.94
Fold 3	0.95
Fold 4	0.96
Fold 5	0.96

Table 11. *Confusion Matrix – the number of true and false predictions using five-fold cross validation with RF10.*

	Actual outcome – No further arrest for sexual offence	**Actual outcome – Further arrest for sexual offence**
Predicted outcome – No further arrest for sexual offence	*True negative* = 645	*False negative* = 12
Predicted outcome – Further arrest for sexual offence	*False positive* = 17	*True positive* = 15

Finally, the Gini Index was analysed to understand variable importance. Figure 21 shows the mean decrease in accuracy when each variable is removed from the model in turn. It reveals the top three most important variables in the current model were previous arrests for contact offences, non-contact offences and CCHI score. The least important variables were gender, previous drugs-related arrests, and ethnicity.

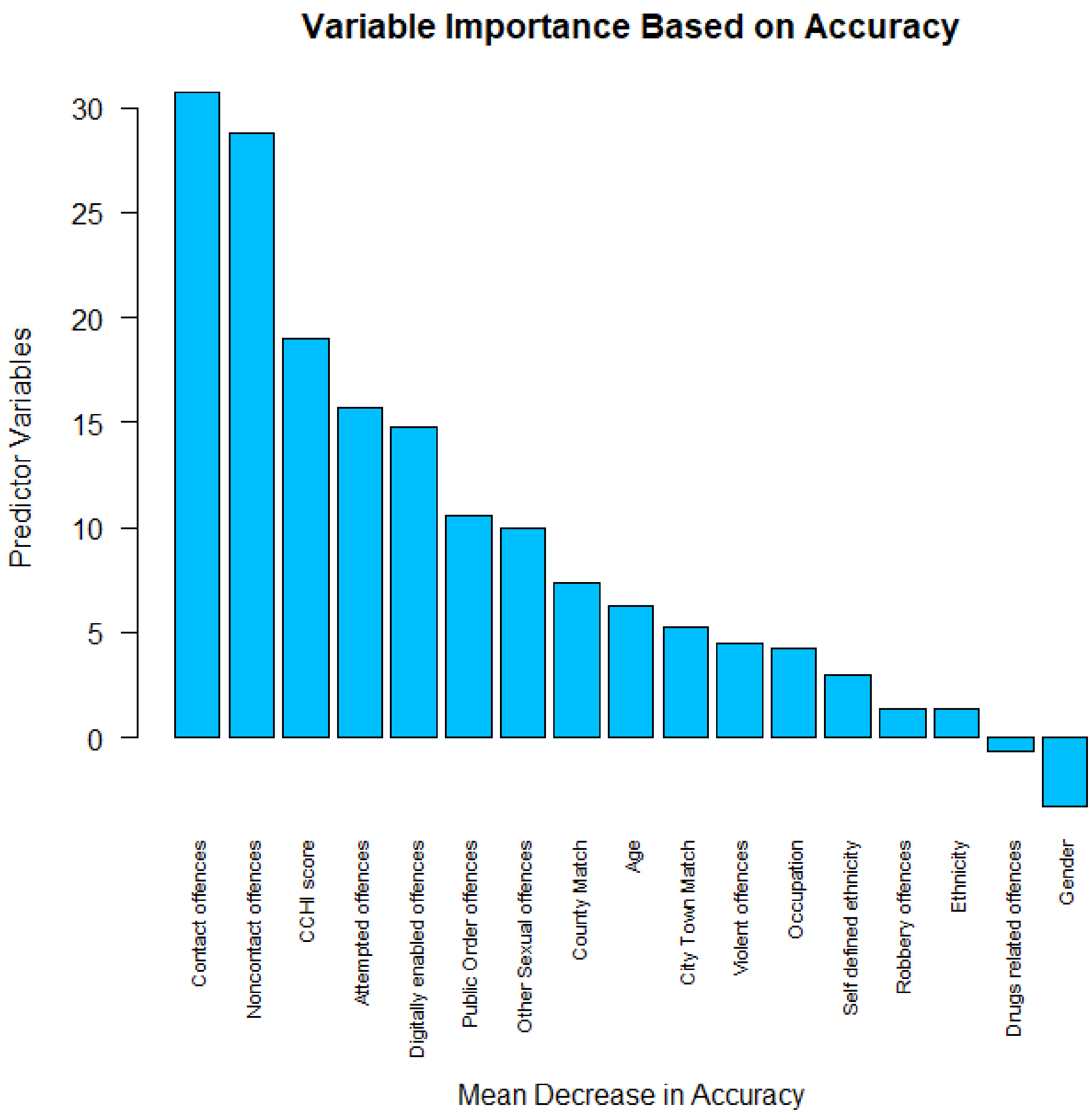

Figure 16. *Gini Index – Variable Importance (Based on Accuracy)*

Chapter Summary

This chapter presented findings for the research questions based on analyses using descriptive statistics and predictive modelling. Regarding RQ1, the findings reveal that the most common type of sexual offence on the rail network is ‘*sexual assault on a female’*, and the rate of sexual recidivism is 3.6% (within a 12-month follow up period), rising to 10.4% over an 8-year period. For RQ2, the results indicate most sexual offenders are young, white males, and this was consistent across both recidivists and non-recidivists. Overall, recidivists had a greater number of total arrests, including public order, violence and drugs-related arrests. Most offenders were arrested for sexual offences within a single subgroup; however the 8-year recidivist sample showed a higher proportion of arrests for offences spanning across two and three different sexual offence subgroups. Regarding RQ3, slight differences were observed in the characteristics and risk factors of offenders across the various sexual offence subgroups. The analysis revealed a positive correlation between previous arrests for one offence type, and a further arrest for the same offence type, however only the correlation for contact offences was statistically significant. Lastly, in response to RQ4, the RF model achieved an overall accuracy of 0.96, a sensitivity score of 0.56 and specificity of 0.97. The most important variables to the model were previous arrests for contact offences and non-contact offences along with the CCHI score. The variables that had the least impact on the model’s accuracy were gender, previous drugs-related arrests and ethnicity.

Chapter 5: Discussion

Introduction

This research set out to understand the nature of sexual offences on the UK rail network and examine the prevalence of sexual recidivism. It sought to identify the characteristics of sexual offenders, risk factors associated with reoffending and evaluate the predictive validity of ML algorithms in forecasting recidivism. This chapter begins by contextualising the findings within the existing body of research. It then discusses potential practical and policy implications, highlighting benefits of further research before addressing limitations of the current study.

Nature of Sexual Offences

The findings reveal the three most common sexual offences on the rail network were '*sexual assault on a female', 'outraging public decency' and 'exposure'*. This aligns with previous research which identified '*unwanted touching and groping'* and '*indecent exposure'* as some of the most common forms of USB on the UK rail network (Ariel et al, 2024b, p.16). Ariel et al (2024b) also highlights sexualised jokes, comments, gestures and mimes as common forms of USB. In the current study, these behaviours may have been recorded under Outraging Public Decency; however, some were likely categorised as 'Public Order', and therefore not classified as sexual. Notably, Public Order was the second most common offence category in this study, which may be attributed to these behaviours; however, without detailed information regarding these offences, it is difficult to draw definitive conclusions. This emphasizes the benefits of the new Sex-based Harassment legislation which will help police to identify these behaviours and improve the understanding of USB in public spaces. Interestingly, the current study and the VAWG STRA (VKPP, 2024) report similar rates of exposure (8% and 10% respectively), however Ariel et al (2024b) reported a rate twice as high (20.5%). One explanation for this discrepancy is that the VAWG STRA and current study both utilised *police* data, whereas Ariel et al (2024b)'s findings were based on *victim survey* data. This suggests not all instances of exposure are reported to police, highlighting well-documented issues of underreporting (Garg, 2023). Furthermore, the current study reported ‘contact’ offences accounted for the largest proportion of sexual arrests, followed by ‘non-contact’ which contributed to over a third of offences. This is particularly note-worthy, as non-contact offences are associated with lower crime harm scores and subsequently may not be prioritised (Sherman, 2007). This underscores the importance of combining crime harm scores with crime counts to provide a more accurate representation of the overall impact of these offences (Sherman, Neyroud and Neyroud, 2016).

Rates of Recidivism

Previous research reported the sexual recidivism rate in the UK was 5.8% over a two-year follow up period, rising to 17.4% over six years or more (Craig et al, 2008). The current study revealed a similarly low sexual recidivism rate of 3.6%, albeit within a shorter follow up period (12 months). When the rearrest data was considered beyond the 12 month follow up period, sexual recidivism rates increased to 10.4% over eight years. This remains lower than the findings of Craig et al (2008) for the open-ended time frame, however, it falls within the range of typically observed sexual recidivism rates (10%-15%) reported by Hanson and Morton-Bourgon (2005). A possible explanation for the slightly lower rates, may be due to the various definitions of ‘recidivism’. In the current study, sexual recidivism was defined as a further arrest for a sexual offence, however it is widely acknowledged that definitions vary across the research (Hanson & Bussière, 1998; Hanson & Morton-Bourgon, 2005; Hanson & Morton-Bourgon, 2009). Additionally, as the current study used only BTP data, observed recidivism rates may be underestimated, as subsequent arrests outside the rail network were not included.

Regarding crime harm, the ‘Pareto Principle’ proposes 20% of offenders are generally responsible for 80% of harm (Sherman, 2007). However, in the current dataset, the Pareto curve revealed just 10% of offenders were responsible for over 80% of the harm. This indicates a small subset of offenders disproportionately contributed to the overall harm and supports that a ‘power few’ exists within the sample (Sherman, 2007). This highlights where BTP should target resources to effectively 'pursue' the most harmful offenders, as outlined in the 4P approach and stresses the importance of an accurate tool to identify these offenders.

Key Characteristics

The current findings support the well-established association in literature between males and the perpetration of sexual offences (Hohl & Stanko, 2022; VKPP, 2024). Over 90% of offenders in the current sample were males, a trend that was consistent amongst recidivists, non-recidivists, and all sexual offence sub-groups. This suggests the proportion of male perpetrators is higher for sexual offences, compared to overall VAWG offences, where males account for 75% of offenders (VKPP, 2024). Previous research reported the most common age group for sex offenders is 30 to 39 years old (Emeagi et al, 2024). In the current study, the most common age group was between 25 to 29 years old (when using the 5-year intervals) (Mean age = 35, SD= 12.6). However, when using 10-year intervals, offenders aged 20 to 29, and 30 to 39 years old, accounted for almost equal proportions of the overall sample. These findings suggest the age distribution in the current sample is slightly wider and incorporates younger offenders. The current study also supports the observation that the prevalence of sexual offences declines with age (Craig et al, 2008). The most common recorded ethnicity within the current sample was White – North European. This was consistent across sexual offence subgroups, recidivists and non-recidivists. Given that existing research regarding ethnicity is mixed, these findings provide support for the report by the MOJ, HO and ONS (2013). Unfortunately, nearly half the sample had no employment details recorded at the time of arrest. This is concerning given police obligations under Common Law Police Disclosure (CLPD), which requires the disclosure of information to enable third parties to mitigate risks in relation to employment or voluntary roles (NPCC, n.d.). Without this employment information, police may miss opportunities to *prevent* harm as set out in the 4P approach.

Risk Factors

Regarding risk factors associated with recidivism, the current study found between 35% to 40% of reoffenders were unemployed, compared to 17.7% of non-recidivists. On average, recidivists also had a higher number of total arrests, including violent, public order and drugs-related arrests. These findings support the existing literature that purports criminal lifestyle and anti-social orientation are associated with sexual recidivism (Hanson & Bussière, 1998; Hanson & Morton-Bourgon 2005; Seto et al, 2023). Notably, the current study reported the 12-month recidivist sample had fewer *sexual* arrests than non-recidivists, contradicting previous research which states individuals with prior sexual convictions were more likely to reoffend (Hanson & Bussière, 1998). However, extending the analysis to an 8-year timeframe, revealed recidivists had the highest number of sexual offences, thereby supporting earlier findings. Furthermore, Hanson and Bussière (1998) identified individuals who committed a diverse range of sexual crimes were more likely to reoffend. The current study revealed that offenders in the 12-month recidivist sample committed a *less* diverse range of offences, compared to those in the non-recidivist sample. However, a more diverse range of sexual offending was present in the 8-year recidivist sample – with a larger proportion of offenders committing offences across two and three subgroups. The inconsistencies with the 12-month sample are likely due to the small sample size.

Differences Across Subgroups

Slight differences were observed in the characteristics and risk factors among sexual offence subgroups, with digitally enabled offenders appearing the most distinct. Regarding age, the number of non-contact offenders in each age group did not decrease as expected, instead remaining relatively stable between 25 to 39 and 40 to 59 years old. This supports previous research suggesting the decline in sexual offending with age varies depending on offence type (Craig et al, 2008). It indicates that although age has previously been described as a strong predictor, the consistency of its predictive validity may vary across different offence types (Rice & Harris, 2014).

Babchishin, Hanson and Hermann (2011) reported online sex offenders were generally younger than offline offenders, however the current study reported digitally enabled offenders were generally older than other subgroups (Mean age = 37.9, SD= 14.4). These differences may stem from varying definitions of 'online' versus 'digitally enabled' offences. In the current study, the digitally enabled subgroup included voyeurism offences that required technology; however, these offences may not necessarily be considered an online offence as they do not require internet access. This reiterates the challenges of comparing studies using inconsistent classifications and supports that establishing clear definitions of sexual offence subgroups could benefit future research. Interestingly, digitally enabled offences had the highest proportion of employed offenders compared to the other subgroups. This builds on the findings from Babchishin, Hanson and VanZuylen (2013) which reported online offenders had greater academic achievements. This may be explained by the technological skills required to commit online offences, or the expense associated with digital offending – e.g. cost of a laptop. Given the existing literature largely identifies *unemployment* as a key risk factor associated with increased recidivism, this finding is important as it suggests unemployment may not hold the same relevance for digitally enabled sex offenders. Consequently, including 'unemployment' as a risk factor in existing risk assessment tools could lead to digital offenders being inaccurately assessed as lower risk. The same meta-analysis also reported online offenders had greater levels of anti-sociality, including a higher number of prior offences (Babchishin, Hanson & VanZuylen, 2013). This is reinforced in the current study as digitally enabled offenders had some of the highest numbers of total arrests, including both sexual and drug-related arrests.

Finally, the current study found contact, non-contact, digitally enabled, and other sexual offences demonstrated positive correlations between previous arrests, and a subsequent arrest for the same offence type. Notably, only the correlation for contact offences was statistically significant (p=0.03), possibly due to the very small sample size of other offence subgroups. This supports earlier research that proposes patterns of like-for-like forecasting (Emeagi et al, 2024). Negative correlations (albeit very low, and) were also found between *different* subgroups of sexual offences. This strengthens the suggestion that a previous arrest for one offence type, is associated with a lower likelihood of a further arrest in a different subgroup (Emeagi et al, 2024). Many existing risk assessment tools rely on the total count of previous sexual offences, however this finding suggests that focusing on specific sexual offence subgroups could provide more practical applications. Similarly, although previous 'attempted' sexual offences did not correlate with further arrests for the same offence type (*namely due to the absence of 'attempted' offences within the follow-up period*), they did reveal a positive correlation with 'contact' offences, indicating a potential pattern of escalation. Incorporating counts of specific sexual offence subgroups into risk assessment tools, could assist police forces in understanding risk, and prioritising robust interventions for offenders with a history of 'attempted' sexual offences, to *prevent* escalation, as outlined in the 4P approach.

RF Model

The initial RF model (RF1) achieved an accuracy score of 0.96, which is higher than the average accuracy score achieved by ML algorithms when forecasting *general* recidivism (Travaini et al, 2022). It also matches the highest accuracy score achieving using ML to predict *sexual* recidivism (Travaini et al, 2022). This initially supports that ML algorithms provide a promising alternative to existing risk assessment tools, which typically only demonstrate moderate predictive validity (Hanson & Morton-Bourgon, 2009). However, further analysis revealed the model was performing well when predicting true positive cases (*sensitivity*), but extremely poorly at predicting true negative cases (*specificity*). This indicates the model produced a large number of false positive predictions, forecasting further arrests for a sexual offence, when none occurred (Kovalchuk et al, 2023). As the recidivism rate within the current sample was only 3.6%, this aligns with previous research suggesting low recidivism rates increase the likelihood of false positives (Craig et al, 2005; Tully, Chou & Browne, 2013). The substantial decline in the specificity score between the training and testing data indicates the model was overfitting, as the scores did not generalise well to the unseen cases in the testing data. Due to the imbalanced nature of the outcome variable in the training data (2322 = 0, 89 = 1), it is likely that the RF model was also overpredicting the minority class (e.g. recidivism) (Gong & Kim, 2017).

Given the potential consequences of false predictions, measures to reduce overfitting and underfitting were implemented, including adjusting the number of trees. For most iterations of the model, performance remained the same, or improved slightly when the number of trees was decreased, suggesting that a larger number of trees did not contribute any new information to the classifications (Oshiro, Perez & Baranauskas, 2012). Therefore, the optimum number of trees was selected as the smallest number that did not impact the model performance, allowing for faster processing and greater *interpretability* (Oshiro, Perez & Baranauskas, 2012). Due to concerns regarding the imbalanced nature of the data, the minority class (1 = recidivists) was artificially inflated using the 'ROSE' function in R Studio. This seemingly improved the model performance, as the specificity score increased from 0.08 to 0.31, without impacting the sensitivity or overall accuracy. However, as the specificity score remained generally low, it was evident the model was still producing numerous false positive predictions. As such, the majority class (0 = non-recidivist), was under-sampled. This increased the specificity score above 0.5 without compromising the sensitivity or accuracy.

Following cross-validation, the model achieved a sensitivity of 0.56, and specificity of 0.97, indicating the model performed well in identifying true negative cases, however it's performance in predicting true positives had declined. This is likely due to the initial model over-fitting when using the single training and testing subset. This highlights the compromise researchers and practitioners must consider when employing ML classification tools. Similarly to Kovalchuk et al (2023), the current study prioritized a high specificity score, as the potential consequences of misclassifying individuals as recidivists were deemed to have the greatest impact. Such misclassifications could subject an innocent individual to prohibitive conditions, and potentially increase their likelihood of reoffending (Andrews & Bonta, 2014; Bonta & Andrews, 2007). However, this prioritisation resulted in an increased number of false negatives. In practical applications, police must carefully weigh the cost and benefits of false predictions and define acceptable error rates tailored to the context. Whilst the overall accuracy and specificity scores indicate good performance (as the model correctly forecasts non-recidivists 97% of the time), the sensitivity score indicates that the current model only identifies 55% of recidivists. This suggests the current ML algorithm may not exceed the overall predictive validity of existing risk assessment tools.

The most important variables for the model's performance, as identified by the mean decrease in accuracy, were arrests for contact and non-contact sexual offences and the CCHI score. This demonstrates using subgroups of sexual offences as a predictor, rather than total sexual arrests, may improve the accuracy of existing risk assessment tools. Of the non-sexual offence categories included in the model, public order was the most important to the overall accuracy. The least important variables to the model were gender, drugs-related arrests and ethnicity. Figure 21 revealed gender and drugs-related offences reduced the model's overall accuracy. Removing these could potentially improve performance, however it is important to consider variable interaction - although these variables reduce the overall accuracy in isolation, they may still contribute positively when combined with other variables. To ensure *fairness* within the model, as emphasized in the "EU Requirements of Trustworthy AI", ethnicity could be removed to minimise racial bias and address ethical concerns, however potential variable interaction should be carefully evaluated (Farayola et al, 2023).

Finally, the results of the five-fold cross-validation indicate the model performed *consistently* and *reliably*, achieving similar accuracy scores across all five folds of the data. This aligns with two of the additional requirements suggested by Farayola et al (2023). However, further analysis would benefit from examining the sensitivity and specificity scores across training and testing data, for each fold as this could provide a more accurate measure of the model's overall reliability.

Implications

Practical

Regarding practical implications, the findings clearly demonstrate the existence of a 'Power Few' amongst sex offenders on the rail network (Sherman, 2007, p299). As such, to effectively reduce harm, police should identify and target these individuals. The risk principle of the RNR model asserts that criminal behaviour is predictable (Bonta & Andrews, 2007). Thus, enhancing the understanding of sexual offences on the rail network, alongside risk factors associated with recidivism, can improve existing risk assessment tools. Current policies often regard all sex offenders uniformly, however this research suggests risk assessment frameworks should distinguish between subgroups of offenders to enable tailored interventions. For example, as digitally enabled offences were positively correlated with further arrests for the same offence type, interventions can focus on limiting offender's access to technology. This aligns with the motivation-facilitation model by limiting the situational factors that may influence reoffending. Additionally, the findings on sexual offence locations indicate where BTP should target patrols and creates a foundation to build upon regarding possible sexual offence hotspots (Ariel et al, 2024a).

A further practical implication is the application of ML algorithms to forecast recidivism, which aligns with NPCC's commitment to utilise existing data and big technology to disrupt repeat offenders (Hohl & Stanko, 2022; COP & NPCC, 2024). Incorporating ML algorithms into risk assessment frameworks, ensures predictions are data-driven and evidence-based. While the current RF model produced a moderate number of false negative predictions, its performance could likely be enhanced by incorporating additional variables or expanding the dataset. Achieving high sensitivity and specificity scores could provide a defensible and robust mechanism to justify police decision making, and assist in effectively targeting interventions. To ensure *fairness* and *accountability*, algorithm outputs should not be considered in isolation and would require human oversight, particularly during the model's development, and in any subsequent decision-making process (Farayola et al, 2023; NPCC, 2024b).

Developing robust RF models relies on high quality, complete datasets. Although police forces often record lots of valuable formation, access can be challenging due to limitations in data governance processes and as it often stored across multiple systems. Additionally, datasets are often incomplete due to recording issues. To assist with this, police should seek to improve data recording processes and could benefit from a centralised system, accessed by all forces nationally. Furthermore, despite having access to large amounts of data, many forces lack the opportunities, skills and resources to construct and evaluate such models. By increasing alliances with educational establishments, and working collaboratively alongside practitioners, forces may be able enhance ML recidivism predictions.

Given the ability of ML to process large datasets quickly, it could be utilised to analyse the evolving nature of sexual offences, and differences between subgroups of offences, enabling earlier identification of emerging patterns. Currently, practitioners rely on numerous risk assessment tools, with varying performance. Implementing ML methods could streamline this process and ensure consistency across UK police forces. However, implementation would not be without its challenges and the requirements of trustworthy and ethical AI would be essential considerations.

Policy

By using *arrest* data to construct a RF model, the current study proposes risk assessments of sex offenders should be conducted earlier in the criminal justice process. Currently, many sex offenders are not subject to a risk assessment until after charge, or conviction, which significantly limits early opportunities to prevent harm. If policymakers were to recommend sex offenders should be subject to a risk assessment at the point of arrest, this would allow the opportunity to implement various interventions sooner. This proactive approach would also bring the risk assessment process in line with other VAWG offences. Whilst this may be contentious, as not all arrests lead to conviction, these risk assessments would not result in automated interventions, but support decision making in selecting the most appropriate preventative measures. Equally, as risk assessments are only conducted following conviction, there is little evidence to suggest current pre-convictions interventions are targeted fairly or effectively. Targeting interventions earlier in the criminal justice process not only has benefits in harm prevention, but also demonstrates a cost-benefit by reducing the costs associated with reoffending, and resource misallocation (Whitten, 2024).

Research

This study sought to expand the understanding of sexual recidivism on the rail network, and to add to the literature regarding the application of ML algorithms to forecast sexual recidivism. To assist with future research, this study suggests defining each sexual offence subgroup clearly to ensure consistency and enable more meaningful comparisons across studies. Though the current study supports that ML algorithms may not demonstrate advanced predictive performance, this is model specific and should not dismiss the potential of ML altogether. Further research should build on this study's foundation, by seeking to improve the model performance, through additional variables, a larger dataset and advanced outcomes. It could also compare the performance of RF models against current risk assessment tools, using the same sample of offenders.

Limitations and Further Research

When interpreting these findings, it is important to acknowledge limitations to ensure the recommended implications can be considered with these in mind. These limitations can serve as a valuable guidance to shape further research and develop understanding of sexual recidivism and ML algorithms.

The current study relied solely on BTP data and did not include any arrests that occurred beyond the jurisdiction of the rail network. This may affect the *external* validity of the findings, by potentially underestimating recidivism rates if offenders were arrested for further offences beyond this jurisdiction. Additionally, restricting the data to BTP jurisdiction, may limit the inclusion of certain offence types. For example, BTP are less likely to encounter digitally enabled offences, due to the limited digital infrastructure on the rail network. To develop this research further, arrest data from PNC could be incorporated into the dataset. This would improve the validity of the study, and enable the ML algorithm to be developed nationally. Furthermore, the use of arrest data could also be considered a limitation of the current study as this may artificially inflate rates of recidivism as not all those who are arrested, are subsequently prosecuted (Whitten, 2024). However, this data was chosen with potential implications in mind. By using arrest data, this study sought to acknowledge the possible benefits of conducting risk assessments for sexual offenders earlier. If high-harm offenders can be identified prior to any charge or conviction, this enables police to target evidence-based, effective interventions to prevent further harm (Whitten, 2024).

Similarly to Emeagi et al (2024), the sample size in the current study was relatively small compared to previous studies of a similar nature. This limitation potentially contributed to an imbalance of the outcome variable within the data. Thus, to prevent overfitting and underfitting of the RF model, both oversampling and under-sampling were applied. Oversampling increases the number of the minority class by generating synthetic examples of minority cases. Whilst this approach improved performance of the model, it can be criticized for relying on synthetic data, rather than actual observations (Lunardon, Menardi & Torelli, 2014). Likewise, to under-sample the outcome variable, a subset of the majority class is removed from the dataset (Lunardon, Menardi & Torelli, 2014). This reduces the number of cases available for the algorithm to learn from and may be criticized for the loss of information from actual observations. Further research could benefit from a larger sample size which could potentially be achieved through extending the follow up period beyond 12 months. Alternatively, since the current study found the average time between the initial arrest and subsequent arrest for a further sexual offence was 545 days, a further study may benefit from using a follow up period of 2 years following the initial arrest.

Given that this study sought to contribute to an emerging field of research, it developed a foundational RF model to establish a baseline understanding of performance. Subsequently, the current model was reliant on static risk factors, which is a common criticism of existing actuarial risk assessment tools (Craig et al, 2005). Whilst static factors help to identify individuals who pose a long-term risk, incorporating dynamic risk factors could enhance the overall accuracy, sensitivity and specificity (Bonta & Andrews, 2007; Craig, Browne, Stringer & Beech, 2005; Seto et al, 2023). Additionally, the current RF model was constructed using a binary outcome variable – i.e. to forecast if an individual would be arrested for a further sexual offence, or not. The classification from the algorithm did not provide any further detail regarding the potential severity of future sexual offences, nor did it provide a risk rating similar to other existing risk assessment tools (Sjöstedt & Grann, 2002). Further research could build on this by developed more advanced outcome variables to include practical information that will assist the police in targeting the appropriate resources, such as the likelihood, imminence or severity of a further offence. By incorporating risk ratings, such as low, medium or high risk, police could prioritise interventions dependent on level of assessed risk.

Chapter Summary

This chapter explored the findings within the context of the existing body of research, considering the significance of these results and possible explanations for any differences. It contemplated potential implications prior to addressing the limitations of this study and highlighting opportunities for further research.

Conclusion

Background

Sexual offences in the UK represent a growing epidemic, with the *nature* of these crimes evolving and the *number* of crimes rising, despite widespread underreporting (Garg, 2023; COP & NPCC, 2024). Concerningly, research indicates public transport environments show a high prevalence of USB, with various situational features that may facilitate offending. Amid growing pressures on police to pursue and disrupt sexual offenders, recent evidence suggests subgroups of sexual offences are becoming increasingly distinct (Emeagi et al, 2024). To effectively reduce the harm caused by sex offenders, police require a robust mechanism to identify repeat and high-harm offenders. However, existing risk assessment tools only demonstrate moderate predictive validity and their performance varies across this evolving landscape of sexual offending (Långström, 2004; Parent, Guay & Knight, 2012) Emerging research indicates AI and ML algorithms offer promising alternatives, by rapidly identifying patterns from existing police data (Farayola et al, 2023; Travaini et al, 2022).

This research set out to address gaps in the available literature regarding the nature of sexual offences and rates of sexual recidivism on the UK rail network. It utilised arrest data and descriptive statistics to understand the key characteristics of sex offenders, risk factors associated with increased reoffending and any differences between sex offender subgroups. Finally, the research developed a Random Forest Model to assess the predictive validity of ML algorithm in forecasting future arrests for sexual offences.

Key Findings

RQ1: What is the nature and prevalence of sexual offences on the UK's rail network, and what are the rates of sexual recidivism among offenders?

The current study revealed the rate of sexual recidivism on the rail network is generally low, with only 3.6% of individuals being arrested for a further sexual offence within a 12-month follow up period, rising to 10.4% over 8 years. These rates align closely with figures reported in earlier research. Contact sexual offences accounted for the largest proportion of sexual arrests, followed by non-contact which contributed to almost one third of all sexual arrests. The most-common sexual offences on the rail network were *sexual assault on a female, outraging public decency* and *exposure.* Collectively, these offences accounted for all three quarters of all sexual arrests. As non-contact offences often carry a lower crime harm score, this highlights the importance of considering crime counts alongside harm scores (Sherman, Neyroud & Neyroud, 2016). Furthermore, sexual offences were recorded twice as frequently in B-Division compared to C-Division, with over half the reported offences occurring onboard a train, compared to other areas of the station. Sexual offences numbers remained relatively consistent throughout the week, though a slight increase was observed with Fridays and Saturdays which aligns with nighttime economy.

RQ2: What are the characteristics and risk factors associated with sexual recidivism on the rail network?

The characteristics of offenders in the current sample largely reflect existing research. Most sex offenders on the rail network are male, aged between 25 to29 and White-North European. However, offenders aged between 20 to 29 and 30 to 39 years accounted for almost equal proportions of sexual offences. Notably, individuals arrested of a further sexual offence were more likely to be unemployed and had a higher number of total previous arrests including sexual arrests. Additionally, recidivists appeared to commit a more diverse range of sexual offences, with higher proportions of offences occurring across two and three different offence subgroups.

RQ3: Do characteristics of sex offenders differ, based on type of offending?

Minor differences were observed amongst the different subgroups of sexual offenders, with those who committed digitally enabled offences appearing to be the most distinct. On average, digitally enabled offenders were slightly older (mean age = 37.9, SD= 14.4), while non-contact offenders maintained consistent offending rates as they aged. Additionally, digitally enabled offenders had the highest proportion of employed individuals and greatest number of sexual arrests. A pattern of like-for-like prediction was observed for all sexual offence subgroups except for attempted sexual offences, though only the correlation for contact offences was statistically significant (p=0.03). This suggests an arrest for one subgroup may predict a further arrest within the same subgroup (e.g. a contact offence predicts a further contact offence). A small negative correlation was also observed between different subgroups, indicating an arrest for an offence in one sexual offence subgroup may decrease the likelihood of a further offence within a different subgroup.

RQ4: What is the predictive validity of a machine learning algorithm in forecasting sexual recidivism on the rail network?

The RF model achieved high accuracy scores across both the training and testing data when predicting a further arrest within a 12-month follow up period (1.00 and 0.96 respectively). The model performed well at predicting true positives, however it produced many false positives. The substantial decrease in the specificity score between the training and testing data indicates the model was overfitting, potentially due to the unbalanced nature of the outcome variable. Given the potential consequences of false positive predictions, whereby an innocent individual could be mislabelled as a recidivist, over-sampling, under-sampling and finetuning methods were employed. The final RF model achieved a sensitivity of 0.56, a specificity of 0.97 whilst the overall accuracy remained at 0.96. This indicates the model correctly identified non-recidivists 97% of the time, however it only classified recidivists correctly 55% of the time. Previous arrests for contact offences, non-contact offences and CCHI score were the top three most important variables to the model. The least importance variables to the model's performance were gender, previous drugs-related arrests and ethnicity.

Implications

Practical

Using the characteristics identified in this study, alongside the risk-factors associated with recidivism, BTP can better identify individuals at higher risk of reoffending. This understanding will assist in enhancing recidivism predictions, support the development of effective risk assessment tools, and enable tailored interventions for different subgroups of sexual offences.

This research proposes that policymakers consider ML algorithms as potential alternatives to existing risk assessment tools, which may not be performing consistently across all subgroups of sexual offences. The current model was able to successfully predict no further arrest for a sexual offence 97% of the time, which demonstrates higher predictive accuracy than existing methods. However, achieving this high specificity score came at the cost of generating a moderate number of false negative predictions. This underscores the importance for police forces and policymakers to consider the cost-benefits and define acceptable error rates, prior to implementing any predictive algorithms operationally. Due to their ability to process large datasets quickly, ML algorithms enable the early identification of emerging patterns and their possible impact on sexual recidivism more quickly than the human equivalent. This capability enables law enforcement the opportunity to adapt more effectively to the evolving nature of sexual offending, whilst also streamlining the existing process, which currently relies on numerous different risk assessment tools, each with varying performance.

The implementation of ML algorithms would not be without its challenges. To do so, requires buy-in from senior leaders and key stakeholders, alongside a suitable digital infrastructure. ML algorithms are complex, and forces would require individuals who understand how to construct the model, regularly evaluate its performance, and make necessary amendments. Additionally, police forces would need to ensure that officers and staff members using the algorithm can effectively interpret the outputs. If implemented, the current research proposes that ML outputs should not be used in isolation, but alongside human expertise (*human-in-the-loop*) in line with EU Requirements (Farayola et al, 2023). Effective implementation would require robust, transparent, and explainable methodology, alongside responsible and accountable decision-making consistent with the NPCC Covenant for using AI in Policing (NPCC, 2024). Whilst the application of ML algorithms to forecast recidivism may initially challenge trust and confidence in policing, due to the lack of transparency and 'black-box' nature, they also hold the potential to increase trust and confidence – provided a fair and ethical model is developed with minimal false predictions. Although the current model is not perfect, it may outperform 'business as usual' and it creates a foundation for further research. Due to the complexity of ML algorithms, this research recommends further collaboration between practitioners and academics to increase the opportunity, skills and resources, enabling police forces to develop algorithms.

Policy

By using *arrest* data to construct the RF model, the current study also proposes that policymakers consider moving the risk assessments for sexual offenders earlier within the criminal justice process. Currently, risk assessments are generally conducted by prison or probation services *following conviction* for a sexual offence. This limits opportunities for interventions and delays the implementation of these. By conducting the risk assessment at the point of arrest, this would align the process with other VAWG offences, such as the DASH risk assessment form for all domestic abuse, stalking and harassment cases. Furthermore, it would enable forces to consider interventions earlier in the process, and potentially prevent further harm. If implemented pre-charge, the outcome of such risk assessment could also be used to inform police decisions regarding disclosures (such as CLPD) or referrals (similar to the MARAC process), which would support the whole-system approach suggested by COP and NPCC (2024). Whilst this approach may be considered contentious, as not all who are arrested are subsequently convicted, the risk assessment outcome would not be proposed to automate any interventions, however could be used as a tool alongside professional judgement to inform police decisions.

Research

The current research adds to emerging literature regarding the use of ML algorithms to forecast sexual recidivism. Though the final model did achieve an impressive specificity score, this was accomplished using under-sampled data, and the result of a trade-off with the sensitivity score. To contextualise these findings, future research could compare the performance of the current model against existing risk assessment tools, using the same sample of offenders to see if this outperforms 'business as usual'. Further research could also potentially improve the overall performance of the model, particularly with the inclusion of PNC arrest data, more dynamic risk factors and expanding the outcome variable to include different levels of risk. As the landscape of sexual offending continues to evolve, it is crucial that all existing risk assessment tools are regularly evaluated to ensure they can still effectively and efficiently forecast risk (Emeagi et al, 2024). Beyond this, further research could evaluate which interventions are most effective in reducing harm, for each assessed level of risk, or each subgroup of sex offender.

Final synthesis

Within a landscape of reduced police funding and declining police resources, alongside increased numbers of sexual offences, this research demonstrates the potential of using data-driven ML algorithms to transform the risk assessment of sex offenders. By enhancing the accuracy of recidivism predictions, particularly in light of the evolving nature of sexual offences, ML can inform a proactive approach to pursue and prevent further harm. By utilising arrest data in the RF model, the current study advocates for earlier risk assessment and intervention, promoting evidence-based reforms that prioritise resource efficiency alongside public safety. These findings contribute to the emerging research towards streamlining risk assessment processes, and provide a foundation for continued innovation in targeting sexual offending.

References

Andrews, D. A. and Bonta, J. (2010) *The psychology of criminal conduct* (5th ed). New York: Routledge.

Andrews, D. A., Bonta, J. and Wormith, J. S. (2006) 'The recent past and near future of risk and/or need assessment', *Crime & Delinquency,* 52(1): 7-27.

Ariel, B., Langton, J., Peters, K., Webster, K. and Assaraf, N. (2024a) 'Private security for curbing unwanted sexual behaviours in train stations: a place-based randomised controlled trial,' *Journal of Experimental Criminology*, 1-29.

Ariel, B., Ceccato, V., McDonnell, A. and Webster, K. (2024b) 'Experiences and reporting of unwanted sexual behaviours on Great Britain's rail network: A survey of victims and witnesses with an embedded randomized vignette experiment on callback effects,' *Victim & Offenders,* 1-30.

Babchishin, K. M., Hanson, R. K. and Hermann, C. A. (2011) 'The characteristics of online sex offenders: A meta-analysis', *Sexual Abuse,* 23(1): 92-123.

Babchishin, K. M., Hanson, R. K. and VanZuylen, H. (2015) 'Online child pornography offenders are different: A meta-analysis of the characteristics of online and offline sex offenders against children', *Archives of sexual behaviour,* 44: 45-66.

Bamford, J., Chou, S. and Browne, K. D. (2016) 'A systematic review and meta-analysis of the characteristics of multiple perpetrator sexual offences,' *Aggression and Violent Behaviour,* 28: 82-94.

Barbaree, H. E., Langton, C. M. and Peacock, E. J. (2006) 'Different actuarial risk measures produce different risk rankings for sexual offenders', *Sexual Abuse,* 18(4): 423-440.

Bengtson, S. and Långström, N. (2007) 'Unguided clinical and actuarial assessment of re offending risk: A direct comparison with sex offenders in Denmark,' *Sex Abuse*, 19: 135-153.

Berk, R. (2012) *Criminal justice forecasts of risk: A machine learning approach.* New York: Springer.

Berk, R. (2013) 'Algorithmic criminology', *Security Informatics*, 2: 1-14.

Berk, R., Sherman, L., Barnes, G., Kurtz, E. and Ahlman, L. (2009) 'Forecasting murder within a population of probationers and parolees: a high stakes application of statistical learning', *Journal of the royal statistical society: series A (statistics in society),* 172(1): 191-211.

Bland, M. (2020) 'Algorithms can predict domestic abuse, but should we let them?', in H. Jahankhani, B. Akhgar, P. Cochrane and M. Dastbaz (eds) *Policing in the Era of AI and Smart Societies,* Switzerland: Springer Cham, pp 139-155.

Boer, D. P. and Hart, S. D. (2012) 'Sex offender risk assessment: Research, evaluation, 'best practice' recommendations and future directions', in J. Ireland, C. Ireland and P. Birch (eds) *Violent and sexual offenders. Assessment, treatment and management,* 1st ed., London: Routledge, pp. 49-64.

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

Breiman, L. (2001) 'Random Forests', *Machine Learning*, 45: 5–32.

British Transport Police Authority (n.d.). BTPA Strategic Plan 2022-2027. Retrieved 30th April 2023 from https://btpa.police.uk/wp-content/uploads/2022/05/9433_BTPA_StrategicPlan2022_Digital.pdf.

Cann, D. (2017) 'Sex offender policies that spin the revolving door: An exploration of the relationships between residence restrictions, homelessness, and recidivsim', [Unpublished

Thesis submitted for master's degree in criminology and criminal justice, College of Arts and Sciences, University of South Carolina].

Caputo, M., Fineman, M. and Khan, S. (2024) 'Sexual assault and the matrix of harm: Sexual assault survivors narrate their whole lives in more negative ways', *PLoS One,* 19(6): 1-27.

Ceccato, V. and Loukaitou-Sideris, A. (2022) 'Fear of sexual harassment and its impact on safety perceptions in transit environments: a global perspective', *Violence against women,* 28(1): 26-48.

College of Policing (2016). First response quick reference guide – domestic abuse. Retrieved 15th April 2024 from https://www.college.police.uk/app/major-investigation-and-public-protection/domestic-abuse/first-response-quick-reference-guide-domestic-abuse.

College of Policing (2017). Identifying, assessing and managing risk. Retrieved 8th November 2024 from https://www.college.police.uk/app/major-investigation-and-public-protection/managing-sexual-offenders-and-violent-offenders/identifying-assessing-and-managing-risk.

College of Policing (n.d.). Briefing note; For police first responders to a report of rape or sexual assault. Retrieved 12th April 2024 from https://assets.production.copweb.aws.college.police.uk/s3fs-public/2023-04/Briefing-note-for-police-first-responders-to-a-rape-or-sexual-assault.pdf.

College of Policing (2023). Embedding the national operating model in your area. Retrieved 21st December 2024 from https://www.college.police.uk/national-operating-model-rasso/embedding-national-operating-model.

College of Policing (2024). Sharing photographs or film of people in an intimate state. Retrieved 20th

November 2024 from https://www.college.police.uk/guidance/sharing-photographs-of-people-intimate-state.

COP & NPCC (2024). Violence Against Women and Girls (VAWG) National Policing Statement 2024. Retrieved 6th August 2024 from file:///C:/Users/6482Fell/Downloads/National%20Policing%20Statement%202024%20For%20Violence%20Against%20Women%20and%20Girls%20(VAWG)%20-%20July%202024%20WEBSITE%20PUBLICATION%20(8).pdf.

Cortoni, F., Babchishin, K. M. and Rat, C. (2017) 'The proportion of sexual offenders who are female is higher than thought: A meta-analysis', *Criminal Justice and Behaviour,* 44(2): 145-162.

Craig, L., Browne, K., Stringer, I. and Beech, A. (2004) 'Limitations in actuarial risk assessment of sexual offenders: A methodological note,' *The British Journal of Forensic Practice,* 6(1): 16-32.

Craig, L. A., Browne, K. D., Stringer, I. and Beech, A. (2005) 'Sexual recidivism: A review of static, dynamic and actuarial predictors', *Journal of Sexual Aggression,* 11(1): 65-84.

Craig, L. A., Browne, K. D., Stringer, I. and Hogue, T. E. (2008) 'Sexual reconviction rates in the United Kingdom and actuarial risk estimates', *Child Abuse & Neglect*, 32(1): 121-138.

Craig, L. A. and Rettenberger, M. (2016) 'A brief history of sexual offender risk assessment' in D. R. Laws and W. O'Donohue (eds) *Treatment of sex offenders: Strengths and weaknesses in assessment and intervention*, 1st ed., Switzerland: Springer Cham, pp. 19-44.

Craik, A., Han, L., Sullivan, L., Landsiedel, J., Travers, Spaull, C. and Howard, P (2024) Revalidation: Risk of recidivism tools. An evaluation of the actuarial instruments developed to assess recidivism risk in England and Wales (Ministry of Justice Analytical

Series). Retrieved 4th October 2024 from https://assets.publishing.service.gov.uk/media/65bbca4ecc6fd6000d5dbeb5/revalidation risk-recidivism-tools.pdf.

Crookes, R. L., Tramontano, C., Brown, S. J., Walker, K. and Wright, H. (2022) 'Older individuals convicted of sexual offenses: A literature review', *Sexual Abuse,* 34(3): 1-31.

Dworkin, E. R., Jaffe, A.E., Bedard-Gilligan, M. and Fitzpatrick, S. (2021) 'PTSD in the year following sexual assault: A meta-analysis of prospective studies', *Trauma, Violence & Abuse,* 24(2): 497-514.

Emeagi, C., Sullivan, L., Landsiedel, J., Craik, A. and Howard, P. (2024) The Actuarial Prediction of Sexual Reoffending: Responding to Changing Offending Patterns (Ministry of Justice Analytical Series). Retrieved 10th October 2024 from https://assets.publishing.service.gov.uk/media/65bbcb0521f73f0014e0bafe/actuarial-prediction-sexual-reoffending.pdf.

Etzler, S., Schönbrodt, F. D., Pargent, F., Eher, R. and Rettenberger, M. (2024) 'Machine learning and risk assessment: Random forest does not outperform logistic regression in the prediction of sexual recidivism', *Assessment,* 31(2): 460-481.

Falshaw, L., Bates, A., Patel, V., Corbett, C. and Friendship, C. (2003) 'Assessing reconviction, reoffending and recidivism in a sample of UK offenders', *Legal and Criminological Psychology,* 8: 207-215.

Farrington, D. P. (1995) 'The development of offending and antisocial behaviour from childhood: Key findings from the Cambridge study of delinquent development', *Journal of Child Psychology and Psychiatry,* 6(36): 929-964.

Farrington, D. P., Tofi, M. M. and Piquero, A. R. (2015) 'Risk, promotive, and protective factors in youth offending: Results from the Cambridge study in delinquent development', *Journal of Criminal Justice,* 45: 63-70.

Farayola, M. M., Tal, I., Connolly, R., Saber., T. and Bendechache, M. (2023) 'Ethics and trustworthiness of AI for predicting the risk of recidivism: A systematic literature review', *Information,* 14(8): 426.

Fellows, E. C. (2023) 'Select a distinct group of either victims or offenders and critically assess Current activity (in your agency or more generally) for reducing harm in this group against available evidence. Suggest and critique possible improvements to this activity using the Triple T model as framework', *MSt Applied Criminology and Police Management Year 1*. University of Cambridge. Unpublished essay.

Fellows, E. C. (2024a) 'Research proposal', *MSt Applied Criminology and Police Management Year 2*. University of Cambridge. Unpublished essay.

Fellows, E. C. (2024b) *Thesis Oversight Committee* [Presentation to Thesis Oversight Committee, MSt Applied Criminology and Police Management]. Cambridge University. 16 September 2024.

Forsdike, K., Ison, J., Hooker, L., Henry, N. and Taft, A. (2024) '"God, whatever you do, don't tell people it's unsafe": Public transport service providers' perspectives on women's safety from sexual violence on public transport', *Transport Policy,* 150: 14-23.

Garg, A. (2023) '*Serious Sexual Offences'* [Lecture to Mst Applied Criminology and Police Management Year 1], University of Cambridge. 6 July 2023.

Gekoski, A., Gray, J. M., Adler, J. R. and Horvath, M. A. H. (2017) 'The prevalence and nature of sexual harassment and assault against women and girls on public transport: an international review', *Journal of Criminological Research, Policy and Practice,* 3(1): 3-16.

Gong, J. and Kim, H. (2017) 'RHSBoost: Improving classification performance in imbalance data', *Computational Statistics & Data Analysis*, 111: 1-13.

Hanson, R. K. (1998) 'What do we know about sex offender risk assessment?' *Psychology, Public Policy and Law,* 4(1-2): 50-72.

Hanson, R. K. (2002) 'Recidivism and age: Follow-up data from 4673 sexual offenders,' *Journal of interpersonal violence,* 17(10): 1046-1062.

Hanson, R. K. and Bussière, M. T. (1998) 'Predicting Relapse: A Meta-Analysis of Sexual Offender Recidivism Studies', *Journal of Consulting and Clinical Psychology,* 66(2): 348-362.

Hanson, R. K. and Harris, A. J. (2000) 'Where should we intervene? Dynamic predictors of sexual offence recidivism', *Criminal Justice and Behaviour,* 27(1): 6-35.

Hanson, R. K., Harris, A. J. R., Helmus, L. and Thornton, D. (2014) 'High-Risk Sex Offenders May Not Be High Risk Forever', *Journal of Interpersonal Violence,* 29(15): 2792-2813.

Hanson, R. K. and Morton-Bourgon, K. E. (2004) '*Predictors of sexual recidivism: An updated meta analysis 2004-02'*, Canada: Public Safety and Emergency Preparedness.

Hanson, R. K. and 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): 1154-1163.

Hanson, R. K., and Morton-Bourgon, K. E. (2007) '*The accuracy of recidivism risk assessments for* *sexual offenders: A meta-analysis 2007-01,'* Canada: Public Safety and Emergency Preparedness.

Hanson, R. K., and Morton-Bourgon, K. E. (2009) 'The accuracy of recidivism risk assessments for sexual offenders: A meta-analysis of 118 prediction studies', *Psychological Assessment,* 21(1): 1-21.

Helmus, L. M. (2018) Sex offender risk assessment: Where are we and where are we going? *Currently Psychiatry Reports,* 20: 1-9.

Hohl, K. and Stanko, E. A. (2022) 'Five Pillars: A framework for transforming the police response to rape and sexual assault', *International Criminology,* 2: 222-229.

Home Office [UK]. (2023) *The Strategic Policing Requirement* [Government document]. Home Office, London. Available at: https://assets.publishing.service.gov.uk/media/64955fb9de8682000cbc8cf0/Strategic_Policing_Requirement_V1.3.pdf. (Accessed: 12 April 2024).

Howard, P. and Wakeling, H. (2021) Comparing two predictors of sexual recidivism: the Risk Matrix 2000 and the OASys Sexual Reoffending Predictor (Ministry of Justice Analytical Series). Retrieved 12th August 2024 from https://assets.publishing.service.gov.uk/media/600eaacfd3bf7f05c06dfc48/comparing-2-predictors-sexual-recidivism.pdf.

Jung, S., Ennis, L., Stein, S., Choy, A. L. and Hook, T. (2012) 'Child pornography possessors: Comparisons and contrasts with contact- and non-contact sex offenders', *Journal of Sexual Aggression,* 19(3): 295-310.

Kewley, S., Osman, S. and McGuinness, Á. (2020) 'How well do police specialists risk assess registered sexual offenders?', *Journal of Sexual Aggression,* 26(3): 302–315.

Kovalchuk, O., Karpinski, M., Banakh, S., Kasianchuk, M., Shevchuk, R. and Zagorodna, N. (2023) 'Prediction machine learning models on propensity convicts to criminal recidivism', *Information*, 14(3): 161.

Labour Party. (2024) *Change. Labour Party Manifesto*. London: Labour Party. Retrieved 13th January 2024 from Available at https://labour.org.uk/wp-content/uploads/2024/06/Labour-Party-manifesto-2024.pdf.

Långström, N. (2004) 'Accuracy of actuarial procedures for assessment of sexual offender recidivism risk may vary across ethnicity', *Sexual Abuse: A Journal of Research and Treatment,* 16: 107-120.

Loukaitou-Sideris, A. and Ceccato, V. (2021) 'Sexual harassment on transit: a global, comparative examination,' *Security Journal,* 35: 175-204.

Lunardon, N., Menardi, G and Torelli, N. (2014) 'ROSE: a Package for Binary Imbalanced Learning.' *R Journal,* 6(1):82-92.

Mannell, J., Lowe, H., Brown. L., Mukeriji, R., Devakumar, D., Gram, L., Jansen, H.A.F.M., Minckas, N., Osrin, D., Prost, A., Shannon, G. and Vyas, S. (2022) 'Risk factors for violence against women in high-prevalence settings: a mixed-methods systematic review and meta-synthesis', *BMJ Global Health,* 7(3): 1-14.

McAlister, R. (2014) 'New technologies, 'risk' and sexual offending.' in K. McCartan (ed.) *Responding to Sexual Offending: Perceptions, Risk Management and Public Protection*, 1st ed., London:

Palgrave Macmillan UK, pp. 72-92.

McNaughton Nicholls, C. and Webster, S. (2014). Sex Offender Management and Dynamic Risk: Pilot evaluation of the Active Risk Management System (ARMS) (Ministry of Justice Analytical Series). Retrieved 12th April 2024 from https://assets.publishing.service.gov.uk/media/5a7e41f1ed915d74e622522d/sex-offender-management-and-dynamic-risk.pdf.

Ministry of Justice, Home Office and Office for National Statistics (2013) An Overview of Sexual Offending in England and Wales. Statistics bulletin. Retrieved 17th September 2024 from https://assets.publishing.service.gov.uk/media/5a7ca66d40f0b65b3de0a47d/sexual-offending-overview-jan-2013.pdf.

Mitchell, R. J., Burns, N., Glozier, N. and Nielssen, O. (2023) 'Homelessness and predictors of criminal reoffending: a retrospective cohort study', *Criminal behaviour and mental health,* 33(4): 261-275.

NPCC (n.d). Common Law Police Disclosures (CLPD) – Provisions to supersede the Notifiable Occupations Scheme (NOS). Retrieved 11th December 2024 from https://assets.college.police.uk/s3fs-public/2022-04/NPCC-2017-Common-Law-Police-Disclosures-CLPD-%E2%80%93-Provisions-to-supersede-the-Notifiable-Occupations-Scheme-NOS.pdf

NPCC (2024) Covenant for Using Artificial Intelligence (AI) in Policing. Retrieved 15th June 2024 from https://science.police.uk/site/assets/files/4682/ai_principles_1_1_1.pdf.

O'Connell, M. (2024) *Leveraging AI Responsibly: What is AI and how does it impact on law* *enforcement and the delivery of justice.* [Webinar – What is AI and how does it impact on law enforcement and the delivery of justice?]. 02 May 2024. Available at https://www.workcast.com/AuditoriumAuthenticator.aspx?cpak=8406237530139160&pak=7273672883586985. Accessed on 02 May 2024).

Oshari, T. M., Perez, P. S. and Baranauskas (2012) 'How many trees in a random forest?', in P. Perner (ed) *Machine Learning and Data Mining in Pattern Recognition*, Berlin: Springer, pp. 154-168.

Office for National Statistics (2024). Crime in England and Wales: year ending June 2024. Retrieved 2nd November 2024 from https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingjune2024#:~:text=The%20trend%20in%20police%20recorded,June%202024%20were%20rape%20offences.

Parent, G., Guay, J. and Knight, R. A. (2012) 'Can we do better? The assessment of risk of recidivism by adult sex offenders', *Criminal Justice and Behaviour,* 39(12): 1647-1667.

Prime Ministers Office (2024). King's Speech 2024: background briefing notes. Retrieved 16th December 2024 from https://www.gov.uk/government/publications/kings-speech-2024-background-briefing-notes.

Protection from Sex-based Harassment in Public Act 2023. Available at: https://www.legislation.gov.uk/ukpga/2023/47/enacted (Accessed: 20 December 2024).

Ratcliffe, J.H. (2022) *Evidence-Based Policing: The Basics.* London and New York: Routledge Taylor & Francis Group.

Rice, M. E. and Harris, G. T. (2014) 'What does it mean when age is related to recidivism among sex offenders?', *Law and Human Behavior,* 38(2): 151-161.

Seto, M. C., Augustyn, C., Roche, K. M. and Hilkes, G. (2023) ‘Empirically-based dynamic risk and protective factors for sexual offending’, *Clinical Psychology Review,* 106*:* 102355.

Sherman, L. W. (2007) 'The Power Few: experimental criminology and the reduction of harm', *Journal of Experimental Criminology*, 3(4): 299-321

Sherman, L. W., Neyroud, P. W. and Neyroud, E. (2016) ‘The Cambridge Crime Harm Index: Measuring total harm from crime based on sentencing guidelines’, *Policing: A Journal of Policy and Practice*, 10(3): 171-83.

Sjöstedt G. and Grann, M. (2002) ‘Risk assessment: What is being predicted by actuarial prediction instruments?’ *International Journal of Forensic Mental Health,* 1(2): 179-183.

Tollenaar, N. and M. van der Heijden, P. G. (2013) ‘Which method predicts recidivism best?: A comparison of statistical, machine learning and data mining predictive models’, *Journal of Royal Statistical Society Series A: Statistics in Society,* 176(2): 565-584.

Travaini, G. V., Pacchioni, F., Bellumore, S., Bosia, M. and De Micco, F. (2022) ‘Machine learning and criminal justice: A systematic review of advanced methodology for recidivism risk prediction’, *International Journal of Environmental Research and Public Health,* 19(17): 10594.

Tully, R. J., Chou, S. and Browne, K. D. (2013) ‘A systematic review on the effectiveness of sex offender risk assessment tools in predicting sexual recidivism of adult male sex offenders,’ *Clinical Psychology Review,* 33: 287-316.

VKPP (2024). VAWG Strategic Threat and Risk Assessment. Retrieved 9th November 2024 from https://www.vkpp.org.uk/assets/Uploads/VAWG-Strategic-Threat-and-Risk-Assessment-underpinning-and-informing-the-2024-VAWG-Statement-v2.pdf.

VKPP & NPCC. (2024) Executive summary, findings and recommendations: Domestic homicides and suspected victim suicides year 3 report (2020-2023) Retrieved 15th January 2024 from https://www.vkpp.org.uk/assets/Files/Domestic-Homicides-and-Suspected-Victim-Suicides-2021-2022/Executive-Summary-Y3-Report.pdf.

Whitten, D. (2024) *'Tracking, testing & targeting sexual risk orders: A tactic to counter violence against women and girls',* [Unpublished Dissertation submitted for Master's degree in Applied Criminology and Police Management, Institute of Criminology, University of Cambridge].

Williams, J. L., Malik, A. A. and McTarnaghan, S. (2020). Gender-based violence on public transportation: A review of evidence and existing solutions. Retrieved 11th January 2025 from https://urban-links.org/tools-resources/gender-based-violence-on-public-transportation-a-review-of-evidence-and-existing-solutions/.

Appendices

Appendix 1 – Ethical Approval

UNIVERSITY OF CAMBRIDGE
Institute of Criminology

Leo Zaibert
Andreas von Hirsch Professor of Penal Theory and Ethics
Director, Centre for Penal Theory and Ethics

Eve Fellows
Institute of Criminology
University of Cambridge
Sidgwick Avenue Cambridge, CB3 9DA
efc38@cam.ac.uk

28 April 2024

Dear Eve Fellows,

I write to confirm that your research proposal entitled "Utilising Machine Learning to forecast risk of recidivism for Sexual Offences and Sexual Harassment in the British Transport Police" has been reviewed and formally approved by the Institute of Criminology's Ethics Committee.

Yours sincerely,
LZaibert

Leo Zaibert

Sidgwick Avenue
Cambridge, CB3 9DA
Assistant Tel: ++01223 335369
Direct Tel: ++01223 335368
Email: lz465@cam.ac.uk
http://www.crim.cam.ac.uk/staff/professor-leo-zaibert

Contact	Non-contact	Digitally enabled	Attempted	Other
Adult abuse of position of trust - cause / incite sexual activity with a girl 13 to 17 - not s.21 premises	Act of outraging public decency - common law	Disclose private sexual photographs and films with intent to cause distress	18 or over attempt to cause / incite a boy 13 to 15 to engage in sexual activity - no penetration	Administer substance with intent to stupefy / overpower to allow sexual activity
Adult abuse of position of trust - sexual activity with a girl 13-17 cared for in s.21 premises - SOA 2003	Cause a public nuisance	Distribute an indecent photograph / pseudo-photograph of a child	18 or over attempt to cause / incite a girl 13 to 15 to engage in sexual activity - no penetration	Adult meet a boy under 16 years of age following grooming
Adult abuse position of trust - sexual activity with a girl U.13 - not s. 21 premises - SOA 2003	Conspire to outrage public decency	Make indecent photograph / pseudo-photograph of a child	Adult attempt to engage in sexual communication with a child	Adult meet a girl under 16 years of age following grooming
Adult sexual activity with a girl under 13 family member - no penetration	Engage in sexual communication with a child	Operate equipment beneath clothing of another without consent	Adult attempt to meet a boy under 16 years of age following grooming	Aid / abet / counsel / procure another to sexually assault a female by penetration
Assault a boy under 13 by touching - SOA 2003	Exposure - SOA 2003	Possess a prohibited image of a child	Adult attempt to meet a girl under 16 years of age following grooming	Aid / abet / counsel / procure the rape a boy aged 13 / 14 / 15 years of age
Assault a female 13 and over by penetration with part of body / a thing - SOA 2003	Intentionally / recklessly cause a public nuisance	Possess an extreme pornographic image / images portraying an act which threatened life	Attempt rape of a girl under 13 - SOA 2003	Aid abet the sexual assault of a female child under 13 by touching
Assault a girl under 13 by penetration with a part of your body / a thing - SOA 2003	Observe a person doing a private act	Possess an extreme pornographic image / images portraying rape	Attempt sexual assault on a female - SOA 2003	Arrange / facilitate commission of child sex offence committed by child / young person

Assault a girl under 13 by touching - SOA 2003	Public Nuisance - indecent exposure	Possess extreme pornographic image / images portraying an act of intercourse / oral sex with a dead / alive animal	Attempt sexual assault on a male person	Arrange / facilitate the commission of a child sex offence - SOA 2003
Cause / incite a boy 13 to 15 to engage in sexual activity - offender 18 or over - penetration	Welsh - Exposure - SOA 2003	Possess extreme pornographic image portraying an act of intercourse / oral sex with a dead / alive animal	Attempt to arrange / facilitate the commission of a child sex offence	Arrange / facilitate the sexual exploitation of a child aged 13 - 17 years
Cause / incite a female child under 13 to engage in sexual activity - offender 18 or over - penetration		Possess indecent photograph / pseudo-photograph of a child	Attempt to breach a restraining order after conviction	Cause / incite a child 13 - 17 to prostitution / pornography - SOA 2003
Cause / incite a girl 13 to 15 to engage in sexual activity - offender 18 or over - penetration		Possess to show / distribute - indecent photograph / pseudo-photograph of a child	Attempt to Breach SHPO / interim SHPO / SOPO / interim SOPO / foreign travel order	Cause / incite the sexual exploitation of a child aged 13 - 17 - SOA 2003
Cause / incite a girl under 13 to engage in sexual activity - no penetration		Publish advert re - indecent photograph / pseudo-photograph of a child	Attempt to cause / incite a female child aged under 13 to engage in sexual activity - no penetration	Cause a child aged 13 to 15 to watch / look at an image of sexual activity - offender 18 or over
Cause a female 13 or over to engage in a non penetrative sexual activity - SOA 2003		Record image under clothing to observe another without consent	Attempt to cause / incite a female child under 13 to engage in sexual activity - offender 18 or over - penetration	Commit an offence with the intention of committing a relevant sexual offence - SOA 2003
Cause a female 13 or over to engage in a penetrative sexual activity - SOA 2003		Show an indecent photograph / pseudo-photograph of a child	Attempt to cause / incite a girl 13 to 15 to engage in sexual activity - offender 18 or over - penetration	Conspire to rape a woman 16 years of age or over
Cause a male 13 or over to engage in a non penetrative sexual activity - SOA 2003		Take an indecent photograph / pseudo-photograph of a child	Attempt to cause a female aged 13 or over to engage in a penetrative sexual activity	Conspire to sexually assault a female aged 13 or over - no penetration

Cause a male 13 or over to engage in a penetrative sexual activity - SOA 2003		Voyeurism - install equipment / construct / adapt a structure - SOA 2003	Attempt to cause a female aged 13 or over to engage in sexual activity - no penetration	Conspire to sexually assault a male aged 13 or over - no penetration
Engage in penetrative sexual activity with a girl 13 to 15 - offender under 18		Voyeurism - operating equipment to observe - SOA 2003	Attempt to cause a male aged 13 or over to engage in sexual activity - no penetration	Engage in sexual activity in presence of a child 13 to 15 - offender 18 or over
Gross indecency with a boy under the age of fourteen years of age		Voyeurism - install equipment / construct / adapt a structure - SOA 2003	Attempt to distribute an indecent photograph / pseudo-photograph of a child	Engage in sexual activity in presence of a child under 13 - offender 18 or over
Indecent assault on a girl under the age of 14 years		Voyeurism - operating equipment to observe - SOA 2003	Attempt to engage in a sexual activity in the presence of a child aged 13 to 15 - offender aged under 18	Kidnap / falsely imprison a person with intent to commit a relevant sexual offence
Indecent assault on boy under the age of 14 years		Voyeurism - recording a private act - SOA 2003	Attempt to observe a person doing a private act	Knowingly / recklessly trespassed on premises with intent to commit a relevant sexual offence - SOA 2003
Offender 18 or over cause / incite a girl 13 to 15 to engage in sexual activity - no penetration - SOA 2003			Attempt to possess an indecent photograph / pseudo-photograph of a child	Offender 18 or over attempt to cause a child under 13 to watch / look at an image of sexual activity
Offender 18 or over engage in non penetrative sexual activity with boy 13 to 15 - SOA 2003			Attempt to rape a boy aged 13 / 14 / 15 - SOA 2003	Offender under 18 engage in sexual activity in presence of a child under 13 - SOA 2003
Offender 18 or over engage in non penetrative sexual activity with girl 13 to 15 - SOA 2003			Attempt to rape a girl aged 13 / 14 / 15 years of age - SOA 2003	Sex offenders register - fail comply with interim notification requirements - SOA 2003

Offender 18 or over engage in penetrative sexual activity with a girl 13 to 15 - SOA 2003			Attempt to rape a man aged 16 or over - SOA 2003	Sex offenders register - fail comply with notification requirements - SOA 2003
Offender of any age cause / incite a boy under 13 to engage in sexual activity - no penetration - SOA 2003			Attempt to rape a woman 16 or over - SOA 2003	Sex offenders register - supply false information in purported compliance with a notification requirement
Offender under 18 engage in non penetrative sexual activity with a boy under 13 - SOA 2003			Attempt to sexually assault a boy under 13 by touching	Sex offenders register - supply false information in purported compliance with interim notification - SOA 2003
Rape a boy aged 13 / 14 / 15 years of age - SOA 2003			Attempt to sexually assault a girl under 13 by touching	Welsh - Engage in sexual activity in presence of a child under 13 - offender 18 or over
Rape a female under 16			Attempt to sexually assault by penetration a female aged 13 and over	
Rape a girl aged 13 / 14 / 15 - SOA 2003			Offender 18 or over attempt to engage in non penetrative sexual activity with girl 13 to 15 - SOA 2003	
Rape a girl under 13			Offender 18 or over attempt to engage in sexual activity in presence of a child 13 to 15	
Rape a male under 16				
Rape a man 16 or over - SOA 2003				
Rape a woman 16 years of age or over - SOA 2003				
Sexual activity in a public lavatory - SOA 2003				

Sexual assault on a female				
Sexual assault on a male				
Unlawful sexual intercourse with a girl under 13 years of age				
Welsh - Sexual assault on a female				