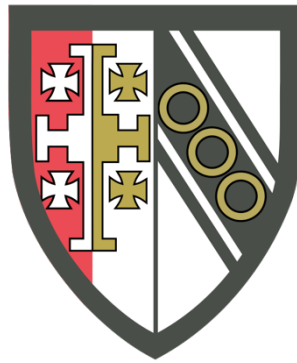


Candidate Number: Pol-2031

Max Braden Ralph

Selwyn College

Supervisor: Dr Sara Valdebenito



## **Tracking Stalking for High Harm Outcomes: A 365-Day**

### **Follow-up Analysis**

Submitted in part fulfilment of the requirements for the Master's Degree in

Applied Criminology and Police Management

February 2022

# Research Contract

## *Research Questions*

The research undertaken in completion of this thesis has sought to answer the following three research questions:

1. What is the profile of stalking cases in a police force in England and Wales that is utilising the Domestic Abuse Risk Assessment (DARA)?
2. What is the concentration of high harm for a population of stalking victims during a 365-day follow-up period?
3. What factors that are known at the time of the initial stalking report are most predictive of a victim suffering high harm in the next 12 months?

## *Research Design*

This study uses descriptive analysis techniques to quantitatively describe the profile of stalking cases within the population. A prospective longitudinal cohort dataset is analysed to identify the concentration of harm both within the victim and suspect populations, with each victim and suspect subject to a full 365-day follow-up period.

## *Data and Methodology*

The dataset for this study is comprised of the population of recorded cases of stalking that occurred within the jurisdiction of Kent Police between 1<sup>st</sup> November 2019 and 31<sup>st</sup> July 2020. The data was subject to cleansing techniques to remove variables with a high percentage of missing data and reconcile the victim and suspect datasets, using the unique person identifiers and crime reference numbers to remove any duplicate records from the dataset. The final dataset is comprised of 2006 unique victims and 1902 unique suspects.

## *Analytic Methods*

Descriptive analysis has been carried out to illustrate the breakdown of different variables within the stalking population, for example the presence of domestic abuse, the age of the victim and suspect, and the prior criminal history of the victim and suspect. Conditional probability, relative risk ratio and random forest modelling are all used to identify the likelihood of an individual becoming a “high harm” victim during the following year after a stalking report being recorded by the police, and to identify what factors are more predictive of that being true.

## *Findings*

Stalking is primarily carried out by male perpetrators against female victims and the significant majority of stalking cases are domestic abuse related. Victims and suspects tend to be younger in age, with 25-34 the most common age range of individuals within the study. Police charge very few stalking cases, with the majority of cases subject to a “no further action” outcome, including the victim declining to support police action. Most stalking cases are risk assessed by police, with a risk grading of “medium” the most frequently observed.

The research quantitatively demonstrates that stalking victims are subjected to both high levels of harm and a high count of offending. When analysing the concentration of crime during the follow-up period, a strong power few of both victims and suspects were identified. Just over 5% of victims and suspects either suffer or perpetrate 50% of total cumulative harm in the following year. 10.5% of victims go on to suffer high harm, defined in this study as a cumulative CCHI score of 500+.

Across both the relative risk ratio analysis and the random forest modelling, prior victimisation (in respect of both high crime count and high CCHI score) was the most predictive variable for a victim going on to suffer subsequent cumulative high harm during the cohort follow-up period. The suspect's offending history (again both crime count and CCHI score) was also a good predictor of future high harm. Variables that would be assumed to be predictors of high harm, such as the presence of domestic abuse, were found to have little effect on the risk of future high harm in a population of this size. Subsequent stalking was found to be correlated to significant levels of future harm for victims, particularly if the stalking had escalated in terms of severity.

### *Policy Implications of the Findings*

These findings indicate that police risk assessments carried out to assess the future risk to stalking victims can be improved by the inclusion of police information, such as prior victimisation and suspect prior count of offending, alongside the current methodology of gaining a first-hand account from the victim. There is some predictability observed in respect of which victims go on to suffer future high harm, with moderate to strong risk ratios observed across many of the variables assessed. Additionally, high harm is concentrated to a small number of both victims and suspects. By utilising the predictive factors, police agencies may be able to more precisely identify the power few victims and suspects allowing for the more effective and efficient deployment of limited resources. Finally, the conditional probability analysis and random forest modelling demonstrated that there is some promise in the prospect of forecasting future high harm in stalking cases. Further research is recommended to further develop this approach.

## **Acknowledgements**

I am incredibly grateful to the many people who have helped to make this research possible.

Thank you to Kent Police for providing me with this extraordinary opportunity, to my amazing team for the support and encouragement you have shown, and to Becky Nairn for your assistance in capturing the data.

My thanks to Dr Sara Valdebenito for your guidance and wisdom at every stage of this process.

To my daughters Lillian and Constance, this, like everything I do, is for you. I hope you can one day use this to inspire you to do something that you never thought you could.

My biggest thanks are reserved for my wife Lauren. This whole experience would not have been possible without your enduring support, patience, and belief.

# Contents

<b>Research Contract .....</b>	<b>2</b>
<i>Research Questions .....</i>	2
<i>Research Design .....</i>	2
<i>Data and Methodology .....</i>	2
<i>Analytic Methods .....</i>	3
<i>Findings.....</i>	3
<i>Policy Implications of the Findings .....</i>	4
<b>Acknowledgements .....</b>	<b>5</b>
<b>Contents .....</b>	<b>6</b>
<b>List of Tables .....</b>	<b>9</b>
<b>List of Figures .....</b>	<b>10</b>
<b>Introduction .....</b>	<b>12</b>
<i>Purpose of This Research .....</i>	13
<b>Literature Review .....</b>	<b>15</b>
<i>Introduction .....</i>	15
<i>Defining Stalking .....</i>	16
<i>Repeat Victimisation .....</i>	17

<i>Impact on Victims</i> .....	18
<i>Typology of Stalkers</i> .....	19
<i>The Link Between Stalking and Domestic Abuse</i> .....	20
<i>The Link Between Stalking and Violence</i> .....	21
<i>The Link Between Stalking and Violence Against Women and Girls</i> .....	22
<i>Limitations in the Literature</i> .....	23
<b>Methodology</b> .....	<b>25</b>
<i>Introduction</i> .....	25
<i>Operational Definitions</i> .....	26
<i>Data</i> .....	28
<i>Research Questions</i> .....	34
<i>Random Forest</i> .....	39
<b>Results</b> .....	<b>42</b>
<i>Introduction</i> .....	42
<i>Research Question 1</i> .....	42
<i>Research Question 2</i> .....	51
<i>Research Question 3</i> .....	63
<b>Discussion</b> .....	<b>72</b>
<i>Introduction</i> .....	72
<i>Theoretical Implications</i> .....	72
<i>Policy Implications</i> .....	75

*Research Implications*..... 78

*Limitations of the Study*..... 81

*Strengths of the Study*..... 82

**Conclusions**..... **84**

**References** ..... **88**

**Appendices** ..... **94**



## List of Tables

<i>Table 1: Total recorded stalking crimes in Kent Police by calendar year.....</i>	<i>29</i>
<i>Table 2: Initially Captured Variables and % of Missing Data .....</i>	<i>30</i>
<i>Table 3: 27 independent variables used for relative risk ratio calculation .....</i>	<i>38</i>
<i>Table 4: Count of subsequent victimisations during 365-day follow-up sorted in descending order by number of recorded crimes.....</i>	<i>52</i>
<i>Table 5: Count of subsequent records of being recorded as a suspect during 365-day follow-up sorted in descending order by number of recorded crimes.....</i>	<i>53</i>
<i>Table 6: Average harm, standard deviation and median for victims categorised under each risk grading.....</i>	<i>59</i>
<i>Table 7: Average subsequent harm and count, standard deviation and median for victims based on their repeat victimisation status.....</i>	<i>61</i>
<i>Table 8: Future victim high harm based on the outcome code assigned to the initial stalking investigation .....</i>	<i>62</i>

## List of Figures

<i>Figure 1: Illustration of rolling 365-day follow-up for each stalking victim .....</i>	<i>25</i>
<i>Figure 2: Proportion of cases recorded as each of the three notifiable stalking offences.....</i>	<i>43</i>
<i>Figure 3: Percentage of stalking cases within the population which contained a domestic abuse marker .....</i>	<i>44</i>
<i>Figure 4: Percentage of female and male victims and suspects observed in the study.....</i>	<i>45</i>
<i>Figure 5: Proportion of victims and suspects in each age group within the study population</i>	<i>46</i>
<i>Figure 6: Frequency with which each of the police recorded outcomes appear within the population.....</i>	<i>48</i>
<i>Figure 7: Number of victims and suspects in the cohort who have previous victimisation or perpetration records of stalking on the Athena database .....</i>	<i>49</i>
<i>Figure 8: Proportion of risk grading categories assigned to the cohort victims.....</i>	<i>50</i>
<i>Figure 9: Power curve to illustrate the concentration victims who suffer harm within the 365-day follow-up period .....</i>	<i>54</i>
<i>Figure 10: Power curve to illustrate the concentration of suspects who perpetrate harm within the 365-day follow-up period.....</i>	<i>55</i>
<i>Figure 11: Quantity of victims subjected to each level of harm during the follow-up period.</i>	<i>56</i>
<i>Figure 12: Quantity of suspects who caused each level of harm during the follow-up period</i>	<i>57</i>
<i>Figure 13: Conditional probability of becoming a high harm victim during the follow-up period given that the original stalking case was assigned a risk grading of standard, medium, high, or no grading.....</i>	<i>58</i>

*Figure 14: Forest plot describing the relative risk ratio of each of the twenty-seven independent variables for a victim suffering a cumulative CCHI score of 500+ during the 365-day follow-up ..... 64*

*Figure 15: Random Forest Model Based on Testing Data..... 68*

*Figure 16: Predictor variable importance ranked based on the Mean Decrease in Gini Value ..... 69*

*Figure 17: Partial Dependency Plot for Historic CCHI Total..... 70*

*Figure 18: Partial Dependency Plot for Victim Age..... 70*

## Introduction

Stalking is a crime that sees a suspect persistently and repeatedly target their victim through behaviours such as following, spying and forcing contact, and can cause significant levels of harm and distress (Stefanska et al, 2021; CPS, 2018; Taylor-Dunn et al, 2018). The term 'stalking' has embedded itself into common public discourse, from throwaway conversational comments related to "Facebook stalking" through to an increased public awareness of stalking behaviours and the harm it can cause at the most severe end of the offending. Google Trends data shows that the term "stalking" is searched four times as frequently in the United Kingdom in 2021 as it was in 2004 when the data was first collected (Google, 2022), indicating an increased public interest in the topic.

In the 1990's, nations such as the United States of America and England and Wales led the way internationally in recognising stalking within legislation introduced to tackle harassment offending. In England and Wales, this was done through the ascension of the Protection from Harassment Act 1997, and stalking has since been subject to increased interest within academia. The same focus and attention have not necessarily been observed in law enforcement and the wider criminal justice system. Despite updates to the legislation through the introduction of the Protection of Freedoms Act 2012, which saw stalking recognised as a criminal offence the first time in England and Wales (having previously been considered part of the offence of harassment), independent inspections have identified that there are continued issues in respect of the identification and prosecution of stalking cases. In their joint inspection, Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS) and Her Majesty's Crown Prosecution Service Inspectorate (HMCPSI) commented "*We found that stalking was misunderstood by the police and the CPS. As a*

*result, it often went unrecognised. The police sometimes mis-recorded stalking offences, or worse, did not record them at all. Prosecutors on occasion missed opportunities to charge stalking offences, instead preferring other offences, particularly harassment.”* A series of recommendations were made to improve the service that the police and courts deliver in respect of stalking (HMICFRS, 2017).

A combination of factors including changes to the crime recording guidelines for stalking offences (Home Office, 2021a) and the introduction of the Domestic Abuse Risk Assessment (DARA), which contains questions that explicitly seek to identify stalking behaviours, have seen the number of stalking cases recorded by Kent Police, a police force in the South-East of England, rise 605% between 2018 (n = 676) and 2020 (n = 4091). Given that research into stalking has demonstrated that it can be a key risk factor in the highest harm crimes such as domestic homicide (Richards, 2003) and often acts as a precursor to other violent acts (McEwan et al, 2009), such a significantly increased level of cases is not something that can be ignored by law enforcement. With a limited number of police resources stretched across many competing priorities, it is essential that police forces seek to gain a better understanding of the profile of stalking cases, including the levels of harm suffered by stalking victims.

### *Purpose of This Research*

This study uses a combination of descriptive analysis techniques and risk factor identification to answer the following three research questions to build on the existing body of research into stalking:

1. *What is the profile of stalking cases in a police force in England and Wales that is utilising the Domestic Abuse Risk Assessment (DARA)?*

2. *What is the concentration of high harm for a population of stalking victims during a 365-day follow-up period?*
3. *What factors that are known at the time of the initial stalking report are most predictive of a victim suffering high harm in the next 12 months?*

Much of the existing research into stalking examines the harm that is caused to victims, but this is generally based on qualitative and/or medical data reported by small numbers of self-selected victims (Taylor et al, 2019; Jason et al, 1984). This study uses a population of victims taken from police recorded data and uses the Cambridge Crime Harm Index (CCHI) (Sherman, Neyroud & Neyroud, 2016) to quantify the future harm suffered by victims of stalking. This quantitative approach will provide a consistent measure of harm for which it is possible to identify which victims suffer most and which factors may be most predictive of future risk.

The thesis is set out across four chapters. The first chapter will examine the existing academic literature in respect of stalking, outlining the research that has been carried out across the world and what that tells us about stalkers and their victims. The second chapter outlines the research questions in more detail and the methods used to answer each of them. This is followed by the results which are presented as a descriptive analysis of the findings relevant to each research question, concluding with the findings of an attempt to create a random forest forecasting model for predicting which victims are most likely to suffer high harm during the follow-up period. The discussion chapter outlines the theoretical, policy and research implications for this research, as well as the strengths and limitations of the study. The thesis draws to a close by outlining the conclusions of the study.

# Literature Review

## *Introduction*

Early research into stalking was often centred on high profile cases involving celebrity victims, linking the offender's behaviour to poor mental health and mental disorders (Rosenfeld, 2004). To this day, stories about high profile individuals being stalked are never too far from the front pages or news websites, such as the stalking campaign perpetrated against BBC presenter Emily Maitliss (BBC News, 2021). However, as time passed and the body of research developed, it became clear that stalking had a greater presence in wider society, rather than being something that was merely restricted to a fascination with celebrity, and the research focus shifted to victim-offender relationships (Sheridan et al, 2001).

This chapter will examine the wide-range of studies that have taken place since the shift in stalking research to exploring the victim-offender relationships. The chapter will firstly focus on the definition of stalking, of which there is no single consensus. Stalking is a repeating crime by its very nature, and as such much of the literature relating to repeat victimisation has relevance and will be examined in this chapter. The chapter will then move on to outlining the impact that stalking has on victims, from the physical and mental strain that it causes through to the lifestyle changes required for victims of the most serious stalking behaviour to safeguard themselves. The literature outlines stalker typologies and the various motivations for a perpetrator to stalk a victim which will be examined. The chapter then looks at the links between stalking and domestic abuse, stalking and violence, and the overall impact of stalking in the contemporary focus on violence against women and girls. Finally, the chapter outlines limitations in the current literature that need addressing.

## *Defining Stalking*

Western nations have steadily introduced anti-stalking legislation during the past 30 years, led by the state of California who formally recognised stalking as a crime in 1990 (Gregson, 1998; Spitzberg & Cupach, 2007). Despite the shift towards criminalising stalking, there is still no single definition agreed within academia and law enforcement (Sheridan et al, 2001). In England and Wales, the offence of stalking is contained within The Protection of Freedoms Act 2012, and wider research similarly recognises that most statutes require a course of conduct by the offender for an offence to constitute stalking (Purcell et al, 2004).

Commonalities between definitions typically include the words “fixated” and “obsessive” to describe the offender’s behaviour. Monckton-Smith et al (2017) defined stalking as “any fixated and obsessive attention designed to make the victim fearful or distressed”. Stalking advocacy services Protection Against Stalking and the Suzy Lamplugh Trust assist victims of stalking across England and Wales and share a definition similar to Monckton-Smith et al’s, although more detailed in respect of both the offenders’ behaviour and the impact on the victim, of “a pattern of fixated and obsessive behaviour which is repeated, persistent, intrusive and causes fear of violence or engenders alarm and distress in the victim”. From a law enforcement perspective, in England and Wales the College of Policing has developed a mnemonic “FOUR”, standing for “fixated”, “obsessive”, “unwanted” and “repeated” (College of Policing, 2020). This mnemonic assists law enforcement professionals in identifying stalking behaviours, which may be especially important if the course of conduct is not immediately apparent in what is being reported. The Home Office Counting Rules (HOCR) (Home Office, 2021a), which provides guidance to police forces in England and Wales on how to record reports of crimes, does not define



stalking, but instead provides examples of the types of behaviours that constitute stalking and how they should be interpreted to record notifiable offences. These behaviours include spying, following and social media hacking (Home Office, 2021a).

There can be discrepancies between what the public, and more specifically victims, see as stalking compared with what academia, law enforcement and third sector agencies such as stalking advocacy services know to be stalking behaviours (McKeon et al, 2014). The word 'stalked' can be used lightly by the public, leading to minimisation of the behaviours and the effects that it has on victims (Boehnlein et al, 2020). There is the risk that if the behaviours are minimised by the victims or their support network and the severity is not appropriately identified, then victims may not be able to access the support that they need (McKeon et al, 2014).

### *Repeat Victimisation*

Stalking is, by its very nature, a repeat offence. The course of conduct required to evidence that stalking is occurring means that a victim has already been subjected to a minimum of two events prior to the initial stalking crime being recorded and stalking and harassment are largely defined by their repetitive nature (Farrell & Pease, 2014).

The literature in respect of repeat victimisation is well established, and it can be expected that most crime occurs against targets that have already been victimised (Grove et al, 2012), with 1% of targets accounting for 59% of all personal crime, including violence (Pease, 1998). Chronic victimisation can occur against victims of certain crime types, stalking included, where the relationship between the victim and offender is traditionally well-established, meaning that the offender has already assessed the risk and reward of carrying out their crime (Farrell & Pease, 2014). Research shows that crimes occurring against repeat

victims tends to be underreported, with this most significantly occurring when the victim is a “supertarget” who suffers a disproportionate amount of harm (Farrell & Pease, 2014). In the context of stalking, this appears to be the case given the inherent nature of repeat victimisation and the backdrop of underreporting of the crime to police agencies (Brady & Nobles, 2017).

Research indicates that stalking tends to occur over a prolonged period, with one meta-analysis of 175 studies providing an average of 22 months (Gill & Brockman, 1996; Turmanis & Brown, 2006; Spitzberg & Cupach, 2007), and 77% of victims will suffer up to one hundred incidents prior to reporting the stalking to police (Paladin, 2017). The stalking can often go unreported to police for some time, particularly if the suspect is an ex-intimate partner (Reyns & Englebrecht, 2010; Jordan et al, 2007). There is also some evidence of escalation, with 75% of victims indicating that the behaviour of the stalker had worsened over time (Sheridan et al, 2001b).

Once it has been reported, research shows that the police are ineffective at identifying stalking, often treating reports as isolated cases which leaves the victims vulnerable to further harm (HMICFRS, 2017; Taylor-Dunn et al, 2021).

### *Impact on Victims*

Research into targeted stalking related violence is extensive and identifies a strong correlation between stalking and significant effects on victims lives and health (Fein et al, 1995; Taylor-Dunn et al, 2021). Common health impacts observed include psychological harms such as post-traumatic stress disorder, depression, and a feeling of helplessness (Reyns & Englebrecht, 2010; Taylor et al, 2019). The fear felt by victims of stalking can result in impactive lifestyle changes to get away from the stalker and reduce the harm, including

moving addresses or changing phone numbers (Jason et al, 1984), changing their job or even stopping working altogether (Pathé & Mullen, 1997; Baum et al, 2009). On top of the health and lifestyle impacts that a victim will suffer because of being stalked, research shows that victims also suffer financially in respect of lost productivity and wages (DuMonthier et al, 2017). In one study, 28% of stalking victims reported over \$1000 in lost income, with 5% reporting that they had lost in excess of \$5000 (Baum et al, 2009).

### *Typology of Stalkers*

Research into the motivations of stalkers has moved away from an obsession over celebrity or high-profile figures and has, more recently, focused on the victim-offender relationships within wider society. Various studies have sought to classify stalkers, whether that be by their mental state or experience of psychosis (Geberth, 1992; Kienlen et al, 1997), or the motivations of the stalker based on their relationship (or lack thereof) with the victim (Mullen, Pathé and Purcell, 2000).

Further research has sought to categorise stalkers in a way that will be most useful to law enforcement agencies in respect of profiling and tackling stalkers, rather than being focused in the world of mental health and psychiatric assessment (Boon & Sheridan, 2001). Boon & Sheridan's research saw the creation of four stalker categories: Ex-partner harassment/stalking (the most common observed in the study); infatuation harassment; delusional fixation stalking; and sadistic stalking. With the intention that these categorisations should help law enforcement, each of them contained suggestions for case management along with the characteristics that an offender in each typology would demonstrate, allowing for law enforcement professionals to easily identify which type of stalker they were dealing with. The limitation with this research, as is the case with much of

the research carried out on stalking as a topic, is the low sample size. These findings were based on a sample of 124 cases collected from members of two prominent UK-based self-help groups which were set up to support victims of stalking and that may not be enough to overcome any categories that were over-represented in the sample, leading to conclusions that certain characteristics are more prevalent than they perhaps would be in the general population.

### *The Link Between Stalking and Domestic Abuse*

Research shows that most stalking occurs between individuals who are in some way acquainted with one another (Cupach & Spitzberg, 2004). Stalking and domestic abuse are highly correlated (Norris et al, 2011) and in addition to Boon & Sheridan's (2001) study into stalker classifications, research has shown that ex-intimate partner stalking is recognised as being the most common form of stalking (Tjaden & Thoennes, 1998). In fact, Spitzberg and Cupach (2007) argued that ex-intimate partner stalking accounted for half of all stalking cases in their sample. Furthermore, there is some evidence to suggest that ex-intimate and intimate partner stalking can be more harmful to a victim and be longer lasting than other forms of stalking (Tjaden & Thoennes, 2000), with ex-intimate stalkers exhibiting more dangerous stalking behaviours than non-ex-intimates (Palarea et al, 1999), and victims of ex-intimate stalking being at a heightened risk of violence than other classifications of stalking victims (Mullen et al, 2000).

More recent research has demonstrated a correlation between coercive/controlling behaviour and stalking. The introduction of the Domestic Abuse Risk Assessment (DARA) in several pilot forces in England and Wales, which was primarily introduced to enable more effective identification of controlling and coercive behaviours, has led to an increase in

disclosure of stalking events by victims (Wire & Myhill, 2018), with two of the DARA questions explicitly focused on stalking (Appendix 1). Positive responses from the victim result in the creation of a stalking crime report, as per the Home Office Counting Rules guidance (Home Office, 2021a). Studies into the links between controlling and coercive behaviour and stalking have found that stalking acts as an extension of the perpetrators' control over the victim once the intimate relationship has finished (Taylor-Dunn et al, 2021).

### *The Link Between Stalking and Violence*

Stalking is itself classified as a violent crime in England and Wales, and even in the absence of any physical violence, research has demonstrated that victims can suffer severe mental trauma because of stalking (Reyns & Englebrecht, 2010; Taylor et al, 2019). It has been observed that it can be difficult to estimate the extent of stalking related violence, with much of the literature focused on high harm offences more serious than stalking, which risks underestimating the rate of severe violence due to cases potentially not being recognised as stalking (Rosenfeld, 2004). Despite this, many studies demonstrate a correlation between stalking and (physical) violence, with the utterance of threats often acting as a precursor to violence (Purcell, Pathé & Mullen, 2001; Sheridan et al, 2001b; Brewster, 2000). Furthermore, a history of violence on the part of the stalker is the strongest predictor of stalking related assault (Mullen, 1994).

When looking at the most serious violent offences involving the murder of female victims, stalking behaviours were present in 94% of the 358 cases of femicide examined by Monckton-Smith et al, 2017. Further research by Monckton-Smith led to the definition of an eight-stage homicide timeline. Stalking is noted as key offender behaviour in five of the

eight documented stages, further indicating the strong correlation between stalking and the most serious violence (Monckton-Smith, 2020).

Some studies have sought to identify a link between certain demographic variables in both victims and suspects and the prevalence of violence. Research in this area has generated mixed findings, with some studies demonstrating a positive correlation between variables such as younger age and lower educational achievement as predictors of violence (Rosenfeld & Harmon, 2002), whilst others fail to demonstrate a relationship (Brewster, 2000; Schwartz-Watts & Morgan, 1998).

### *The Link Between Stalking and Violence Against Women and Girls*

Potentially anyone can become a victim of stalking, however research shows that stalking is a crime that is predominantly perpetrated against women, with females being three times more likely than males to report being a victim of stalking (Senkans, McEwan & Ogloff, 2017). One meta-analysis showed that the average proportion of female victims across 103 studies examined was 75% (Spitzberg, 2002), however, it should be noted that much of the research into stalking is gender biased with wholly female victim samples being obtained and studied (Sheridan et al, 2000; Sheridan et al, 2001a).

In the wake of several high-profile tragic cases of violence against women and girls in the United Kingdom, there has been a renewed spotlight on offences which predominantly impact on women and girls such as domestic abuse, sexual offences and female genital mutilation. The UK Government launched the Tackling Violence Against Women and Girls strategy in July 2021 and stalking features prominently within the strategy (Home Office, 2021b). The strategy outlines the discrepancy in the number of police recorded stalking crimes (81,995 in the calendar year of 2020) against the projected number of stalking crimes

based on the Crime Survey of England and Wales (1.3 million victims, with 892,000 of these being female).

### *Limitations in the Literature*

In preparing to undertake this study, the primary limitation of the stalking literature that has been observed is the lack of representativeness of studied samples that sits behind much of the research. Many of the studies into stalking are made up of small, non-representative samples of often self-selected participants. Victim research tends to use surveys as the primary methodological technique, often based within a university or college setting (Reyns et al, 2011; Patton et al, 2010). The self-selecting nature of surveys, particularly in the usual volume of responses observed, means that respondents are not likely to be representative of an entire population which creates issues in respect of external validity. Suspect research has often focused on individuals from a clinical perspective or within a clinical setting, for example, psychiatric assessment centres (Geberth, 1992; Mullen et al, 2000). There is a risk that research of this type overestimates the links between stalking behaviour and psychiatric issues, again presenting problems with external validity to the wider population of stalkers.

There is a positive to the self-selecting nature of much of the stalking research. It is widely accepted that stalking is under-reported to law enforcement, and even when it is reported police agencies are not always effective at joining the dots and treating cases as a coherent chain of linked events, often recording and investigating cases in isolation. By removing the third party of the police agency and seeking to understand the views and experiences of victims directly, it means that researchers may have been uncovering more examples of crimes that would not have been either reported to, or recorded by, the police.

Furthermore, the lack of consistently applied definitions across studies creates a limitation in the interpretation and generalisability of the findings. The term “stalking” is interpreted differently, with varying thresholds of what behaviours constitute stalking.

This study seeks to address many of the current limitations observed in stalking research. Firstly, by using the entire population of stalking cases recorded in Kent Police within the specified time frame, a sample has been obtained that is far larger than those traditionally observed in stalking research. The study does not rely on unique or bespoke definitions, and instead utilises measures that are consistent across all police forces in England and Wales which helps to improve the external validity of the study. Finally, whilst it is accepted that the population may not be fully representative of all stalking cases that occurred within the policing jurisdiction during the specified time frame due to the known under-reporting of stalking, with all recorded cases being captured there is no element of self-selection bias in either the victims or suspects observed.

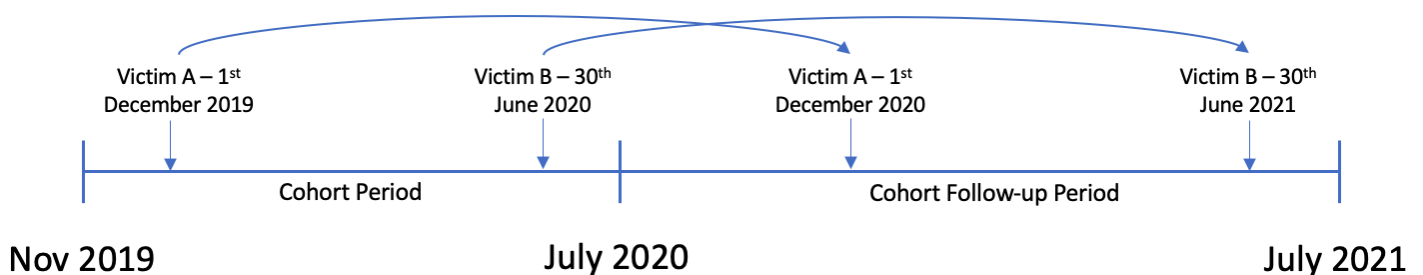


# Methodology

## Introduction

This thesis carries out a longitudinal study (Caruana et al, 2015) that allows for the victimisation of a cohort of stalking victims (n = 2006) to be observed in the year following the initial stalking offence. As demonstrated in Figure 1, all stalking cases recorded by Kent Police during the recruitment cohort period of 1<sup>st</sup> November 2019 and 31<sup>st</sup> July 2020 have been captured for this study. Each unique victim within that cohort has been tracked for 365-days following the date of the initial stalking offence being reported, with the follow-up period extending to the 31<sup>st</sup> July 2021 to ensure that all victims could be tracked for the same amount of time. This study measured the prevalence of future victimisation and severity of harm subjected to during the follow-up period using the Cambridge Crime Harm Index (CCHI) (Sherman, Neyroud & Neyroud, 2016).

**Figure 1: Illustration of rolling 365-day follow-up for each stalking victim**



The longitudinal design of this study has many advantages. It allows a consistent cohort of victims to be tracked over the study period, establishing a clearly defined sequence of events from the initial stalking case through to the end of the follow-up period. Additionally, by examining the cohort prospectively using quantitative data, there is no

recall bias on the part of the victim in having to recount the events as per their memory of the events (Caruana et al., 2015).

The study will attempt to establish whether any factors known at the time of the original stalking offence being reported are correlated to subsequent victim high harm following that stalking event. Critical variables include prior victimisation both in respect of crime count and crime harm, perpetrator prior offending in respect of crime count and crime harm, the presence of a domestic abuse relationship between the victim and perpetrator, and several demographic factors including age and gender.

### *Operational Definitions*

#### *Stalking*

For the purposes of this thesis, a stalking event has been defined as an incident reported to Kent Police during the cohort recruitment period which has met the threshold, as per Kent Police's interpretation of the Home Office Counting Rules (Home Office, 2021a), to be recorded as one of the following three notifiable offences:

1. Harassment – Pursue a course of conduct in breach of Sec 1 (1) which amounts to stalking
2. Harassment – Stalking involving fear of violence
3. Harassment – Stalking involving serious alarm/distress

#### *High Harm*

The Cambridge Crime Harm Index (CCHI) (Sherman, Neyroud & Neyroud, 2016) has been used as a consistent measure of crime harm throughout this research, both for historic harm caused to victims and by suspects, as well as the future harm identified

during the 365-day cohort follow-up period. The CCHI is founded on the basis that all crimes are not created equally and that some offences cause a greater degree of harm to those it affects than others, as well as a disproportionate level of police resources to detect (Sherman, Neyroud & Neyroud, 2016). Despite being focused on stalking, this study has measured *all* crime harm for three reasons. Firstly, stalking is by its very definition a “repeat” crime. As such, it was hypothesised that the identified victims are highly likely to have been victims prior to the stalking event, post the stalking event, or indeed both. Secondly, as outlined above, stalking can sometimes present as if it were an isolated case (Taylor-Dunn et al, 2018) and as such there is potential that Kent Police may not record subsequent offences as part of the same stalking course of conduct. By measuring all crime harm, this has ensured that any offences whether related or unrelated are captured in order that the true effect of the harm against the victim can be identified. Thirdly and finally, the close link between stalking and domestic abuse (Norris et al, 2011) meant that it is likely that within many of the victim-offender dyads present in the study, other linked offences were likely to occur that may not be recorded as part of the stalking course of conduct.

This thesis has a particular focus on high harm. The CCHI does not provide a clear definition of what score constitutes high harm, therefore allowing subjectivity in its practical and academic use. In this research, high harm has been defined as a CCHI score of 500 or more, which is the equivalent of 500 or more imprisonable days. High harm during the follow-up period is considered by taking the cumulative total of all harmful events during that period, however examples of single offences that exceed a harm score of 500 include grievous bodily harm, kidnapping and rape. This score has primarily been chosen due to the concentration of both victims and suspects who meet that criterion,

representing a “power few” (Sherman, 2007) of victims within the population that account for the majority of subsequent harm suffered in the follow-up period.

### *Risk Assessment*

Where the thesis refers to findings concerning a risk assessment, this is specifically referring to any of the three current risk assessments completed by Kent Police for stalking offences, dependent on whether they are domestic abuse related or not. In October 2019, Kent Police moved to a new domestic abuse risk identifier, the Domestic Abuse Risk Assessment (DARA) tool (Appendix 1). Kent Police were one of the pilot forces chosen to test the tool, however a subsequent review by the College of Policing has found that the DARA tool led to an increase in the disclosure of coercive and control and stalking offences (Wire & Myhill, 2018). DARA replaced the Domestic Abuse, Stalking and Harassment and Honour-Based Violence (DASH) (Richards, 2009) (Appendix 2) risk model as the predominant domestic abuse risk assessment tool in Kent Police. DASH is a checklist of 27 questions that was used by frontline officers during the initial response to a call involving a domestic abuse incident and still had some use by the police force during the stalking reporting period for this study, as well as its usage by partner agencies. The final risk assessment is Stalking-DASH (S-DASH) (Appendix 3), a subset of eleven questions specifically related to stalking that should be asked of all stalking victims whether the case is domestic abuse related or otherwise.

### *Data*

#### *Data Management*

A Kent Police Analyst was tasked with creating a new query within the Business Objects universe to extract all crimes recorded as one of the three notifiable stalking

offences<sup>1</sup> between 1<sup>st</sup> November 2019 and 31<sup>st</sup> July 2020 from the Athena crime recording system. Athena is the first policing software of its kind to integrate recorded crime data from nine separate police forces, however the records used for this research were limited to those recorded within the Kent jurisdiction only. This extraction accounted for the entire population of stalking cases recorded by Kent Police during the stated time period. This date range was specifically chosen to ensure that only cases where the risk assessment would have been completed using the new DARA tool would be included, therefore creating a more accurate profile of the current stalking in Kent Police. The significance of the introduction of DARA relates to the hypothesis that the use of DARA is a critical factor in the increase of stalking crimes recorded in Kent Police. Between the calendar years of 2018 and 2020, Kent Police has seen a 605% increase in recorded stalking cases (Table 1), with most of that increase occurring between 2019 and 2020, coinciding with the introduction of DARA. The upper limit of 31<sup>st</sup> July 2020 was selected as this was the latest date on which all identified victims could receive a full 365-day follow-up.

**Table 1: Total recorded stalking crimes in Kent Police by calendar year**

	<b>2018</b>	<b>2019</b>	<b>2020</b>
<b>Recorded cases of stalking</b>	676	916	4091

The data was initially captured across multiple Excel spreadsheets due to constraints on the number of unique identification points of data that Business Objects can extract in a single query. The first spreadsheet contained all the cohort data for both victims and

---

<sup>1</sup> Harassment with stalking, Stalking involving fear of violence, Stalking involving serious alarm/distress

suspects for those offences recorded during the study period. It had captured the following criteria for each recorded crime (Table 2):

**Table 2: Initially Captured Variables and % of Missing Data**

<b>Data Heading</b>	<b>% Missing Data</b>
Crime reference number	0%
Full offence title and related Home Office classification codes	0%
Cambridge Crime Harm Index (CCHI) score	0%
District in which the crime occurred	0%
Closure outcome code assigned to the crime (e.g. charged)	0.008%
Text explanation of the assigned outcome code	0.008%
Crime created date	0%
Reported date	0%
Upper and lower committed dates	0.006%
Domestic abuse (true / false)	0%
Unique victim identification number (identifier)	0%
Victim date of birth	0.003%
Victim age at time of the crime	0.009%
Victim gender	0.005%
Victim self-defined ethnicity	16.5%
Offence location	62%

As Table 2 demonstrates, except for self-reported ethnicity and location, the data extracted was complete with greater than 99% of data present across all other headings. Due to the quantity of missing ethnicity data, coupled with the fact that 50% of self-reported ethnicities captured were “white” which left just 52 victims across the other 12 ethnicity categories, the decision was taken to exclude ethnicity data from the analysis due to its low statistical representativeness.

An additional tab within that spreadsheet captured the location type data for each of the cases. This data would typically identify whether the stalking event occurred at the victim’s home address, in a public place, or online. It was anticipated that there may be

some problems in obtaining location data, mainly because the location field in Athena does not require mandatory completion and is prone to human error and subjectivity in how it is recorded. This proved to be the case with 62% of data missing. A decision was therefore taken not to include location data in this study. The second spreadsheet contained both the historic victimisation data and subsequent recorded victimisation data for each of the study victims during the 365-day follow-up. The final spreadsheet contained the equivalent data for each of the study suspects. Each of these spreadsheets captured the same criteria as outlined in the bullets above. At this point, all data was in its raw form and no cleansing had taken place.

On starting the cleansing process and inspecting the data more closely, it was identified that the data in the first spreadsheet had been captured based on committed dates as opposed to reported. A new set of data was therefore extracted based on reported dates. Further inspection of the new data set identified that there were issues with the data from the initial stalking period, with 181 person iteration numbers appearing two or more times in the dataset, where there should have been no duplicate records found. The initial assumption was that the data extraction had simply spilled over into the follow-up period and included further stalking events that had occurred during the study time. However, further inspection found that a total of 112 records were present that were suspects of stalking, not victims. All 181 records were subject to a manual review, cross referring the Excel spreadsheet with the Athena crime recording system to ensure that victims and suspects had been captured appropriately. Where suspects or subsequent crimes for the victim were found, those records were transferred to the appropriate data set.

Inspection of the subsequent victimisation and suspect data identified similar issues in respect of dates. In total, 613 subsequent victimisations were found to have occurred outside of the 365-day follow-up period and were removed, whilst 895 were found to have occurred prior to the initial stalking date. These records were transferred to the “historic” crimes tab.

A final check of the initial stalking victimisation data was completed to ensure that all suspect records had been cleansed. A VLOOKUP was created, searching on the person iteration numbers against the suspect data for the initial reporting period, with a new column created headed “Suspect same as victim?”. This used an IF formula to show as “Yes” if the same person iteration number appeared in both the victim and suspect spreadsheets against the same crime reference number. A total of 25 further suspect cases were identified and following a manual review of Athena, all 25 cases were removed from the victim data. Finally, manual checks were carried out on crime reference numbers that were duplicated in the data set. One of these records was removed as the crime report showed two suspects, one of which had been logged as a victim in the research data. The other 14 were all legitimately duplicated as they contained more than one victim. The final total of unique victims remaining in the study was 2006.

A similar process was conducted on the suspect data. 220 duplicate crime references were identified and subject to a manual review. A total of 100 records were moved to the subsequent suspect data set, with the other records remaining due to having multiple suspects listed on the same crime report. This left a total of 1902 unique suspects, with 42 of these suspects appearing more than once in the data set due to their targeting of multiple victims. Finally, the subsequent suspect data was interrogated. A total of 278



records were identified where the suspect listed had no link to the initial stalking data and as such were removed from the data set. Inspection of the reported dates identified a further 1129 records that were moved to the historic suspect dataset.

Once all incorrect records had been removed from the victim and suspect data, the separate spreadsheets were merged into one master document using the crime reference number for the stalking crime report as the common identifier to match the datasets.

During the analysis, the following further columns were added to the merged sheet:

- Days from reported date to crime report created date
- Length of offence (in days) before reporting (i.e. difference in time between the lower committed date and the reporting date)
- Victim subsequent crime harm index score
- Victim subsequent crime count
- Victim historic crime harm index score
- Victim historic crime count
- Suspect of crime
- Suspect subsequent crime harm index score
- Suspect subsequent crime count
- Suspect historic crime harm index score
- Suspect historic crime count
- Suspect prior stalking (Yes/No)
- Suspect follow-up stalking (during the 365-day follow-up period)
- Risk grading (based on S-DASH)
- Victim prior stalking

- Victim follow-up stalking
- Victim follow-up stalking Home Office sub-code (to identify the stalking offence recorded in the follow-up period)
- Stalking escalated? (comparing the Home Office sub-codes of the initial stalking and the subsequent stalking)

With the data mostly complete, no records have been removed from the study. “Location” and “ethnicity” are the only variables not included in the analysis, leaving a final sample size of 2006 stalking cases.

### *Research Questions*

This study seeks to answer three principal research questions with the analytical methods used to do so set out below.

1. *What is the profile of stalking cases in a police force in England and Wales that is utilising the Domestic Abuse Risk Assessment (DARA)?*

Prior to the introduction of DARA into a small number of pilot forces in England and Wales, the single risk assessment framework for cases involving domestic abuse was the Domestic Abuse, Stalking and Harassment and Honour-Based Violence (DASH) checklist. A review in three forces identified that DASH was inconsistently applied by attending officers, and that the checklist did not adequately identify patterns of behaviour that amount to coercion and control by a suspect (Robinson et al., 2016). As a result, the College of Policing, in liaison with survivors of domestic abuse, charities, police practitioners and academics, developed the DARA to improve the quality and consistency of risk assessments completed by frontline officers in relation to domestic abuse (Wire & Myhill, 2018). The questions were

primarily written to help officers uncover previously hidden offences such as coercive and controlling behaviour.

Two of the questions are directly related to stalking<sup>2</sup>, with a positive response triggering the completion of a S-DASH risk assessment. Given the overt nature of the stalking questions within the DARA, it is therefore hypothesised that the significant increase in stalking cases recorded in Kent Police is, at least in part, down to the use of the new risk assessment framework. With 87% of cases in this study being domestic abuse related, this appears to be a highly plausible explanation. With such a large increase in cases, it is probable that many stalking offences are now being reported to Kent Police that would never have come to police attention previously, where it has been previously identified that stalking is a crime that often goes unreported to police (Jerath et al, 2020). There is potential that this large influx in reported crimes has changed the profile of stalking, leaving many of the assumptions of law enforcement in respect of stalking outdated.

Several descriptive techniques, including mean and standard deviation, have been used to analyse the cohort of the 2006 unique victims and 1902 unique suspects identified during the initial stalking cases to identify the profile of stalking cases in a police force using the DARA. This information is vital if police agencies are to effectively safeguard victims and tackle stalkers.

*2. What is the concentration of high harm for a population of stalking victims during a 365-day follow-up period?*

---

<sup>2</sup> DARA Q.5 “How often does [...] make you account for where you have been, or monitor your phone, email or social media to check up on you?”  
DARA Q.6 “How often does [...] follow or stalk you, or try to contact you when you do not want them to?”

Following on from the initial profile of stalking cases within the police force, it is important that law enforcement practitioners understand the future crime outcomes for both the victims and suspects of stalking. Crime tends to be concentrated amongst a small proportion of the population, with prior victimisation acting as a good predictor of future victimisation (Pease, 1998). The purpose of the completion of a DARA or S-DASH risk assessment is to identify which victims are most likely to suffer future harm and therefore require different degrees of safeguarding to prevent that harm. By understanding which victims go on to be re-victimised across all crime types, police forces will be able to identify those victims most at risk, as well as those stalkers most likely to perpetrate further harm.

A power curve has been produced to identify the concentration of crime harm (assessed using the Cambridge Crime Harm Index) amongst the population of 1500 victims from the initial 2006 who are victims of at least one crime during the follow-up period. The same power curve has also been produced to identify the concentration of harm caused by the 1427 suspects from the initial 1902 who commit at least one crime during the follow-up period.

Finally, analytical techniques including conditional probability (in this study the probability of a victim becoming a subsequent high harm victim in the follow-up period given that a particular variable is present), mean and standard deviation have been used to identify the concentration of victimisation against many variables including risk grading, crime outcome and stalking escalation. The equation used to calculate conditional probability is:

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$$

3. *What factors that are known at the time of the initial stalking report are most predictive of a victim suffering high harm in the next 12 months?*

For the 1500 individuals that were victimised during the 365-day follow-up period, a total of 211 suffered high harm as per the definition outlined earlier in this chapter. This represents 14% of those that suffered any degree of repeat victimisation, and 10.5% of the entire victim population. Set against the dependent variable of high harm, a total of 27 independent variables were assessed to identify which were most predictive of a victim suffering high harm in the future. The relative risk ratio of each of the 27 independent variables was calculated using the following formula:

	High Harm Outcome = Yes	High Harm Outcome = No
Independent Variable Present (Study Group)	A	B
Independent Variable Not Present (Comparison Group)	C	D

$$\text{Relative Risk Ratio} = \frac{\text{Rate of outcome in study group } (A/(A+B))}{\text{Rate of outcome in comparison group } (C/(C+D))}$$

The risk ratio calculated describes the probability of the victim being subjected to high harm during the follow-up period against all possible outcomes (Ranganathan, 2015). The presence of a higher risk ratio for an independent variable does not infer that it causes the high harm outcome for the victim, but instead points to a correlation between the two. The independent variables are defined in Table 3 below.

**Table 3: 27 independent variables used for relative risk ratio calculation**

<b>Independent Variable</b>	<b>Definition</b>
Both victim and suspect have historic CHI score of 500+ and stalking is DA related	The initial stalking case identified as part of the study has a DA marker present, and both the victim and suspect (of any gender) have suffered harm totalling 500+ on the Cambridge Crime Harm Index in the past
Female victim	The victim is identified as a female
Female victim historic CHI score of 500+	The victim is identified as a female and has suffered harm totalling 500+ on the Cambridge Crime Harm Index in the past
Female victim of DA related stalking	The victim is identified as a female and the initial stalking case has a DA marker present
Length of time of offence before reporting (31+ days)	The range between the lower committed date and the reported date is a minimum of 31 days
Male victim	The victim is identified as a male
Male victim historic CHI score of 500+	The victim is identified as a male and has suffered harm totalling 500+ on the Cambridge Crime Harm Index in the past
Male victim of DA related stalking	The victim is identified as a male and the initial stalking case has a DA marker present
Suspect prior stalking offence	The identified suspect (of any gender) for the initial stalking case is shown as a suspect on Athena for at least one prior stalking case
Victim age range 0-17	Victim (of any gender) is aged between 0-17 at the time of the offence being reported
Victim age range 18-24	Victim (of any gender) is aged between 18-24 at the time of the offence being reported
Victim age range 25-35	Victim (of any gender) is aged between 25-35 at the time of the offence being reported
Victim age range 36-45	Victim (of any gender) is aged between 36-45 at the time of the offence being reported
Victim age range 46-55	Victim (of any gender) is aged between 46-55 at the time of the offence being reported
Victim age range 56-65	Victim (of any gender) is aged between 56-65 at the time of the offence being reported
Victim age range 66+	Victim (of any gender) is aged 66+ at the time of the offence being reported
Victim and suspect both have historic CHI score of 500+	Both the victim and suspect (of any gender) have suffered harm totalling 500+ on the Cambridge Crime Harm Index in the past

Victim historic CHI score of 500+	Victim (of any gender) has suffered harm totalling 500+ on the Cambridge Crime Harm Index in the past
Victim historic crime count of 0	Victim (of any gender) has no recorded prior victimisation on Athena
Victim historic crime count of 1 to 9	Victim (of any gender) is shown as a victim on between 1 and 9 crime reports on Athena prior to the date of the initial stalking
Victim historic crime count of 10 to 19	Victim (of any gender) is shown as a victim on between 10 and 19 crime reports on Athena prior to the date of the initial stalking
Victim historic crime count of 20 to 29	Victim (of any gender) is shown as a victim on between 20 and 29 crime reports on Athena prior to the date of the initial stalking
Victim historic crime count of 30 to 39	Victim (of any gender) is shown as a victim on between 30 and 39 crime reports on Athena prior to the date of the initial stalking
Victim historic crime count of 40+	Victim (of any gender) is shown as a victim on 40+ crime reports on Athena prior to the date of the initial stalking
Victim of DA related stalking	Victim (of any gender) and initial stalking case has a DA marker present
Victim prior stalking victim	The identified victim (of any gender) for the initial stalking case is shown as a victim on Athena for at least one prior stalking case
Victim with a suspect who has a historic CHI score of 500+	Suspect (of any gender) has suffered harm totalling 500+ on the Cambridge Crime Harm Index in the past

*Random Forest (modelling completed by Dr Valdebenito, 2022)*

To further answer the research question, the data from this study has been imputed into a random forest model. A classification version of the random forest algorithm has been used for this study, which utilises an ensemble learning algorithm to rank order predictor variables in respect of their correlation to harm, in order to forecast the outcome of “high harm” for the victim population. The algorithm creates 2200 decision trees and outputs the most frequent class predicted by them (Breiman, 2001). When there is confidence in the effectiveness of the random forest in its capability to predict which victims from the past go on to suffer high harm, there is potential to use the algorithm in the live

operational environment to predict which victims who are being stalked now are most likely to suffer high harm in the future and therefore target resources and safeguarding interventions appropriately.

The cases within the population were randomly allocated into two subsets of data: the training data and the testing data. The *training* data was made up of 1433 cases (70% of the population) and was used to build up the random forest model for predicting which victims within the population would go on to suffer high harm within the subsequent 365-days based on the 15 predictive variables built into the model. The *testing* data was made up of 573 cases (30% of the population) and was kept entirely separate from the training data to allow for a rigorous test of the model when presented with new data. The algorithm calculated the overall accuracy of the forecasting model, as well as the sensitivity and specificity of its ability to classify victims into the low harm and high harm categories. The importance of each variable to predict future high harm was tested exploring the Gini Index (Berk, 2012). All analyses were run using the package randomForest (Liaw & Wiener, 2002) in R studio, version 1.2.5033.

The starting point for the random forest analysis is the date of recording of the initial stalking case. Just as with the relative risk ratio calculations, only information known to police at the time of the offence being recorded could be used as a predictor variable, given that anything that occurred during the subsequent 365-day follow-up period would be unknown at the time of recording the offence. The same consistent follow-up period was applied to this analysis as has been used throughout the rest of the study to protect the internal validity of the findings. A total of 15 predictor variables were included in the model.



A full data dictionary which provides a description of each of these variables can be found at Appendix 4.

## Results

### *Introduction*

This chapter begins with a descriptive analysis of the population of 2006 stalking cases extracted from the Athena database for this study to provide a baseline that gives context for the subsequent findings in relation to future harm outcomes. This descriptive analysis also provides the opportunity to gain an insight into the prevalence and types of stalking being recorded by a police force in England and Wales, as well as showing the demographics present within the victim and suspect populations. It then moves on to an analysis of the concentration of future harm within the population during the 365-day follow-up analysis, establishing the severity and frequency of future harm and identifying which victims most suffer that harm, and which suspects most cause it. There is then a description of the risk factors identified as being correlated to future high harm outcomes for victims, using relative risk ratios to rank order twenty-seven independent variables assessed against the single dependant variable of high harm, constituting a cumulative Cambridge Crime Harm Index score of 500 or greater in the 365-day follow-up period. The chapter concludes with the results of a “random forest” analysis of the data, outlining the predictability of future high harm based on 15 predictor variables.

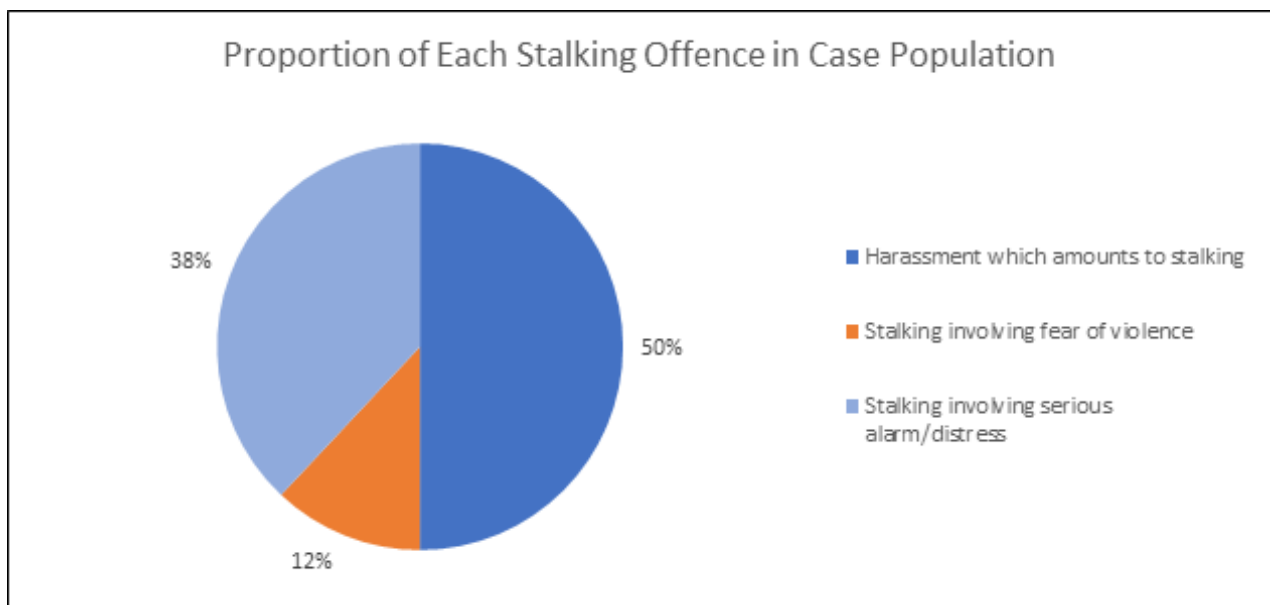
### *Research Question 1*

***Q1. What is the profile of stalking cases in a police force in England and Wales that is utilising the Domestic Abuse Risk Assessment (DARA)?***

### *Initial Stalking Offence*

There are three notifiable stalking offences, each allocated a CCHI score. Harassment with Stalking is the lowest scoring stalking offence, with a CCHI score of 10. The next is Stalking with Fear of Violence, which receives a CCHI score of 84. Finally, the most severe stalking crime recorded is Stalking Involving Serious Alarm/Distress with a CCHI score of 252. Figure 2 outlines the proportion of cases recorded as each of these offences, indicating the harm level at which each victim entered the cohort.

**Figure 2: Proportion of cases recorded as each of the three notifiable stalking offences**

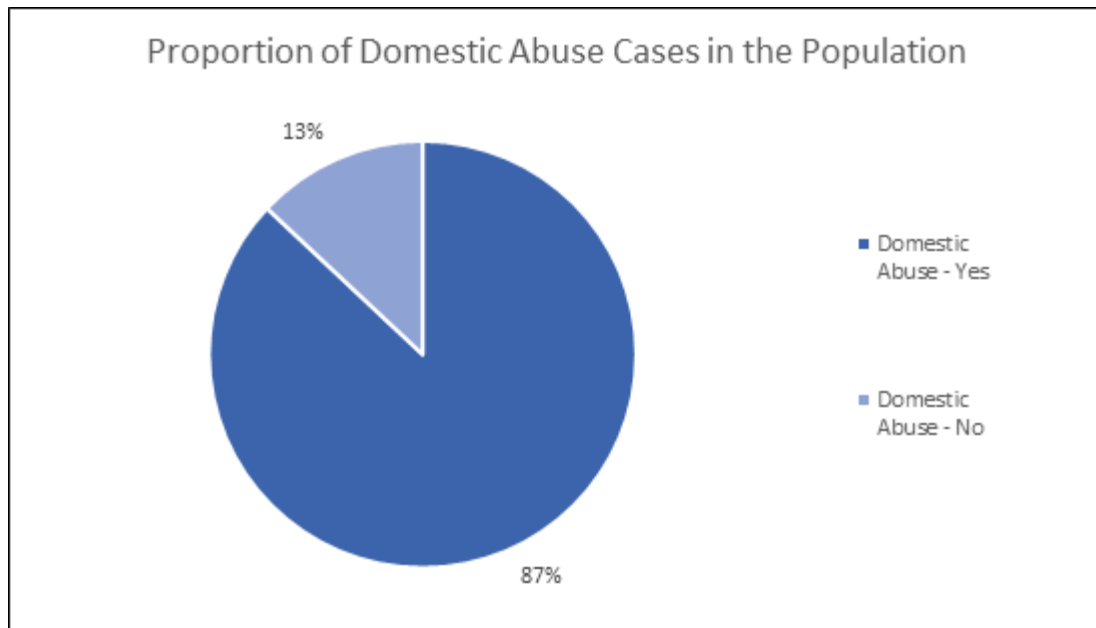


It is interesting to note that a significant proportion of victims (38%) came into the cohort at the highest harm level for stalking, far more so than at the medium harm level (12%). Unsurprisingly, the majority of victims had experienced the lowest level of stalking, without any fear, alarm or distress (50%).

### *Prevalence of Domestic Abuse*

The literature review chapter outlined the strong link between stalking and domestic abuse. Figure 3 illustrates the breakdown between stalking cases where domestic abuse was a factor versus those cases that were not.

**Figure 3: Percentage of stalking cases within the population which contained a domestic abuse marker**

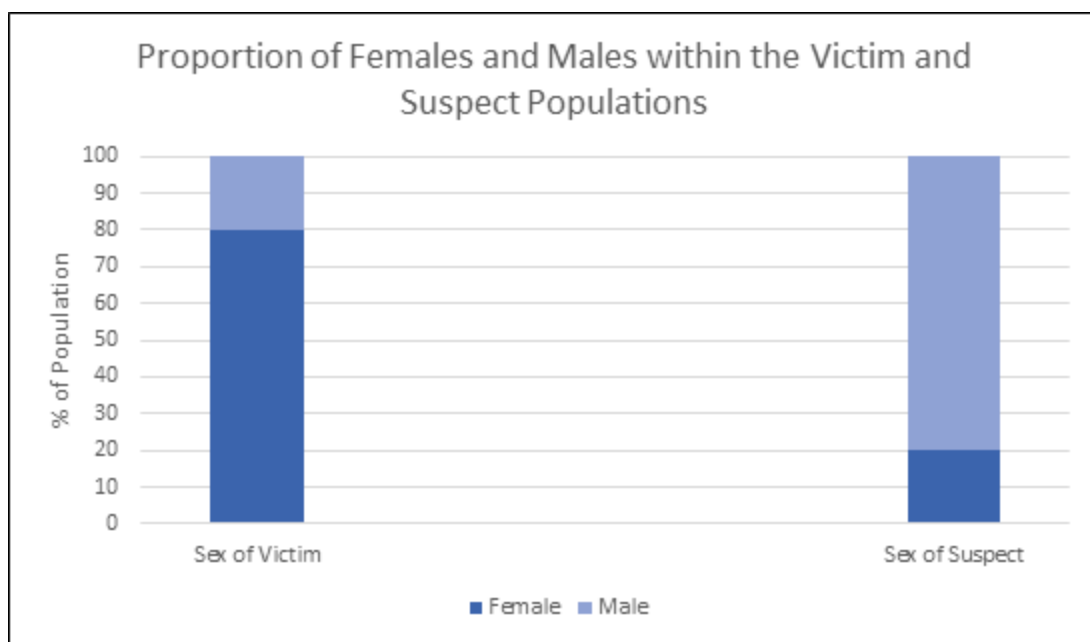


The volume of cases with a domestic abuse marker is high at 87%, reinforcing the existing research that indicates that the most prevalent type of stalking is ex-intimate partner stalking. Even in the context of it being recognised as the most common form of stalking, the figure of 87% is perhaps surprisingly high. In a meta-analysis by Spitzberg and Cupach (2007) data indicated that across 175 studies into stalking, half of cases were between a couple that had been romantically acquainted. The increased volume of domestic abuse related cases in this study may be due to the DARA risk assessment framework, which has potentially picked up most of the stalking cases recorded where the primary reason for the victim coming forward to police may not have been the stalking behaviours they were experiencing.

### *Difference Between the Sexes*

Whilst the estimates vary across the studies, from 60-80% (Spitzberg & Cupach, 2007) to 92% (Sheridan, 2001), the literature agrees that females are more likely to be victims of stalking than males. Figure 4 displays the breakdown of female to male victims and suspects observed in this study.

**Figure 4: Percentage of female and male victims and suspects observed in the study**



There is a symmetry of female victims (80%) to male suspects (80%) within this study, broadly in line with what has been identified in other studies across multiple settings (Spitzberg & Cupach, 2017). This further reinforces the notion that stalking is a crime that disproportionately affects women and girls and the findings from this study must be viewed alongside the police's strategy to tackle violence against women and girls.

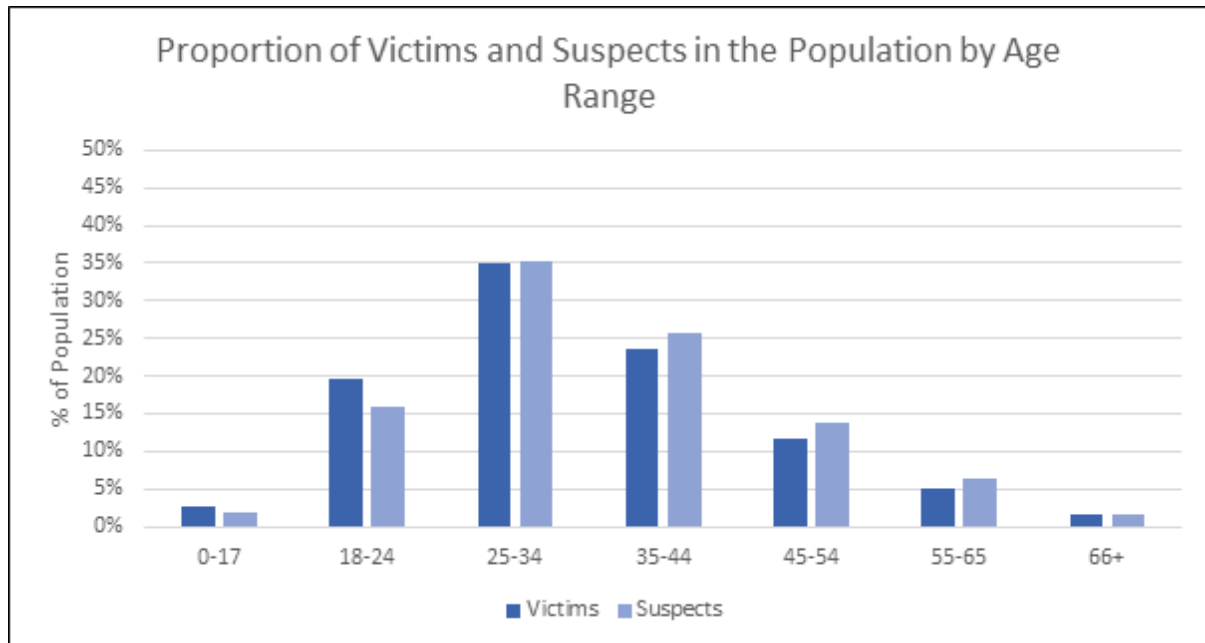
### *Breakdown of Ages*

Stalking is generally believed to be a crime that is both committed by and affects younger people, with an average age of victims observed to be in the early to mid-30's

(Brady & Nobles, 2015) and suspects slightly older at mid to late-30's (Kropp et al, 2002).

Figure 5 illustrates the breakdown in age groups amongst the victims and suspects within the study.

**Figure 5: Proportion of victims and suspects in each age group within the study population**



The graph demonstrates that the most common age range of both victims and suspects within the study is 25-34, making up 35% of the total victims and 35% of the total suspects. The average age of a victim in the study is 34 (SD = 12, range 8 – 88, Median = 32) and the average suspect age was 36 (SD = 12, range 13 – 82, Median = 34), reinforcing what has been observed in previous research.

### *Length of Offence*

It is said that a victim can suffer up to 100 incidents of stalking before reporting it to the police (Paladin, 2017), indicating that the crime can take place over a prolonged period before coming to police attention. Whilst this study has not been able to quantify the number of individual occurrences that took place before the initial stalking was recorded, by

measuring the number of days between the lower committed date and the reported date, it is possible to identify how long each of the victims within the population were living with stalking prior to making a report to police. The average number of days that the stalking had been going on prior to being reported to police in this study is 159 days (SD = 593, range 0 – 12973, median = 11). Whilst these figures suggest that 100 incidents are unlikely to have occurred before reporting to the police, the findings do confirm that victims live with stalking for a significant period of time before seeking help from law enforcement.

### *Police Assigned Outcomes*

There is very little in the literature to indicate the breakdown of different police outcomes assigned to stalking investigations. Tracking the police outcomes assigned to stalking cases has some significance, as similarly sharp recording increases in other crimes in recent years, for example assault and public order offences, has seen no increase in the absolute number of “positive outcomes”, for example charges. In fact, due to the substantial increase in recorded crime and the absolute number of positive outcomes remaining static, the net effect is that the proportion of “solvable” cases has notably dropped (McFadzien et al, 2020). It is therefore important to consider the number of stalking cases receiving a positive outcome since the volume of recorded crime has increased six-fold in two-years to assess the impact that positive police interventions have on subsequent victim harm and suspect behaviour. Figure 6 illustrates the frequency of occurrence of the fifteen different police recorded outcomes assigned to cases within the population.

**Figure 6: Frequency with which each of the police recorded outcomes appear within the population**

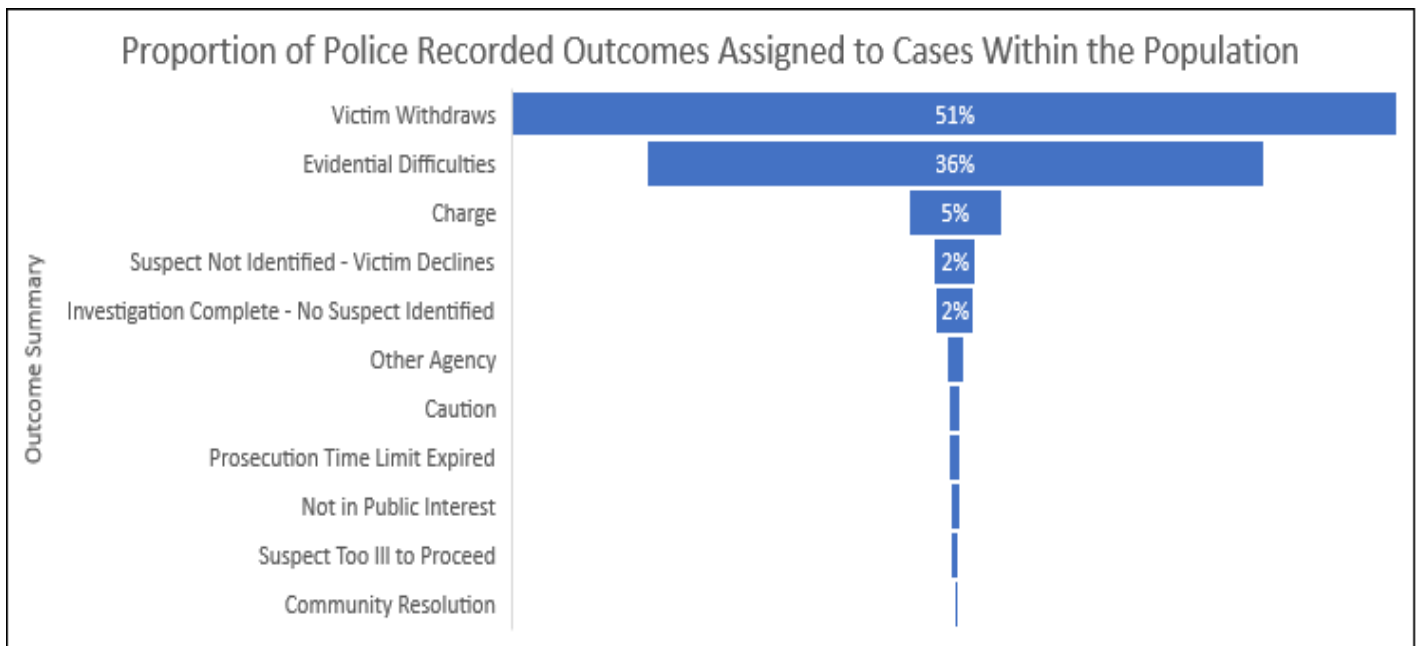


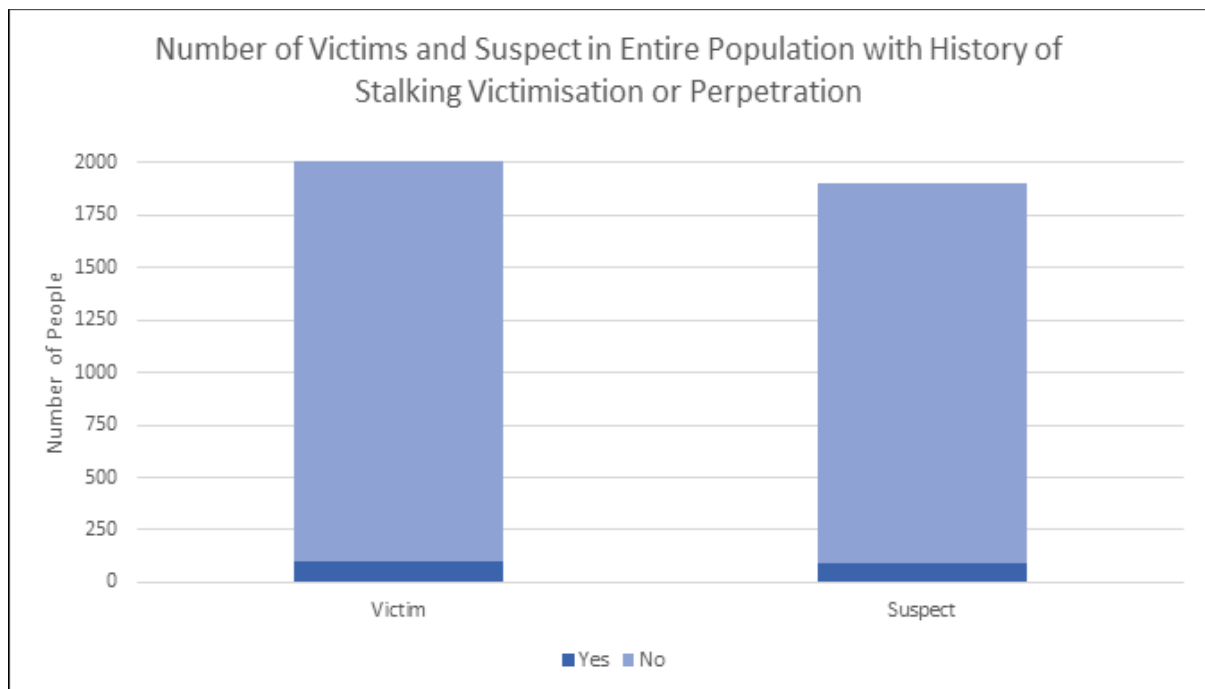
Figure 6 shows that just 5% of the 2006 cases (n=106) received a charge, whilst 93% of cases received one of the “no further action” (NFA) outcomes (n=1866). The outcome assigned most frequently was Outcome Type 16 – Named Suspect Identified: Evidential Difficulties Prevent Further Action: Victim Does Not Support (Or Has Withdrawn Support From) Police Action.

#### *Victim and Suspect Crime History*

The literature review outlined how repeat victimisation is common within stalking cases, given that the course of conduct by its very definition requires multiple events to have taken place. Figure 7 outlines the number of victims and suspects who had a history of stalking victimisation or perpetration recorded on the Kent Police database.



**Figure 7: Number of victims and suspects in the cohort who have previous victimisation or perpetration records of stalking on the Athena database**

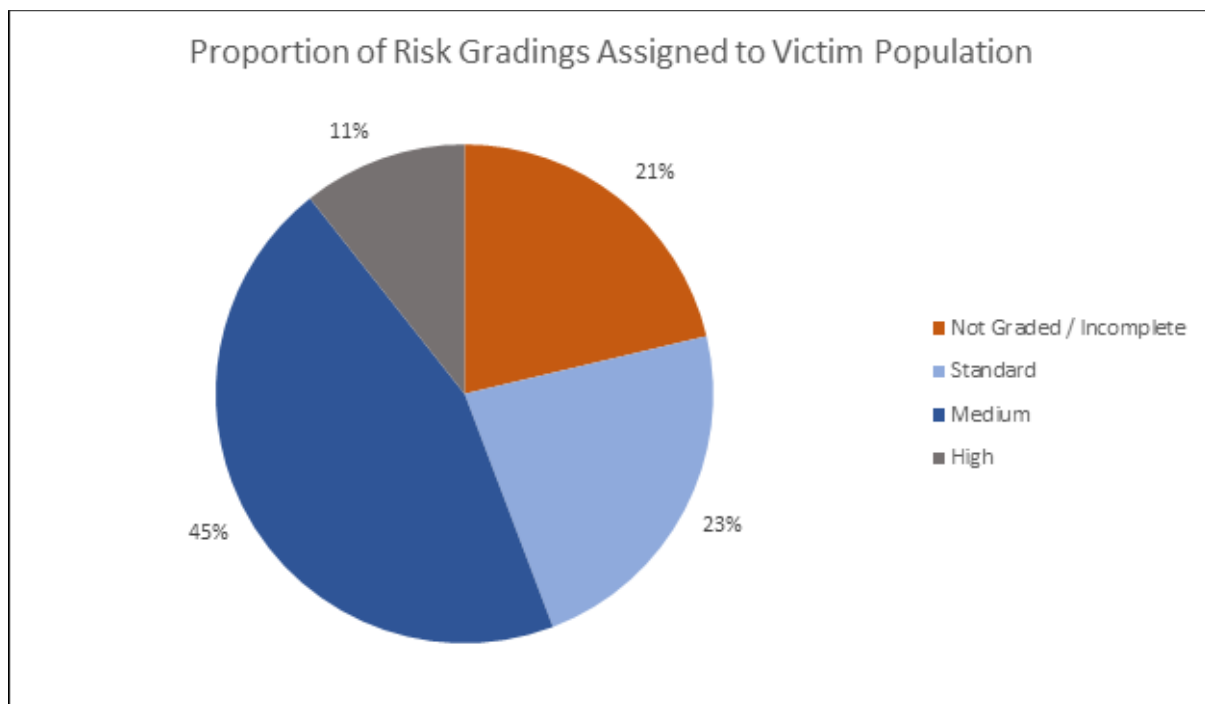


97 of the victims within the cohort (5%) have a history of stalking victimisation on Athena. Further symmetry is seen in the number of suspects who have a history of stalking on the system at 92 (5%). For those 97 victims who had been recorded as previous victims of stalking, only 57 of their stalkers had a history of stalking perpetration. This indicates that 40 of the victims (41% of those who had been a stalking victim previously) had been stalked by more than one perpetrator during their lifetime. Conversely, 35 of the stalkers (38% of those who had stalked before) had stalked a minimum of two separate individuals.

### *Risk Gradings*

This thesis has outlined the current risk assessment processes deployed by Kent Police for both domestic abuse and stalking. The study has captured the rates at which victims are graded as standard (the lowest risk grading), medium and high risk, outlined in Figure 8.

**Figure 8: Proportion of risk grading categories assigned to the cohort victims**



Just under half (45%) of victims were assigned a risk grading of medium (n=905), almost double that of the next most selected grade of standard at 23% (n=460). 11% of victims were graded as being at high risk of future harm (n=214). Just over one fifth of victims (21%) received either an incomplete risk assessment or no risk assessment whatsoever (n=427). This is of particular significance when looking at the breakdown between risk assessments for those cases with a domestic abuse marker versus those which are not domestic abuse related. Of the population of 2006 cases, 260 did not contain a domestic abuse marker (13%). Of those 260 cases, just 18% were subject to a risk assessment, leaving the future risk posed to 213 victims unassessed. It should be noted that towards the latter part of the cohort period the use of S-DASH in all stalking cases, not just those that were domestic abuse related, was mandated in Kent Police, and as such any follow-up study may find a greater risk assessment completion rate.

## *Research Question 2*

### ***Q2. What is the concentration of high harm for a population of stalking victims during a 365-day follow-up period?***

Clinical case reviews into the most harmful offences such as the murder of women and girls indicates that stalking is a critical factor in the build-up to the acute event (Monckton-Smith et al, 2017). There are however limitations to the clinical case review approach, specifically the difficulty in generalising the findings, the difficulty in establishing a cause-effect relationship, and the danger of over-interpretation (Nissen & Wynn, 2014). As such, this study has focused on tracking the prevalence and severity of future victimisation for the entire cohort of 2006 victims across a 365-day follow-up period. It was anticipated that this would provide an important insight into the repeat victimisation rates for stalking victims, as well as identifying whether there is a concentration of victims who go on to suffer most of the harm, and suspects who go on to perpetrate most of the harm.

#### *Subsequent Crime Count*

During the 365-day follow-up, victims went on to suffer a total of 476,263 units of harm, based on the CCHI. This score was calculated from a total count of 5270 subsequent crimes. Table 4 illustrates the number of victims who suffered each count of subsequent crimes, with the highest figure observed being one victim who was subjected to a further 46 recorded crimes. Suspects went on to commit a further 5969 subsequent crimes, totalling 463,749 units of harm. Table 5 illustrates the number of suspects who perpetrated each quantity of subsequent crimes, up to the highest figure observed of 42.

**Table 4: Count of subsequent victimisations during 365-day follow-up sorted in descending order by number of recorded crimes**

<b>Subsequent Crimes</b>	<b>Number of Victims</b>	<b>Cumulative Victims</b>	<b>% Victims</b>	<b>Cumulative % Victims</b>
46	1	1	0.0%	0.0%
41	1	2	0.0%	0.1%
40	1	3	0.0%	0.1%
39	1	4	0.0%	0.2%
37	2	6	0.1%	0.3%
35	1	7	0.0%	0.3%
31	1	8	0.0%	0.4%
29	1	9	0.0%	0.4%
27	2	11	0.1%	0.5%
26	1	12	0.0%	0.6%
25	1	13	0.0%	0.6%
24	1	14	0.0%	0.7%
23	1	15	0.0%	0.7%
22	1	16	0.0%	0.8%
21	1	17	0.0%	0.8%
20	4	21	0.2%	1.0%
19	4	25	0.2%	1.2%
18	5	30	0.2%	1.5%
17	5	35	0.2%	1.7%
16	5	40	0.2%	2.0%
15	6	46	0.3%	2.3%
14	6	52	0.3%	2.6%
13	8	60	0.4%	3.0%
12	11	71	0.5%	3.5%
11	10	81	0.5%	4.0%
10	17	98	0.8%	4.9%
9	22	120	1.1%	6.0%
8	34	154	1.7%	7.7%
7	36	190	1.8%	9.5%
6	53	243	2.6%	12.1%
5	73	316	3.6%	15.8%
4	119	435	5.9%	21.7%
3	179	614	8.9%	30.6%
2	335	949	16.7%	47.3%
1	551	1500	27.5%	74.8%
0	506	2006	25.2%	100.0%

**Table 5: Count of subsequent records of being recorded as a suspect during 365-day follow-up sorted in descending order by number of recorded crimes**

<b>Subsequent Crimes</b>	<b>Number of Suspects</b>	<b>Cumulative Suspects</b>	<b>% Suspects</b>	<b>Cumulative % Suspects</b>
42	1	1	0.1%	0.1%
40	1	2	0.1%	0.1%
39	3	5	0.2%	0.3%
37	1	6	0.1%	0.3%
30	1	7	0.1%	0.4%
29	1	8	0.1%	0.4%
28	1	9	0.1%	0.5%
27	6	15	0.3%	0.8%
26	1	16	0.1%	0.8%
25	2	18	0.1%	0.9%
24	1	19	0.1%	1.0%
23	4	23	0.2%	1.2%
21	3	26	0.2%	1.4%
20	5	31	0.3%	1.6%
19	2	33	0.1%	1.7%
18	5	38	0.3%	2.0%
17	8	46	0.4%	2.4%
16	7	53	0.4%	2.8%
15	6	59	0.3%	3.1%
14	11	70	0.6%	3.7%
13	14	84	0.7%	4.4%
12	16	100	0.8%	5.3%
11	19	119	1.0%	6.3%
10	24	143	1.3%	7.5%
9	32	175	1.7%	9.2%
8	38	213	2.0%	11.2%
7	61	274	3.2%	14.4%
6	71	345	3.7%	18.1%
5	86	431	4.5%	22.7%
4	125	556	6.6%	29.2%
3	190	746	10.0%	39.2%
2	287	1033	15.1%	54.3%
1	440	1473	23.1%	77.4%
0	429	1902	22.6%	100.0%

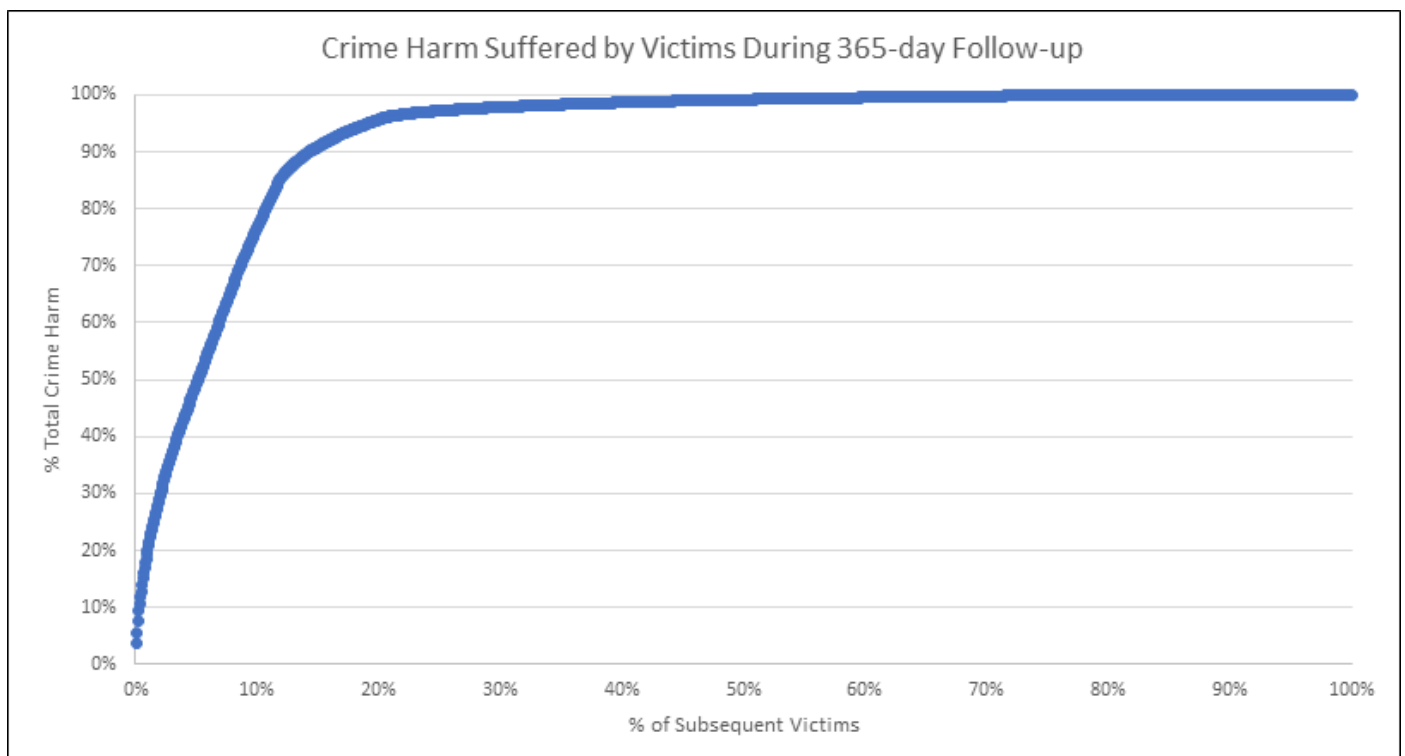
The first point to note from the two tables above is that 25.2% of victims are not recorded as victims and 22.6% of suspects are not recorded as suspects on any crimes in Kent Police in the following twelve months after the initial stalking case. Conversely, that

means that most individuals either go on to suffer or perpetrate future crime events. If one further event per month, totalling twelve subsequent crimes in the full year, is taken to represent a “high” level of repeat victimisation or re-offending, then a clear power few (Sherman, 2007) victims and suspects can be identified. 3.5% of victims, totalling, 71 people, suffer 12 or more further victimisations, whilst 5.3% of suspects, totalling 100 people, commit 12 or more offences in the 365-day follow-up period. The slight discrepancy in the totals indicate that suspects may be carrying out unrelated offending against different targets, whereas victims tend to be victimised by a single perpetrator.

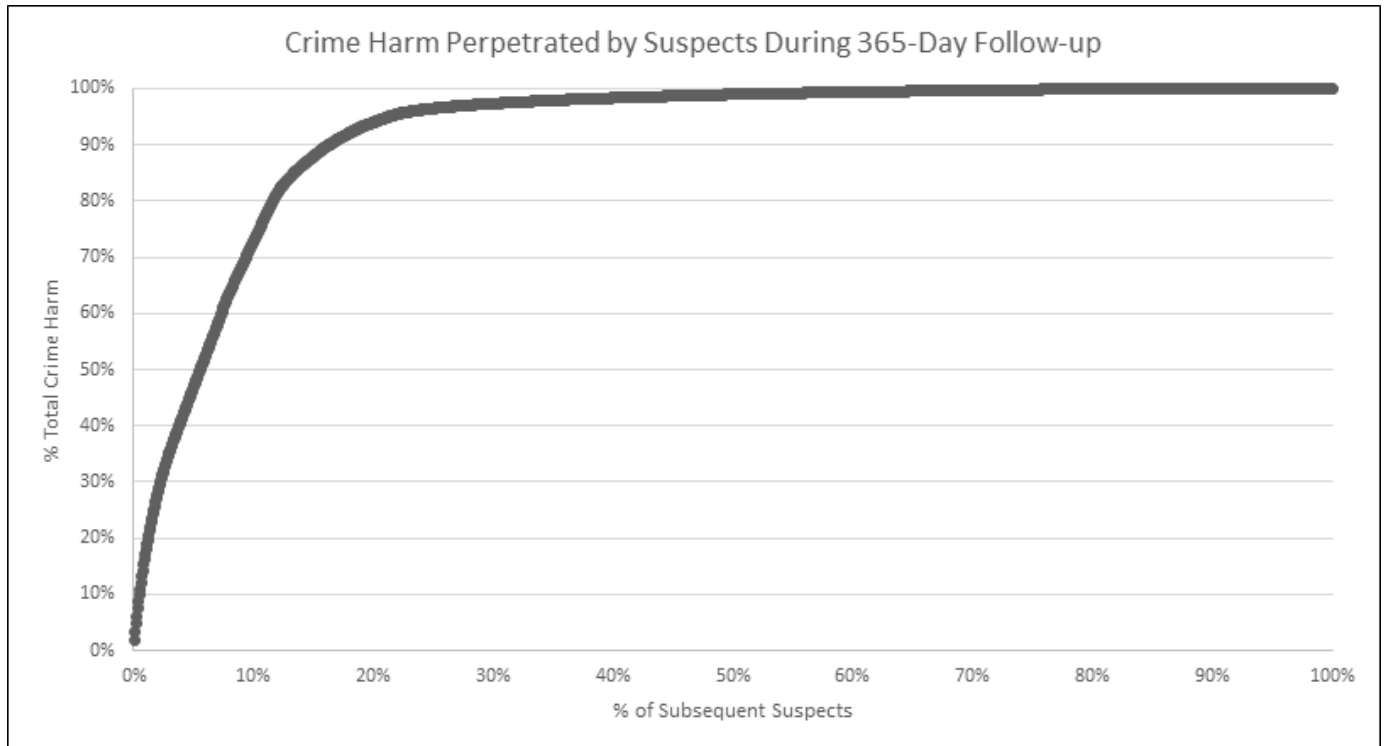
### *Concentration of Harm – The ‘Power Few’*

The power few phenomena is equally observable in the concentration of harm during the follow-up period. Figures 9 and 10 depict the power curves for crime harm, as measured using the CCHI, for victims and suspects.

**Figure 9: Power curve to illustrate the concentration victims who suffer harm within the 365-day follow-up period**



**Figure 10: Power curve to illustrate the concentration of suspects who perpetrate harm within the 365-day follow-up period**



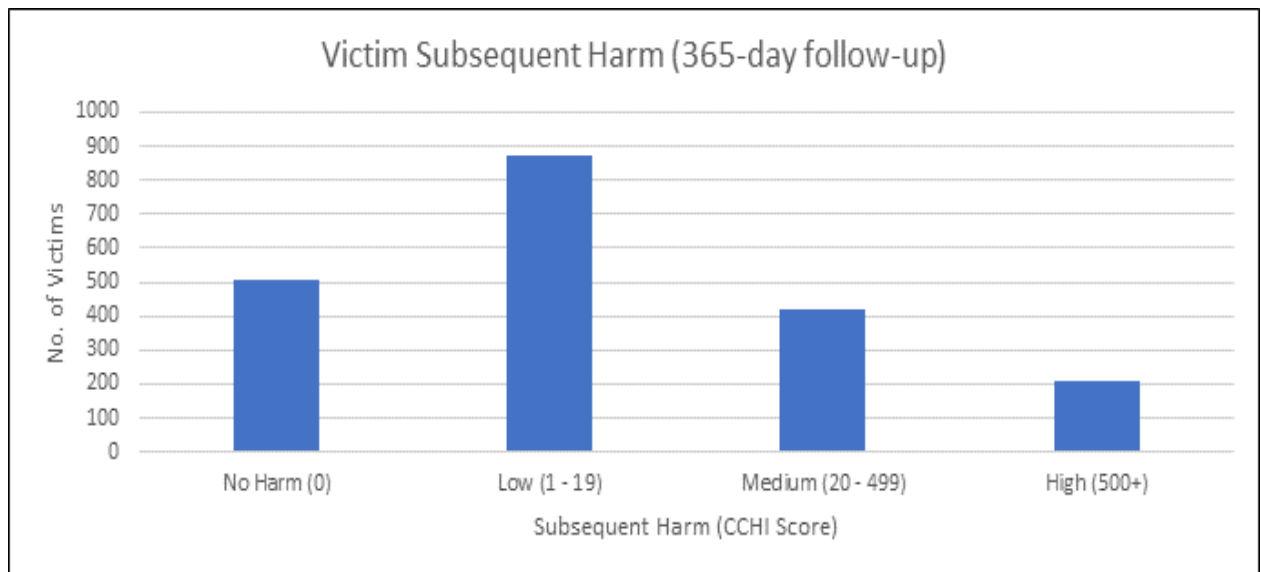
Figures 9 and 10 demonstrate that for both victims and suspects, stalking harm is highly concentrated in a small number of individuals during the cohort follow-up period. For victims, just 5.2% of victims (n=78) suffer 50% of the total recorded harm in the 365-day follow-up period. The steep trajectory continues, with 10% of victims suffering almost 77% of the total harm, and 14% of victims suffering 90% of the recorded harm.

As the power curve chart demonstrates, there is great symmetry in the victim and suspect data. For suspects, 5.5% of suspects (n=79) cause 50% of the harm, with 10% of suspects causing 73% of harm and 16% of suspects causing 90% of harm. These findings have potentially significant policy implications as they demonstrate that a large amount of harm can be prevented from occurring through the effective targeting of a small number of suspects and the effective safeguarding of a small number of victims.

### Average Harm

Deeper analysis of the data regarding harm illustrates the levels of harm suffered by victims during the follow-up period. Figure 11 illustrates the quantity of victims sat within each harm range following the 365-day follow-up period.

**Figure 11: Quantity of victims subjected to each level of harm during the follow-up period**

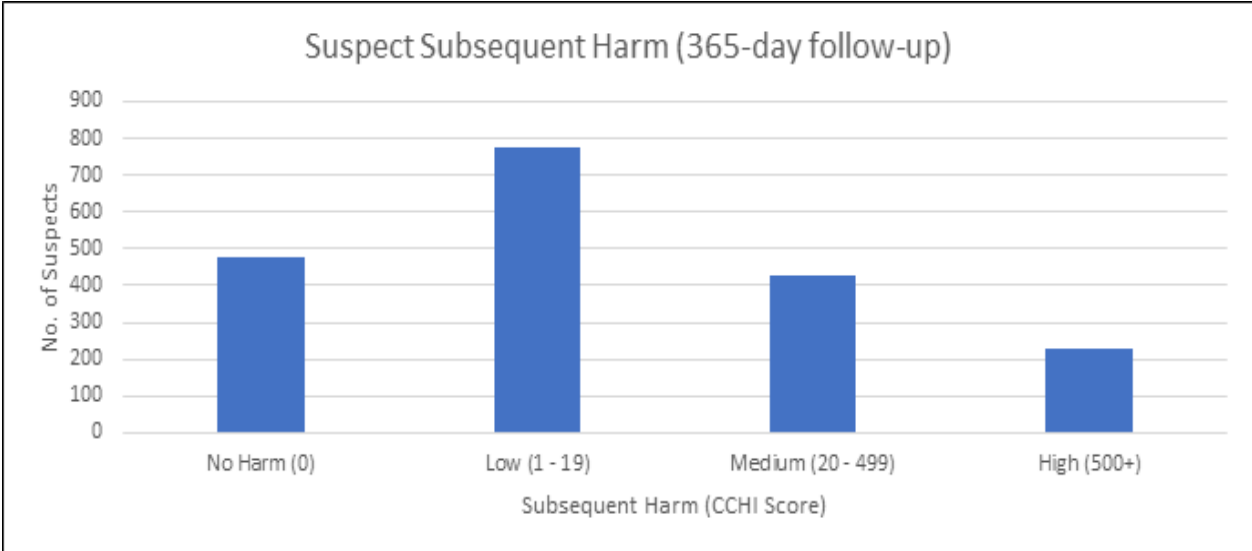


A quarter of victims were not subjected to any further recorded crimes during the follow-up period and as such came to no known harm. The most common level of harm suffered by victims was “low harm” (43%), experiencing one or more incidents that did not amount to a cumulative harm score of 20 or above. The upper end of harm then concentrates to fewer victims, with 21% in the medium harm category and 11% suffering high harm.

The suspect data is strikingly similar once again, as is demonstrated in Figure 12.



**Figure 12: Quantity of suspects who caused each level of harm during the follow-up period**



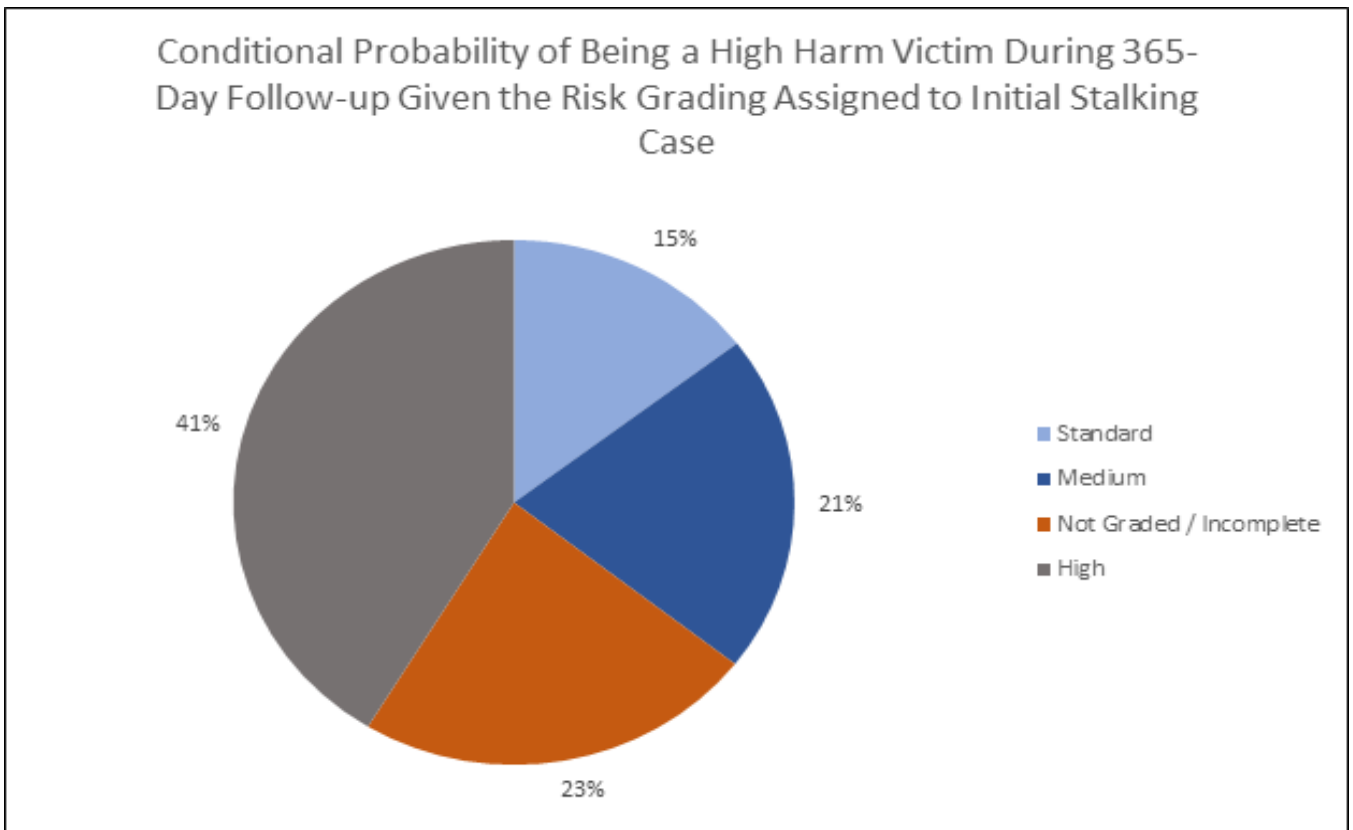
A quarter of suspects were not recorded as being a suspect on any crime in the following 12 months and as such caused no known harm. 41% of suspects perpetrated crimes totalling a cumulative CCHI score below 20 meaning that they fall in the low harm category. There is a greater concentration of suspects at the upper end of harm with 22% causing medium harm and 12% causing high harm. The higher proportion of suspects perpetrating higher harm offending than victims suffering higher harm offending further demonstrates that suspects are carrying out offences against individuals other than the stalking victim.

*Risk Grading and Future High Harm and Crime Count*

The risk grading (S-DASH) assigned to a case by the initial attending officer is intended to outline the potential risk to the victim of coming to further harm in the future. The conditional probability of going on to be a high harm victim during the follow-up period, given the risk grading assigned, has been calculated for each of the three risk gradings, as

well as no grading assigned due to a lack of risk assessment. Figure 13 illustrates these conditional probabilities.

**Figure 13: Conditional probability of becoming a high harm victim during the follow-up period given that the original stalking case was assigned a risk grading of standard, medium, high, or no grading**



The chart shows that the police conducted risk assessments are broadly accurate as an indicator of the likelihood of the victim coming to harm in the future, in so much as that those victims graded as “high” risk had the highest probability of becoming a high harm victim in the study population. Importantly this data highlights a significant probability of high harm for those victims who have not been risk assessed, or where the risk assessment was incomplete, marginally higher than those victims deemed to be at medium risk of future harm. This suggests that police are “blind” to a substantial amount of potential harm.

The average amount of harm has also been assessed at each risk grading and is illustrated in Table 6.

**Table 6: Average harm, standard deviation and median for victims categorised under each risk grading**

Risk Grading	Number Assigned	No. Future High Harm	Average Harm (CCHI)	Standard Deviation	Median Harm (CCHI)	Average Count	Standard Deviation	Median Count
Not Graded / Incomplete	427	47	253	861	3	3	4	1
Standard	460	32	137	477	2	2	3	1
Medium	905	90	226	710	10	3	4	2
High	214	42	468	1446	19	4	4	2

The average harm levels suffered by victims during the follow-up period is significant, with the average sitting at a medium harm level as per the definition for this study at all risk gradings. The large standard deviations for all risk gradings indicate the massive range of harm scores observed. These extremities are flattened in the median harm scores observed, once again illustrating that those victims graded high risk suffer the most harm, followed by medium risk and then standard risk. It is only in the median harm score that a significant difference between medium risk and no risk grading is observed.

Subsequent crime count against victims by risk grading follows the same pattern as subsequent crime harm, with victims assessed as high risk going on to suffer the most crime incidents in the cohort follow-up period, averaging 4 (SD = 4). Standard risk victims go on to suffer an average of 2 subsequent victimisations (SD = 3), whilst both medium risk victims and victims who received no risk grading experience an average of 3 subsequent victimisations (SD = 4).

#### *Repeat Stalking and Escalation*

Whilst repeat victimisation in the follow-up period has generally been assessed based on all crimes, the analysis has also identified the prevalence of repeat stalking events against the victims. A total of 324 victims went on to be a recorded victim of stalking during the 365-day follow-up period, equating to 16% of the total victim population. This figure far exceeds the number of victims in the population who have a history of stalking victimisation (5%), perhaps indicating a future trend in the volume of repeat stalking victimisation now that a greater number of stalking crimes are being recorded by Kent Police.

The stalking was found to have escalated to a more harmful stalking offence than that which was originally reported, as measured by the CCHI score, in 45 cases. There is therefore a 14% conditional probability of a stalking victim going on to become a victim of a more harmful stalking offence during the follow-up period given that they are victim of at least one further recorded stalking. This somewhat conflicts with the existing literature, where it is claimed that 75% of victims felt that the stalking had worsened (Sheridan et al, 2001b). It should however be noted that the key difference in these findings is that it is a comparison between police recorded data, recorded in accordance with the Home Office Counting Rules, versus victim perception.

The average amount of harm suffered by victims who had been subjected to escalating stalking is extremely high at 1049 CCHI (SD = 2671, range 84 – 17790, median = 286), whilst for victims where there was further stalking but no escalation the average harm was slightly lower, but still very high at 520 (SD = 1164, range 10 – 9222, median = 41). These findings suggest that any further stalking, whether it is deemed to have escalated or not, is a strong indicator of a victim suffering significant levels of future harm. This is especially true when comparing those victims who are further victims of stalking with those

that are victimised but by other offences, with those victims having an average subsequent harm score of 241 (SD = 693, range 1 – 8791, median = 10). Table 7 illustrates the average and median harm levels for victims based on their repeat victimisation status.

**Table 7: Average subsequent harm and count, standard deviation and median for victims based on their repeat victimisation status**

Repeat Victimisation Status	Number of Victims	No. Future High Harm	Average Harm (CCHI)	Standard Deviation	Median Harm (CCHI)	Average Count	Standard Deviation	Median Count
All Victims	2006	211	169	591	2	2	3	1
Repeat Stalking No Escalation	279	55	520	1165	41	7	7	4
Repeat Stalking With Escalation	45	15	1049	2671	286	8	9	5
Repeat Victim (Not Stalking)	1176	141	241	694	10	3	3	2

Looking at repeat victimisation from a crime count perspective also yields some interesting findings. Victims who are subjected to further stalking during the follow-up period, whether it escalates or not, are found to have been subjected to significantly more crime events than those who are re-victimised by crimes other than stalking.

These findings are significant as they demonstrate that when measured against the variables of both crime count and harm, victims are more likely to experience worse outcomes if the stalking continues beyond one event, whether or not the subsequent stalking is an escalation of the prior event. If the stalking has escalated, then the median subsequent harm suffered in the follow-up period is almost seven times greater than if there is no escalation.

#### *Police Recorded Outcome and Subsequent High Harm*

Within this dataset, the police outcome code assigned to the initial stalking investigation appears to have little impact on the future harm that a victim suffers. Table 8

outlines the conditional probability of being a high harm victim during the cohort follow-up period given the police recorded outcome assigned to the original stalking investigation, as well as the average harm suffered by victims who had their cases assigned each outcome. Given the high number of possible outcome codes to be assigned and the lack of many of them appearing in the dataset, for the purpose of this analysis the “no further action” (NFA) outcomes have been grouped together, with only “victim withdraws” separated out for individual analysis. Furthermore, the only “positive” outcome analysed is that of charge given the lack of prevalence of any of the other outcomes.

**Table 8: Future victim high harm based on the outcome code assigned to the initial stalking investigation**

<b>Outcome</b>	<b>Number Assigned</b>	<b>No. Future High Harm</b>	<b>Conditional Probability High Harm</b>	<b>Average Subsequent Harm Score</b>	<b>Standard Deviation</b>	<b>Average Subsequent Crime Count</b>
Charge	106	8	8%	180	594	3
All NFA Outcomes	1868	201	11%	242	832	3
Victim Withdraws	1016	100	10%	224	746	3

Table 8 demonstrates that the presence of a charge outcome to the original investigation has a positive effect on the conditional probability of becoming a high harm victim within the follow-up period. Whilst charging the suspect does not impact on the number of crimes that the victim is on average subjected to during the following year, it does appear to decrease the average of level of cumulative harm from those crimes. The higher standard deviation figure for both the “All NFA Outcomes” and “Victim Withdraws” variables indicate a wider range of CCHI scores experienced by these victims and as such the average is likely to be more heavily influenced by the highest harm outliers. It should also be noted that with only 106 charges assigned and just 8 of those victims going on to suffer high harm during the follow-up year, this represents a small sample of the overall population and

any findings here should be treated with some caution with further research recommended to corroborate whether this holds true in a larger sample.

### *Research Question 3*

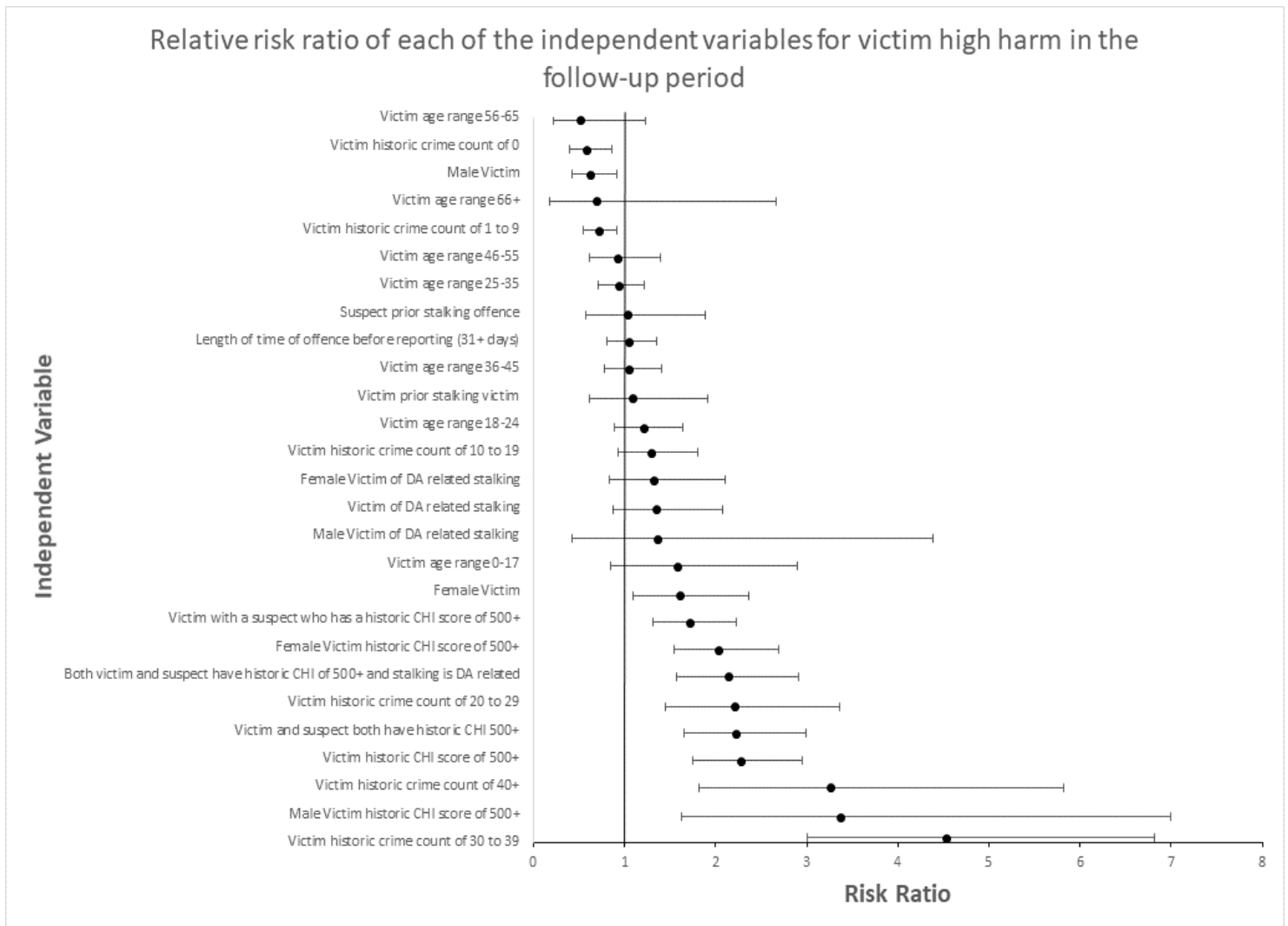
#### ***3. What factors that are known at the time of the initial stalking report are most predictive of a victim suffering high harm in the next twelve months?***

One of the primary motivations for carrying out this research was the massive increase in the number of reported stalking cases in Kent Police over a three-year period. With so many more cases of stalking to investigate, it is more important than ever that officers can effectively predict which victims are most at risk of suffering high harm, and which suspects are most likely to perpetrate high harm in the future. This will allow for the effective targeting of limited police resources to those victims and suspects for which police intervention will have the greatest impact on the reduction in overall harm suffered.

#### *Relative Risk Ratio*

To assist and enhance the pre-existing S-DASH risk assessment procedure, twenty-seven independent variables have been assessed to identify their impact on the likelihood of a victim suffering cumulative high harm during the 365-day follow-up period. This has been calculated using relative risk ratios. A risk ratio of “1” indicates no change in risk. A score lower than 1 indicates a decreased risk, whilst a score over 1 indicates an increased risk, with this risk increasing the higher the risk ratio is. Figure 14 illustrates the relative risk ratio (RRR) of each of the independent variables.

**Figure 14: Forest plot describing the relative risk ratio of each of the twenty-seven independent variables for a victim suffering a cumulative CCHI score of 500+ during the 365-day follow-up**



Fourteen of the variables assessed displayed not statistically significant estimates (95% confidence interval), with the “whiskers” crossing the line of no effect (1 on the x axis) (Ellis, 2010). Most of these variables also display a weak effect size and as such are unlikely to be predictors of a victim suffering future high harm.

*Victim and Suspect Crime History*

It has been said that prior crime is the best predictor variable to assist police agencies in identifying where or to whom future crime may occur (Pease, 1998) and this assertion is once again demonstrated in this study, with prior victimisation showing as the



one common feature between most of the statistically significant risk factors observed. The only variable to show a statistically significant increased risk that is not linked to a prior high count of victimisations or prior high cumulative CCHI score is “female victim” with a risk ratio of 1.6, further highlighting that when taking no other factors into account other than the sex of the victim, stalking is a crime that causes a disproportionately high level of harm to women and girls.

The two most powerful risk factors of future high harm for victims observed in this study are “Victim and suspect both have historic CCHI score of 500+” (RRR = 2.2) and “Victim has historic CCHI score of 500+” (RRR = 2.2). Each of these variables have strong effect sizes and they are statistically significant. Given that this is a cumulative score across all prior victimisation and offending, the indication that these victims are 120% more likely to suffer crimes amounting to a cumulative CCHI score of 500+ in a single year following a stalking event is striking.

The count of historic victimisations has been broken down into intervals of ten, with historic crime counts of 30-39, 20-29 and 40+ making up three of the top six variables for effect size. Conversely, a historic crime count of zero for victims was found to be the most statistically significant variable for a decreased risk, with a risk ratio of 0.58. This finding indicates that a victim is 42% less likely to become a high harm victim in the follow-up period than all other victims that have been subjected to at least one prior victimisation. A prior victimisation count of 1-9 also produced a ratio indicating a decreased risk (RRR = 0.7).

### *Professional Judgement – Domestic Abuse and Prior Stalking*

Variables that may be instinctively assumed to carry a heightened level of risk by police practitioners have not necessarily been found to do so in this study. The link between stalking and domestic abuse is well known, yet the presence of a domestic abuse marker on the initial stalking record produced a weak, non-statistically significant effect size. The only exception to this is when the presence of domestic abuse was combined with both the victim and suspect having historic CCHI scores of 500+, in which case an effect size of 2.1 was observed. Similarly, it could be expected that any history of stalking is likely to indicate an increased risk of future high harm, especially so considering what this study has found in respect of the harm levels suffered by victims who go on to be stalked at least once more during the follow-up year. For both “victim prior stalking victim” and “suspect prior stalking offence”, the effect size is 1.08 and 1.03 respectively. It should be noted, however, that the confidence interval margin for both variables is large, with an upper margin of 1.9 and 1.88 respectively, indicating that further research with a larger sample may be of benefit.

### *Length of Offence Before Reporting*

Stalking victims often suffer for some time prior to reporting the matter to the police. One of the variables assessed sought to identify whether this delay in reporting led to more victim harm in the future. This study found that a delay of 31+ days has no effect on the risk of the victim coming to future high harm with a non-statically significant effect size of 1.04.

### *Victim Age*

Finally, victim age was assessed at varying intervals. Despite all these findings being non-statistically significant, different effect sizes across the age range intervals were observed. This followed a general trend of older victims being at less risk, with 56-65 year olds (RRR = 0.51) and 66+ (RRR = 0.69) making up two of the four lowest risk indicators, and younger victims being at greater risk with 0-17 year olds the highest risk age-based variable (RRR = 1.57). This variation in effect size is of interest and future research is recommended to examine this further, particularly in respect of the differing modus operandi that may be deployed by suspects of different ages and the impact this has on harm.

### *Random Forest (calculations by Valdebenito, 2022)*

A random forest model was built using the study data to develop a model for forecasting which individual victims would go on to become a high-harm victims during the follow-up period. If successful, this model could be operationalised in a live policing environment to assist officers in making decisions on which “power few” victims and suspects to prioritise. Figure 15 illustrates the forecasting accuracy of the random forest model in the testing data, whilst the results from the training data can be found at Appendix 5.

**Figure 15: Random Forest Model Based on Testing Data**

	Actual high	Actual low	Totals	Percent
Forecast high	29 true positive	105 false positive	134	
Forecast low	34 false negative	405 true negative	439	
Totals	63	510		
Percent	46% (sensitivity)	79%(specificity)		
Total per cent of forecasts that are accurate in each box within 365 days				76%
Cautious error/Dangerous error ratio				1:3

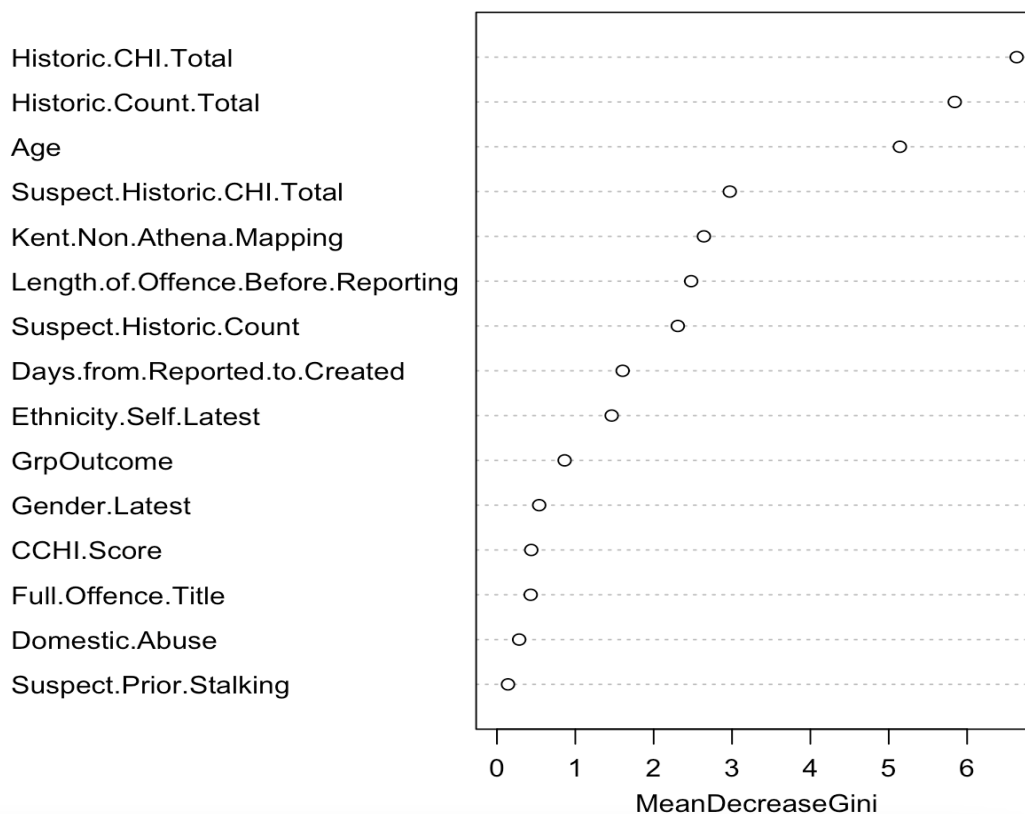
The model based on testing data forecast victim future harm within the follow-up period at a rate of 76%. Future harm was categorised differently for this analysis than for the rest of the analyses in this study, with only two categories: high harm (cumulative harm CCHI score of 500+); and low harm (cumulative CCHI score of 499 and below). The random forest model was more accurate at predicting low harm cases with a 79% specificity when compared with high harm cases with a 46% sensitivity, the reason for this being the small number of high harm cases (n=211). Imbalanced data sets are common; however, machine learning approaches have the greatest potential in overcoming this by down-sampling the majority classification within the data by weighting the minority classification, meaning that the decision trees are grown on a more balanced dataset (Huh et al, 2021).

When looking at the types of errors made for either of the harm categories, the model forecast false positives, i.e. cautious errors that are likely to lead a police force to provide an unnecessarily high level of service to a victim based on their actual risk of future harm, at a rate of 3 to 1 of the false negatives, i.e. the dangerous errors where a victim is assessed as being low risk, when in fact they are high risk. This ratio is more favourable than the alternative, however the high degree cautious errors would lead to a significant amount

of misplaced effort for the limited police resources available. This does mean that this is not currently a viable predictive model for operationalising in a live police environment.

Further to the model itself, the Mean Decrease Gini has been assessed to measure the strength of each of the 15 predictors in respect of their importance in estimating the risk of future victim high harm in the random forest model. Figure 16 shows the 15 predictor variables used to develop the random forest model ranked in order of importance.

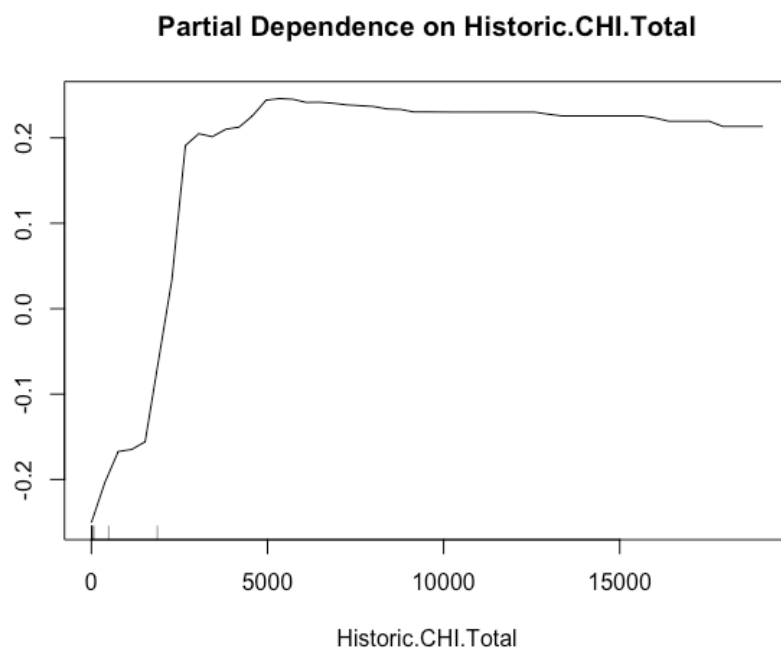
**Figure 16: Predictor variable importance ranked based on the Mean Decrease in Gini Value**



This chart illustrates that the random forest model mirrors the risk ratio forest plot by considering the historic cumulative CCHI score for the victim and the historic victim crime count as the two most important predictors of future high harm. In this model, those two variables were closely followed by age of the victim. A significant gap is then observed

between those three variables and the other twelve, indicating that the importance of those twelve variables as a predictive factor is a great degree lower than the first three. The impact of the strongest predictors (Victim historic CCHI total and Age) is illustrated in Figures 17 and 18 below.

**Figure 17: Partial Dependency Plot for Historic CCHI Total**



**Figure 18: Partial Dependency Plot for Victim Age**

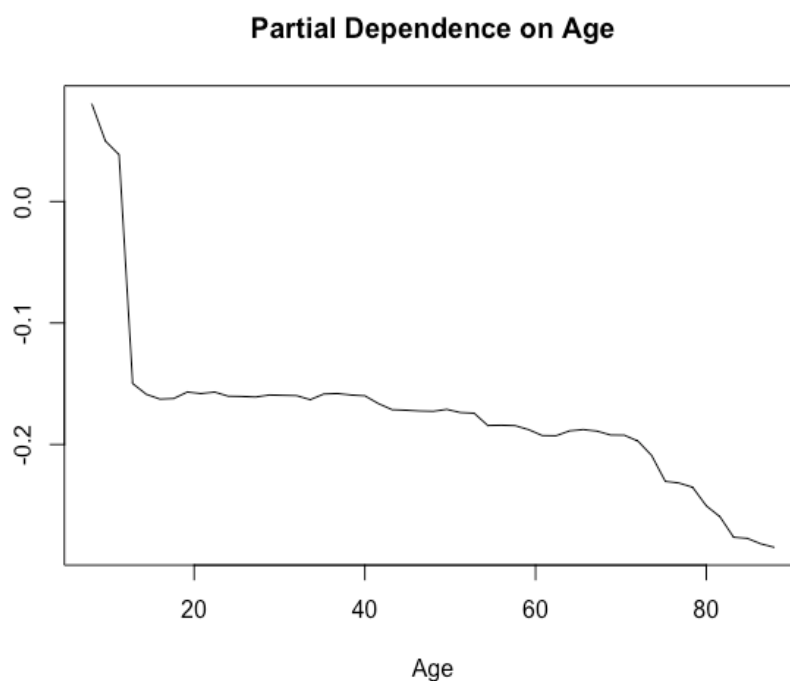


Figure 17 demonstrates that the once a threshold historic CCHI score for a victim is reached, the risk of future high harm remains consistently high. Figure 18 demonstrates that in respect of age, the risk of future high harm reaches a consistent threshold at late teens, initially decreasing with a gentle incline on the chart before a more significant reduction in risk is observed from approximately age 70 onwards. This indicates that younger victims are generally at a higher risk of future high harm than older victims.

The two least important variables in respect of predicting the risk of future harm were if there was a domestic abuse marker present on the initial stalking case, and if the suspect had a history of being a suspect of a prior stalking case in Kent. This once again challenges the intuitive professional judgements of police professionals.

## Discussion

### *Introduction*

The results presented in the previous chapter provide some interesting discussion points regarding what academia and law enforcement knows about stalking and the impact it has on victims. In many ways, the findings reinforce what is already known. However, there are some points of note where the evidence differs from the academic theory and professional judgments held by police officers and other protective agencies that support stalking victims.

The chapter starts by looking at the theoretical implications of these findings, discussing how this research contributes to the existing body of academic literature. It next examines the policy implications and how the evidence presented can be used within an operational context to enhance the way police identify which victims may be most at risk of high harm. It then discusses the research implications for this study and how it may influence future research into the crime of stalking. Finally, the strengths and limitations of the study are examined.

### *Theoretical Implications*

Stalking is a crime defined by its repetitive nature (Farrell & Pease, 2014) and as such one would expect to see a high degree of repeat victimisation across a population of victims. Just 5% of both the victim and suspect populations had a marker for “prior stalking” when entering this cohort, which is perhaps indicative of the low levels of stalking traditionally reported to the police. Whilst repeat victimisation was most certainly a feature within this cohort of victims, with 75% of the population going on to be victimised on at



least one occasion or more during the 365-day follow-up period, the incidence of repeat stalking was only 16%. There are many hypotheses as to why this may be the case, including that the initial police action was sufficient to cause the stalker to desist their behaviour. However, it is equally plausible that further reported crimes may have been part of the same course of conduct as the stalking but not linked within the police database and therefore not recorded as such, or that despite the large increase in reported stalking cases many victims are still not willingly coming forward to report the crime to police.

This study has demonstrated that in a victim population of this size, most victims go on to suffer zero to low recorded harm in the subsequent year post the stalking event. Despite this, whilst no qualitative data was obtained from victims, the quantitative CCHI scores show that medium to high levels of cumulative harm is experienced by almost a third of victims in the follow-up. This is especially true where victims go on to experience stalking that escalates throughout the follow-up year, with a strong association observed between stalking escalation and significant harm. The median cumulative harm score for a victim of escalating stalking in the follow-up year is 286, a staggering twenty-eight times higher than for repeat victims of any crime. Additionally, the cumulative impact of multiple offences can be difficult to quantify in respect of the mental toll it takes on a victim. Victims of escalating stalking were subjected to a mean of five further offences during the follow-up year, 150% more than repeat victims of all other offences. The fear that this repeated victimisation generates cannot be quantified using CCHI but it known from other literature to be significant.

The link between stalking and domestic abuse is well established and this study adds further credence to that theory. With 87% of cases within the population domestic abuse

related, this exceeds some assumptions within the literature (Spitzberg & Cupach, 2007) and supports the findings of Tjaden & Thoennes 1998 study that found domestic abuse stalking is the most common form. The findings from this study do however challenge the notion that victims of domestic abuse related stalking are subjected to a greater degree of harm (Tjaden & Thoennes, 2000), with the presence of a DA marker generating a weak effect size in respect of the risk ratio of being a future high harm victim. It is surmised that the high domestic abuse rate in this study is resultant of the explicit stalking-based questions within the DARA risk assessment which has potentially led to the recording of a significant volume of these offences. Given the frequency with which domestic abuse is now found to be a factor in stalking cases, it is perhaps unsurprising that the links between domestic abuse and high harm are weaker than has been observed in previous studies.

Finally, this research provides further evidence of females being disproportionately affected by stalking, with 80% of victims being recorded as female. This exceeds the figures described in multiple studies, including females being three times more likely to be a victim of stalking than males (Senkans, McEwan & Ogloff, 2017) and females accounting for 75% of victims across a meta-analysis of 103 studies (Spitzberg, 2002). This is a significant finding as much of the research into stalking to date has been gender-biased towards female victims, whereas this research was representative of all victims who came forward to Kent Police to report stalking. In addition to the frequency with which females reporting being victims of stalking, in the assessment of the risk of being a future high harm victim, being female led to a 60% increased risk of being a high harm victim. Conversely, males were found to have a reduced risk.

## *Policy Implications*

The findings for this study are perhaps most relevant to those police forces that have already adopted the DARA for assessing the risk in domestic abuse cases. For those police forces yet to transition from DASH to DARA, this study can act as an indicator of what may be found once the new risk assessment framework is in operational use.

S-DASH is the stalking risk assessment tool used by police forces in England and Wales and is often overlaid with either DASH or DARA. The S-DASH risk gradings assigned to the cases within the study population were a decent indicator of the future risk for victims. Victims assessed as high risk went on to suffer the most harm and most subsequent crimes proportionately, whilst victims assessed as low risk suffered the least harm and least subsequent offences. Despite this, the accuracy of the risk assessments is poor with high levels of false negatives observed within the standard and medium risk victims, and false positives observed within the high-risk victims. This therefore somewhat discredits S-DASH as a tool that is effective at allowing police agencies to identify which victims are at most risk and put the necessary prevention and safeguarding measures in place.

The random forest model developed using this stalking data (Valdebenito, 2022) demonstrated that algorithmic prediction of which victims are most likely to suffer high harm in the subsequent twelve months is possible to some degree. Despite moderate sensitivity of 46% accuracy in forecasting future high harm victims in the testing data, this fares well when compared with the accuracy of human made risk assessments. Table 6 shows that of those victims graded as “high risk” by the attending officer during the initial stalking cases, 42 went on to become high harm victims during the follow-up period, a “true positive” rate of 20%. In this study, the algorithm was therefore 2.3 times more accurate

than the human risk assessments. This raises an interesting discussion related to the acceptance of machine-based predictions in policing, and police forces unwillingness to accept anything other than “perfection”, or at least very close to perfection, despite the inconsistency and inaccuracy of human based risk assessments and the potential improvement that even a model with moderate sensitivity delivers. It should be noted that due to a lack of a control cohort, it cannot be assessed what impact police and other agency interventions had on reducing the level of harm experienced by those victims assessed as high risk.

The variables generating the strongest risk ratios for future high harm for victims in this study almost all related to prior crime victimisation or perpetration. Whilst there may be some information gaps if victims or suspects have moved between force borders, this information will generally be held on local police systems for most individuals. The S-DASH risk assessment process does not consider any police held information and instead is reliant on the information presented by the victim at the time of questioning by the attending officer. Naturally, much of the suspect information may be unknown to the victim and will therefore not be considered when assessing the risk. The findings of this study call into question the reliability of the victim account for being the sole basis of assessing risk and indicates that the addition of police information may bolster the risk assessment to improve its forecasting accuracy.

Furthermore, this study identified a “risk assessment lottery” based on whether the case was domestic abuse related or not. Just 18% of non-domestic abuse cases were subject to a S-DASH risk assessment, meaning that the potential risk of more than 200 victims was left unassessed. This becomes significant when the study has shown that victims who were

not subject to a risk assessment suffered an equivalent amount of harm as those assessed as medium risk. Any discrepancy in police actions between domestic abuse related and non-domestic abuse related stalking should be immediately addressed to ensure that all victims have the same safeguarding and prevention opportunities.

Recorded in the volumes historically seen in police forces in England and Wales, finding resources to attend and investigate stalking cases has not necessarily proven itself to be particularly challenging, especially given that most cases are domestic abuse related which is a national policing priority (HMICFRS 2014). However, with the increase in recorded cases observed since the introduction of the DARA risk assessment in Kent, coupled with the impact of lockdown measures because of the Covid-19 pandemic, forces now face a more challenging decision regarding the resourcing of officers to investigate stalking cases. Prioritisation of those victims who will benefit most from safeguarding and targeting those suspects who will cause the most harm is more important than ever when it comes to stalking.

This study has identified a clear “power few” (Sherman, 2007) of victims and suspects who go on to either be subjected to or perpetrate the most harm within the follow-up period. Within this population, just 78 victims (5.2%) were subject of 50% of the total subsequent harm. Additionally, 79 suspects (5.5%) caused 50% of the total subsequent harm. Considering the overall population of 2006 victims and 1902 suspects, prioritising the “power few” 5% of each of those is more manageable for police forces with limited resources, and the effective targeting of those individuals has the potential to reduce the overall harm suffered by the entire population by 50%. A more nuanced and evidence-based approach could yield significant benefits for police agencies in reducing harm.

## *Research Implications*

This is the first study of its kind to use police recorded data to identify a cohort of stalking victims and suspects and follow those individuals for a consistent 365-day follow-up period, using the CCHI to measure subsequent harm. Most studies into stalking are more qualitative in nature and have provided a rich picture of the impact of stalking on victims, as well as the motivations and drivers of stalking behaviour in suspects. Further research of a quantitative nature such as this would be welcomed, particularly over a longer follow-up period and with a larger sample that is representative of multiple police force areas. By tracking more victims and extending the follow-on period, it will be possible to identify whether a greater number of victims enter the “high harm” cohort as time moves on, or whether there is a “cut off” point at which the chances of being re-victimised significantly diminishes, as research indicates that stalking continues for an average of 22 months (Spitzberg & Cupach, 2007). By looking across county borders, it will allow the identification of stalkers who carry out their crimes against multiple victims, particularly if they do so “remotely” utilising means such as the internet rather than more physical stalking behaviours such as following and leaving gifts. Additionally, further research should be carried out to develop the accuracy of a machine-based forecasting model. Early research into this area utilising the data from this study has shown some promise. With a larger population of victims tracked over a longer period, the effect of the imbalance between low harm and high harm cases may be reduced which will improve the sensitivity of the model.

Due to the small number of positive police outcomes present within the case population (93% cases received an NFA outcome) it has not been possible to draw many conclusions as to the effect that police action has on future harm.

This study cannot explain why the rate of victim non-support for police action is so high for stalking investigations. One hypothesis could be that many cases are now being identified through the overt questioning of stalking type behaviours within the DARA risk assessment. There is a strong possibility that victims are not making contact to police to report a stalking, and they may not see themselves as a victim of stalking. However, on receipt of a positive response to the two stalking based questions, the Home Office Counting Rules dictate that a crime of stalking must be recorded in addition to any other crime being reported, leading to a course of action by police that the victim neither asked for or consented to. The London Rape Review 2021 examined the high victim withdrawal rates in rape investigations and coded the reasons that victims gave for withdrawing their support of the investigation. The joint most common reason for victim withdrawal was found to be “Did not intend to report rape” (MOPAC, 2021). It may be that something similar is occurring in cases of stalking, but rather than rapes being discovered because of DASH questioning, it is instances of stalking being uncovered by DARA questioning.

In addition, the Stalking Protection Act 2019 introduced a new civil order, the Stalking Protection Order (SPO), to allow for early police intervention in stalking cases. It had been intended that the impact of SPOs would be analysed as a variable in this study, however only a handful of SPOs had been issued in Kent by the end of the follow-on period. It would be interesting to include SPOs into any further research into the effectiveness of police interventions on stalking cases, and this will be of particular importance as research from Italy indicates that the prevalence of suspects breaching protective orders is higher than in the official data and occurred in 50% of cases within the first year of the order being granted (Baldry et al, 2016). Further research with a focus on police outcomes in stalking

cases is important as it will help to determine the most evidence based and appropriate police policies for how to effectively deal with stalkers.

This study has provided mixed findings in respect of the impact that repeat and prior stalking victimisation has on the level of future harm suffered by a victim. Acting as a predictor variable, “prior stalking” generated a non-statistically significant weak effect size. However, within the 95% confidence interval a wide range was observed, with the upper range a risk ratio of 1.9. It has been observed that those victims who go on to be subsequent victims of stalking in the follow-up period go on to suffer, on average, more harm than all victims and more harm than repeat victims of all other offences. This is particularly true where the stalking has escalated. One possible hypothesis for the difference between prior stalking as a predictor and subsequent stalking as a measure of actual harm suffered is the lack of police recorded stalking cases in the past. Given the rarity in which stalking was recorded on its own independent crime report, it may be that many more of the victims in this cohort have in fact been stalked in the past, but that it was wrapped up within other recorded crimes under the “Principal Crime Rule” of the Home Office Counting Rules (Home Office, 2021c). With a renewed baseline of stalking crime recording, research into this specific area may find that over time prior stalking victimisation becomes a far stronger predictor of future high harm.

Finally, great value could be gained from combining the qualitative detail understood about stalkers, their motivation, their behaviours, and their typologies with a quantitative study. Whilst suspects have been considered as part of this study and data has been gathered on the impact that their demographics and prior offending has on future harm, this study is primarily victim focused. Victims do not cause stalking. Stalking is purely in the



hands of the perpetrator. It would be of theoretical interest to understand whether different stalker types favour certain modus operandi (MO), and whether the presence of any MO increases the risk of future victimisation. When combined with the risk factors identified within this study, it may create a powerful predictor for identifying who the power few suspects are that warrant the greatest degree of intervention, which will ultimately safeguard more victims from harm.

### *Limitations of the Study*

Whilst Kent Police has a strong crime data integrity rating following independent inspection by Her Majesty's Inspectorate of Constabulary and Fire and Rescue Services (HMICFRS, 2019), it is known that stalking is an under-reported crime to police and as such this study is not representative of all stalking victims in the Kent Police jurisdiction during the study period. Given the significant presence of domestic abuse cases within the population, it is likely that this study is biased towards the domestic abuse variable more so than a self-selecting sample would be.

This study has looked at only Kent Police data, meaning that for both the victims and suspects within the population, the data may not be fully reflective of their police involvement both prior to and after the original stalking case in this study. It is not possible to determine the transience of the victim and suspect populations, however given that the follow-up period was restricted to one year, for which a significant proportion the UK population was "locked down" because of Covid-19 restrictions, any movement of people is expected to have been limited.

Most police forces in England and Wales are not currently utilising the DARA for assessing the risk to domestic abuse victims, and DARA is only in use in England and Wales.

As a result, the external validity of these findings to non-DARA forces and forces outside of British policing may be weakened. It is hypothesised that DARA has in-part led to the sharp increase in the number of stalking cases recorded in Kent Police, and other policing jurisdictions may have an even higher number of unknown victims that would not be represented in a study sample in that force area.

The analysis is only as good as the quality of data. Whilst the number of missing variables on records was generally low, for some variables such as ethnicity, the data quality is poor, and no analysis could be carried out. Additionally, with human crime recording decisions being made, there is the potential that subsequent crimes recorded against the victim may have been part of the stalking course of conduct, without the link being made by the member of staff recording the crime.

### *Strengths of the Study*

This study was not subject to many of the biases that frequent stalking research. Victims did not self-select to be part of a study; they were instead coming to police to report a crime had occurred. This meant that victims, and in turn suspects, would be representative of a variety of demographics which may make the findings more externally valid to other jurisdictions. Many of the existing stalking studies have taken place in colleges and universities which biases towards younger people, whilst many of the studies into suspects have taken place within psychiatric or other clinical environments which biases towards those suspects where mental ill health will likely be a significant factor in their offending.

The study had strong internal validity. By conducting a prospective longitudinal analysis of the entire population of victims, no cases or individuals were discriminated

against and all data from the study period was captured and analysed. There was no opportunity for victims or suspects to “opt out” midway through the study which means that all individuals were subject to a consistent 365-day follow-up period.

Finally, given that it is police data that has been used which is subject to a consistent set of recording rules and principles, these findings are at least in part externally valid to all police forces in England and Wales. This is particularly true of those other police forces who have adopted the DARA, where a consistent risk assessment framework as well as consistent crime recording principles are in use.

## Conclusions

Stalking is a crime that is rapidly being recorded more and more frequently in England and Wales. In Kent, where this study was undertaken, a 605% increase in recorded stalking cases was observed between 2018 and 2020. What was previously 'hidden' demand, often going unreported by victims (Jerath et al, 2020), is now more out in the open than ever before. With independent inspections identifying that the police and the wider criminal justice system need to be more effective in tackling stalking (HMICFRS, 2017), it is vitally important that further research is carried out to improve the service that victims receive.

Early academic research into stalking tended to focus more on high profile victims and suspects with psychiatric illness (Rosenfeld, 2004), but over time the focus has shifted to victim-offender relationships and the community level harms that stalking causes (Sheridan et al, 2001a). Stalking is strongly associated with domestic abuse (Tjaden & Theoness, 1998) and violence (Purcell, Pathé & Mullen, 2001) and research shows that it can have a significant impact on victims' lives, causing them psychological and financial harms (Reyns & Englebrecht, 2010; DuMonthier et al, 2017). The impact of the stalking can result in victims making big lifestyle decisions such as moving address or changing their job (Jason et al, 1984; Pathé & Mullen, 1997). These findings were often discovered in studies biased towards certain victim or suspect types brought about either by the self-selecting nature or location of the study.

There is no single set definition of stalking across academia or within law enforcement, and this may be part of the cause of the challenges that the wider criminal justice system faces in dealing with stalking. The joint HMICFRS and HMICPSI inspection in

2017 noted *“There is not an exhaustive definition of stalking in legislation. This is one of the main reasons why there is a lack of common understanding about which actions can be counted as stalking. At present, identifying stalking is frequently a matter of subjectivity, which can lead to error and/or omission...”* With these words in mind, along with the ever-increasing number of stalking cases recorded by police forces in England and Wales, it is important that police forces understand the profile of stalking victims and suspects and more effectively forecast which victims are at most risk of high harm as a result of the stalking they are subjected to, in order to use the limited police resources more wisely to prevent crimes and reduce the amount of harm that victims suffer.

This study sought to improve the knowledge within policing to provide a baseline profile for stalking within a police force in England and Wales that utilises the DARA to risk assess victims of domestic abuse, improving the external validity of the findings to those police forces comparable to Kent Police. The total population of stalking cases recorded between 1<sup>st</sup> November 2019 and 31<sup>st</sup> July 2020 (n=2006) were captured, with a full profile completed to show factors such as the initial stalking case recorded, the prevalence of domestic abuse present within the population, the police recorded outcome assigned to the case and the number of days that had passed between the crime first occurring and it being reported to the police.

This profiling reinforced many of the factors that existing stalking research has identified. 87% of cases were domestic abuse related, which is a stronger association with domestic abuse than is generally identified elsewhere within the literature and may point towards the proactivity of Kent Police to identify stalking cases by using the DARA. 80% of the victims were female, confirming that stalking is a crime which disproportionately

impacts on women and girls. The stalking had been occurring for a median of 11 days prior to being reported to police, with an average observed of 159 days. This reinforces the notion that stalking is not immediately reported to police and victims are likely to suffer several incidents prior to making the initial report, however it does indicate that it is unlikely that the 100 incidents said to occur prior to reporting (Paladin, 2017) is an accurate figure.

A strong power law (Sherman, 2007) was identified, demonstrating that just over 5% of victims and suspects either suffer or perpetrate 50% of total cumulative harm in the following year. 10.5% of victims go on to suffer high harm, defined in this study as a cumulative CCHI score of 500+, i.e., the equivalent to 500 or more imprisonable days for the suspect.

Across both the relative risk ratio analysis and the random forest modelling, prior victimisation was the most predictive variable for a victim going on to suffer subsequent cumulative high harm. The suspect's offending history was also a good predictor of future high harm. Variables that would be assumed to be predictors of high risk, such as the presence of domestic abuse, were found to have little effect on the risk of future high harm in a population of this size. Subsequent stalking was found to be correlated to significant levels of future harm for victims, particularly if the stalking had escalated in terms of severity.

These findings have both theoretical and research implications. Whilst many of the current theoretical assumptions have been borne out in this study, these findings do indicate that further research is required with larger victim samples to identify the most prevalent risk factors. It is also recommended that further research combines a qualitative approach towards suspect typology and modus operandi combined with the quantitative

approach demonstrated within this study. This will allow for a layered approach to risk profiling that combines intuitive police judgement, recorded police intelligence and academic theory.

Finally, the findings from this study should influence police policy. An immediate improvement to stalking risk assessments can be made by combining the victim account given to the initial attending officer during S-DASH questioning with the recorded police information and intelligence on the victim and suspect. This study indicates that prior victimisation and offending are the strongest predictors of future high harm and by including this information within the risk assessment, police will be able to predict more effectively which victims are at the highest risk and are in most need of enhanced safeguarding measures. Additionally, police forces should immediately ensure that all victims of stalking, whether domestic abuse related or otherwise, receive a consistent risk assessment. This study found that victims who receive no risk assessment go on to suffer medium levels of harm in the subsequent year (n = CCHI 20-499), some of which may be prevented with an effective risk assessment allowing for earlier interventions to be made.

## References

- BBC News (2021) 'Emily Maitlis stalker says late night visit was clumsy'. Retrieved from <https://www.bbc.co.uk/news/uk-england-nottinghamshire-59085509> on 21/01/2022.
- Baldry, A. C., Cinquegrana, V., Cacace, S. and Crapollicchio, E. (2016) Victim's perception of quality of help and support by the police issuing warning orders in ex-intimate partner stalking cases in Italy. *Policing: A Journal of Policy and Practice*, 10:432-445.
- Baum, K., Catalano, S. and Rand, M. (2009) Stalking victimisation in the United States. *Bureau of Justice Statistics Special Report*. US Department of Justice.
- Berk, R. (2012) 'Tree Based Forecasting Methods' in *Criminal Justice Forecasts of Risk: A Machine Learning Approach*, SpringerBriefs in Computer Science, Springer, New York, NY, pp. 59-79.
- Boon, J. C. W and Sheridan, L. (2001) Stalker typologies: A law enforcement perspective. *Journal of Threat Assessment* 1: 75-97.
- Boehnlein, T., Kretschmar, J., Regoeczi, W. and Smialek, J. (2020) Responding to stalking victims: Perceptions, barriers, and directions for future research. *Journal of Family Violence*, 35:755-768.
- Brady, P. Q. and Nobles, M. R. (2017) The Dark Figure of Stalking: Examining Law Enforcement Response. *Journal of Interpersonal Violence* 32(20) 3149–3173.
- Breiman, L. (2001) Random Forests. *Machine Learning*, 45:5-32.
- Brewster, M. P. (2000). Stalking by former intimates: Verbal threats and other predictors of physical violence. *Violence and Victims* 15:41-54.
- Caruana, E. J., Roman, M., Hernández-Sánchez, J. and Solli, P. (2015) Longitudinal studies. *Journal of Thoracic Disease* 7(11) 537-540.
- Churcher, F. P. and Nesca, M. (2013). Risk factors for violence in stalking preparation: A meta-analysis. *FWU Journal of Social Sciences* 2013 7(2) 100-112.
- College of Policing (2020). Stalking or harassment. Advice for investigators on effective investigation. Retrieved from <https://library.college.police.uk/docs/appref/Stalking-or-harassment-advice-for-investigators.pdf> on 04/08/2021.
- Crown Prosecution Service (CPS) (2018) Stalking and Harassment, Retrieved from <https://www.cps.gov.uk/legal-guidance/stalking-and-harassment> on 06/02/2022
- Ellis, P. D. (2010). The Essential Guide to Effect Sizes. Statistical Power, Meta-Analysis, and the Interpretation of Research Results. Cambridge University Press.



Farrell, G. and Pease, K. (2014) Repeat victimisation. *Encyclopedia of Criminology and Criminal Justice* 1:4371-4381.

Fein, R., Vossekuil, B. and Holden, G. (1995). *Threat assessment: An approach to prevent targeted violence*. Washington, DC: U.S. Department of Justice.

Geberth, V. J. (1992) Stalkers. *Law and Order*, 40(10), 138-143

Gill, R. and Brockman, J. (1996) *A review of Section 264 (criminal harassment) of the Criminal Code of Canada*. Ottawa, Canada: Department of Justice Canada.

Google (2022) 'Google Trends, Stalking'. Retrieved from <https://trends.google.com/trends/explore?date=all&geo=GB&q=stalking> on 21/01/2022.

Gregson, C. B. (1998) California's antistalking statute: the pivotal role of intent. *28 Golden Gate University Law Review* 28(2) 221-263.

Grove, L., Farrell, G., Farrington, D. and Johnson, S. (2012) Preventing Repeat Victimization: A Systematic Review. Produced for Brå – The Swedish National Council for Crime Prevention.

Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS) (2014) Everyone's business: improving police response to domestic abuse. Retrieved from <https://www.justiceinspectors.gov.uk/hmicfrs/publications/improving-the-police-response-to-domestic-abuse/> on 16/01/2022.

Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS) (2019). Kent Police: Crime data integrity re-inspection 2018. Retrieved from <https://www.justiceinspectors.gov.uk/hmicfrs/publications/kent-crime-data-integrity-re-inspection-2018/> on 05/08/2021.

Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services (HMICFRS) and Her Majesty's Crown Prosecution Service Inspectorate (HMPCSI) (2017). Living in fear – the police and CPS response to harassment and stalking, a joint inspection by HMIC and HMPCSI. Retrieved from <https://www.justiceinspectors.gov.uk/hmicfrs/wp-content/uploads/living-in-fear-the-police-and-cps-response-to-harassment-and-stalking.pdf> on 15/07/2021.

Home Office (2021a) Home Office counting rules for recorded crime: Violence against the person, Retrieved from [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/992839/count-violence-jun-2021.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/992839/count-violence-jun-2021.pdf) on 21/01/2022.

Home Office (2021b) *Tackling Violence Against Women and Girls Strategy*, London, H.M.S.O.

Home Office (2021c) Home Office counting rules for recorded crime: General Rules, Retrieved from

[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/992833/count-general-jun-2021.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/992833/count-general-jun-2021.pdf) on 06/02/2022

Huh, K. (2021) Surviving in a random forest with imbalanced datasets, Retrieved from <https://medium.com/sfu-csmp/surviving-in-a-random-forest-with-imbalanced-datasets-b98b963d52eb> on 06/02/2022

Jason, L. A., Reichler, A., Easton, J., Neil, A., & Wilson, M. (1984). Female harassment after ending a relationship: A preliminary study. *Alternative Lifestyles*, 6, 259–269.

Jerath, K., Tompson, L. and Belur, J. (2020). Risk Management in Stalking Victims: A Multi-Agency Approach to Victim Advocacy. *Journal of Interpersonal Violence*.

Jordan, C. E., Wilcox, P. and Pritchard, A. J. (2007) Stalking acknowledgment and reporting among college women experiencing intrusive behaviours: Implications for the emergence of a “class stalking case”. *Journal for Criminal Justice*, 35:556-569.

Kienlen, K. K., Birmingham, D. L., Solberg, K., O’Regan, J. T. and Meloy, J. R. (1997) A comparative study of psychotic and nonpsychotic stalkers. *The Journal of the American Academy of Psychiatry and the Law*, 25(3): 317-334

Kropp, P. R., Hart, S. D. and Lyon, D. R. (2002). Risk assessment of stalkers - Some problems and possible solutions. *Criminal Justice and Behavior* 29(5): 590-616.

Liaw, A. and Wiener, M. (2002) Classification and regression by randomForest, *R News*, 2(3) 18-22.

Mayor of London Office for Police and Crime (MOPAC) (2021). The London Rape Review 2021: An examination of cases from 2017 to 2019 with a focus on victim technology. Retrieved from [https://www.london.gov.uk/sites/default/files/final\\_rr\\_victimtech\\_61221.pdf](https://www.london.gov.uk/sites/default/files/final_rr_victimtech_61221.pdf) on 09/02/2022

McKeon, B., McEwan, T. E. and Luebbers, S. (2015) It’s not really stalking if you know the person: Measuring community attitudes that normalize, justify and minimise stalking. *Psychiatry, Psychology and Law*, 22(2) 291-306.

McEwan, T. E., Mullen, P., MacKenzie, R., D. and Ogloff, J. R. P. (2009) Violence in stalking situations. *Psychological Medicine* 39(9):1469-78.

McFadzien, K., Pughsley, A., Featherstone, A. M. and Phillips, J. (2020) The evidence-based investigative tool (EBIT): a legitimacy-conscious statistical triage process for high-volume crimes, *Cambridge Journal for Evidence Based Policing*, 4:218-232.

Monckton-Smith, J., Szymanska, K. and Haile, S. (2017). Exploring the relationship between stalking and homicide. *Homicide Research Group, University of Gloucestershire*.

Monckton-Smith, J. (2020). Intimate partner femicide: using foucauldian analysis to track an eight-stage progression to homicide. *Violence Against Women* 26(11):1267-1285.

Mullen, P. E., Pathé, M. and Purcell, R. (2000). *Stalkers and their victims*. Cambridge, UK: Cambridge University Press.

Mullen, P. E., Pathé, M. and Purcell, R. (2009). *Stalkers and their Victims, 2nd Edition*. Cambridge, UK: Cambridge University Press.

Nissen, T. and Wynn, R. (2014) The clinical case report: a review of its merits and limitations, *BMC Res Notes*, 7:264.

Norris, S. M., Huss, M. T. and Palarea, R. E. (2011). A pattern of violence: analysing the relationship between intimate partner violence and stalking. *Violence and Victims* 26(1):103-115.

Office for National Statistics (2020). Domestic abuse in England and Wales overview: November 2020. Retrieved from <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/domesticabuseinenglandandwalesoverview/november2020> on 05/08/2021.

Paladin (2017) *Stalking, The Legislation and Advocacy*. Presentation delivered to the Northern Ireland Assembly by Laura Richards on 19<sup>th</sup> January 2017. Unpublished, retrieved from <http://www.niassembly.gov.uk/globalassets/documents/justice/19-january-2017-presentation-by-laura-richards-paladin.pdf> on 01/02/2022

Palarea, R. E., Zona, M. A., Lane, J. C. and Langhinrichesen-Rohling, J. (1999). The dangerous nature of intimate relationship stalking: Threats, violence, and associated risk factors. *Behavioural Sciences & the Law* 17(3) 269-283.

Pathé, M. and Mullen, P. E. (1997). The impact of stalkers on their victims. *British Journal of Psychiatry* 170 12-17.

Patton, C. L., Nobles, M. R. and Fox, K. A. (2010). Look who's stalking: Obsessive pursuit and attachment theory. *Journal of Criminal Justice* 38(3) 282-290.

Pease, K (1998) Repeat victimisation: Taking stock. Produced for the Home Office Police Research Group. *Crime Detection and Prevention Series*, Paper 90.

Purcell, R., Pathé, M. and Mullen, P. (2001). A study of women who stalk. *American Journal of Psychiatry* 158:2056-2060.

Purcell, R., Pathé, M. and Mullen, P. (2004) Editorial: When do repeated intrusions become stalking?. *Journal of Forensic Psychiatry and Psychology* 15(4) 571-583.

Ranganathan, P., Aggarwal, R. & Pramesh, C. S. (2015) Common pitfalls in statistical analysis: odds versus risk. *Perspectives in Clinical Research* 6(4), 222-224.

Reyns, B. W. and Englebrecht, C. M. (2010). The stalking victim's decision to contact the police: A test of Gottfredson and Gottfredson's theory of criminal justice decision making. *Journal of Criminal Justice* 38(5) 998-1005.

Reyns, B. W., Henson, B. and Fisher, B. S. (2011). Being Pursued Online Applying Cyberlifestyle-Routine Activities Theory to Cyberstalking Victimization. *Criminal Justice and Behavior* 38(11) 1149-1169.

Richards, L. (2003). Findings from the Multi-Agency Domestic Violence Murder Reviews prepared for ACPO Homicide Working Group. London: Metropolitan Police Service.

Richards, L. (2009) Domestic Abuse, Stalking and Harassment and Honour Based Violence (DASH) risk checklist. Retrieved from <https://www.dashriskchecklist.co.uk> on 02/12/2021.

Rosenfeld, B. and Harmon, R. (2002). Factors associated with violence in stalking and obsessional harassment cases, *Criminal Justice and Behavior* 29(6) 671-691.

Rosenfeld, B. (2004). Violence risk factors in stalking and obsessional harassment - A review and preliminary meta-analysis. *Criminal Justice and Behavior* 31(1) 9-36.

Schwartz-Watts, D. and Morgan, D. W. (1998). Violent versus nonviolent stalkers. *Journal of the American Academy of Psychiatry and the Law* 26(2) 241-245.

Senkans, S., McEwan, T. E. and Ogloff, J. R. P. (2021) Assessing the Link Between Intimate Partner Violence and Postrelationship Stalking: A Gender-Inclusive Study. *Journal of Interpersonal Violence* 36(1-2) 772-802.

Sheridan, L., Gillett, R. & Davies, G. (2000) Stalking - seeking the victim's perspective. *Psychology Crime & Law* 6(4) 267-280.

Sheridan, L., Davies, G. M. and Boon, J. C. W. (2001a) Stalking - Perceptions and prevalence. *Journal of Interpersonal Violence* 16(2) 151-167.

Sheridan, L., Davies, G. and Boon, J. (2001b) The Course and Nature of Stalking: A Victim Perspective. *The Howard Journal of Criminal Justice* 40:215-234.

Sherman, L. W. et al. (1991). From initial deterrence to long-term escalation – short-custody arrest for poverty ghetto domestic violence. *Criminology* 29(4) 821-850.

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): 171-183.

Spitzberg, B. H. (2002) The Tactical Topography of Stalking Victimization and Management. *Trauma, Violence, & Abuse* 3(4) 261-288.

Stefanska, E. B., Longpre, N. and Harriman, R. S. (2021). The Wheel of Change Moves On: Assessing the Severity of Stalking Behavior. *Journal of Interpersonal Violence*.

Taylor, S., Barnes, J. and Short, E. (2019) Health care responses to stalking: Implications and recommendations. Produced by Sussex Stalking Support and the National Centre for Cyberstalking Research, University of Bedfordshire, in association with the National Stalking Consortium. Retrieved from <https://www.suzylamplugh.org/health-care-responses-to-stalking-implications-recommendations> on 14/12/2021.

Taylor-Dunn, H., Bowen, E. and Gilchrist, E. A. (2021). Reporting Harassment and Stalking to the Police: A Qualitative Study of Victims' Experiences. *Journal of Interpersonal Violence* 36(11-12) NP5965-NP5992.

Tjaden, P., & Thoennes, N. (1998). Stalking in America: Findings from the National Violence Against Women Survey. Washington, DC: U.S. Department of Justice.

Tjaden, P., & Thoennes, N. (2000) The role of stalking in domestic violence crime reports generated by the Colorado Springs Police Department. *Violence and Victims*, 15:427-441.

Turmanis, S. A. and Brown, R. I. (2006). The Stalking and Harassment Behaviour Scale: Measuring the incidence, nature, and severity of stalking and relational harassment and their psychological effects. *Psychology and Psychotherapy-Theory Research and Practice* 79 183-198.

Valdebenito, S. (2022) *Random Forest Project: Tracking Stalking for High Harm Outcomes: A 365-Day Follow-Up Analysis*. Internal University of Cambridge report. Unpublished.

Wire, J. and Myhill, A. (2018) Piloting a new approach to domestic abuse frontline risk assessment: Evaluation report. College of Policing. Retrieved from [https://whatworks.college.police.uk/research/documents/da\\_risk\\_assessment\\_pilot.pdf](https://whatworks.college.police.uk/research/documents/da_risk_assessment_pilot.pdf) on 30/11/2021.

# Appendices

## Appendix 1 – Domestic Abuse Risk Assessment (DARA)

	Never	Occasionally	Often	All the time	Not stated							
1. How often does (...) make threats to harm you or things you care about such as people, pets or property?												
2. How often does (...) call you names, humiliate or degrade you?												
3. How often does (...) control your daily activities, such as how you dress or how perform household tasks?												
4. How often does (...) deny you access to money, or control what you can spend it on?												
5. How often does (...) make you account for where you have been, or monitor your phone, email or social media to check up on you?												
6. How often does (...) follow or stalk you, or try to contact you when you do not want them to?												
7. How often do you feel isolated or like you have no one to turn to for support?												
	Never	Once only	Occasionally	Often	Not stated							
8. How often does (...) use physical violence towards you, such as pushing, slapping, punching or kicking?												
9. How often does (...) attempt to strangle, choke, suffocate or drown you?												
10. How often does (...) use or threaten to use weapons such as household items, knives or guns to hurt you?												
11. How often does (...) threaten or attempt suicide?												
			Yes	No	Not stated							
12. Have you recently separated from (...), or do you plan to separate from them?												
<b>[If yes at Q12]</b> 12a. Has this/will this put you in danger?												
13. Is the abuse you are experiencing from (...) getting worse?												
14. Has (...) ever threatened to kill you and you believed they were capable of doing it?												
15. Has (...) ever hurt the children?												
16. Does (...) use child contact arrangements to control you or continue to abuse you?												
	0	1	2	3	4	5	6	7	8	9	10	Not stated
17. On a scale of 0 to 10, where '0' is not at all likely and '10' is very likely, how likely do you think it is that (...) will seriously injure you in the future?												

## General assessment of the threat posed to the victim

Using the information you've collected and know about the victim and perpetrator, and your professional judgement, please provide a summary of the threat posed to the victim by the perpetrator. Specifically consider:

- How scared you perceive the victim to be and the immediate threat posed by the perpetrator
- Whether you think the victim feels they cannot, or is being prevented from, disclosing the abuse
- The level of control the perpetrator has over the victim
- Any aggravating factors that might lead the perpetrator to commit serious harm against the victim e.g. they have recently separated or the victim plans to leave; the victim is pregnant or has a very young child; an injunction has expired or is due to expire; any **significant** mental health, alcohol or substance abuse issues
- Evidence of harmful cultural practices such as Honour Based Abuse (HBA); forced marriage.

## Initial assessment of threat

Standard  Medium  High

<b>Standard</b>	There appears to be no pattern of abusive behaviour, or control of one person by another. Current evidence does not indicate likelihood of causing serious harm.
<b>Medium</b>	There appears to be a pattern of abuse/control of one person by another, and/or frequent physical violence. The offender has the potential to cause serious harm but is unlikely to do so unless there is a change in circumstances, such as the victim attempting to leave.
<b>High</b>	There is an extreme level of control of one person by another and/or very frequent and severe physical violence. There is a serious threat of harm posed to the victim by the perpetrator. The potential event could happen at any time and the impact would be serious.

Risk of serious harm (Home Office 2002 and OASys 2006): 'A risk which is life threatening and/or traumatic, and from which recovery, whether physical or psychological, can be expected to be difficult or impossible.'

## Appendix 2 – Domestic Abuse, Stalking and Harassment and Honour Based Violence

### (DASH) Question Set

1. Has the current incident resulted in injury? (Please state what and whether this is the first injury)
2. Are you very frightened?
3. What are you afraid of? Is it further injury or violence? (Please give an indication of what you think (name of abuser(s) ..... might do and to whom))
4. Do you feel isolated from family/friends ie: does (name of abuser(s) ....) try to stop you from seeing friends/family/Dr or others?
5. Are you feeling depressed or having suicidal thoughts?
6. Have you separated or tried to separate from (name of abuser(s).....) within the past year?
7. Is there conflict over child contact?
8. Does (.....) constantly text, call, contact, follow, stalk or harass you?
9. Is anyone in the household pregnant? (Y/N) If no, has anyone in the household had a child in the last 18 months? (Y?N)
10. Are there any children, step-children or dependents either in the household or elsewhere?
11. Has (.....) ever hurt the children/dependents?
12. Has (.....) ever threatened to hurt or kill the children/dependents?



13. Is the abuse happening more often?
14. Is the abuse getting worse?
15. Does (.....) try to control everything you do and/or are they excessively jealous?
16. Has (.....) ever used weapons or objects to hurt you?
17. Has (.....) ever threatened to kill you or someone else and you believed them?
18. Has (.....) ever attempted to strangle/choke/suffocate/drown you?
19. Does (.....) do or say things of a sexual nature that makes you feel bad or that physically you or someone else?
20. Is there any other person that has threatened you or that you are afraid of?
21. Do you know if (.....) has hurt anyone else? (inside or outside of the family)
22. Has (.....) ever mistreated an animal or the family pet?
23. Are there any financial issues? For example, are you dependent on (...) for money/have they recently lost their job/other financial issues?
24. Has (.....) had problems in the past year with drugs (prescription or other), alcohol or mental health leading to problems in leading a normal life?
25. Has (.....) ever threatened or attempted suicide?
26. Has (.....) ever breached bail/an injunction and/or any agreement for when they can see you and/or the children?
27. Do you know if (.....) has ever been in trouble with the police or has a criminal history?

## Appendix 3 – S-DASH Question Set

### S-DASH Questions to be asked in addition to DASH.

<p>This risk identification can be used in ALL cases of stalking and harassment. It should be completed by professionals if there are two or more incidents of stalking and harassment (reported or unreported) and/or if the victim is extremely frightened. These questions direct you to specific areas that will give you an indication of the victim(s) risk of future violence/harm. Most the behaviours will be about coercive control. Do not think it is any less serious if there has been no physical violence. The more 'yes' answers you have, the higher the risk that the suspect could physically attack the victim at any time. Please ensure that you write the additional notes about the context of what is going on and link the risk identification responses to a risk management/safety plan.</p>		
	<b>Y</b>	<b>N</b>
<b>1. Is the victim very frightened?</b>		
<b>2. Has (Insert name of abuser(s) ..... ) engaged in harassment on previous occasion(s)?</b> (this victim and/or other victims)		
<b>3. Has (Insert name of abuser(s).....) ever destroyed or vandalised the victim's property?</b>		
<b>4. Does (name of abuser(s).....) visit the victim at work, home, etc., more than three times per week?</b>		
<b>5. Has (.....) loitered around the victim's home, workplace etc?</b>		
<b>6. Has (.....) made any threats of physical or sexual violence in the current harassment incident?</b>		
<b>7. Has (.....) harassed any third party since the harassment began? (e.g. friends, family, children, colleagues, partners or neighbours of the victim)</b>		
<b>8. Has (.....) acted out violently towards people within the current stalking incident?</b>		
<b>9. Has (.....) persuaded other people to help him/her? (wittingly or unwittingly)</b>		
<b>10. Is (.....) known to be abusing drugs and/or alcohol?</b>		
<b>11. Is (.....) known to have been violent in the past? (This could be physical or psychological. Intelligence or reported)</b>		
<p>Other relevant information/additional observations made by Practitioner (e.g. level of fear in victim, details of threats and violence, duration of harassment, various harassing behaviours engaged in by abuser, victim's beliefs concerning abuser's motives, weapons owned by abuser, nature of unwanted 'gifts'/items left for victim, attitude/demeanour of abuser including mental health issues and whether victim has responded in any way to the abuser)</p>		

## Appendix 4 – Random Forest Data Dictionary

Variable	Role	Description	% Missing
SubChiTotal	Outcome	Total cumulative CCHI score that victim suffered during 365-day follow-up in Kent	0%
Sub Count Total	Outcome	Total cumulative crime count that victim suffered during 365-day follow-up in Kent	0%
C.CHI.Score	Predictor	CCHI score of the original stalking offence	0%
FullOffenceTitle	Predictor	Original stalking offence title	0%
KentNonAthenaMapping	Predictor	District in which the crime occurred	0%
GrpOutcome	Predictor	Description of the outcome code assigned to the investigation on completion	.8%
DaysfromReportedtoCreated	Predictor	Days that passed between the victim/other reporting the crime to Kent Police to Kent Police creating a substantive crime report on Athena	0% missing
LengthhofOffenceBeforeReporting	Predictor	Days that passed between the lower committed date (date on which offending first occurred) and upper committed date (date on which offending last occurred)	.6%
DomesticAbuse	Predictor	True / False of domestic abuse flag being present on stalking crime report	0%
Age	Predictor	Victim Age at Time of Incident Upper Committed Date	.9%
Gender	Predictor	Recorded gender of victim on Athena	.5%
Ethnicity	Predictor	Self-defined ethnicity of victim recorded on Athena	16.5%
SuspectSubCHITotal	Possible Outcome?	Total cumulative CCHI score that suspect caused during 365-day follow-up in Kent	0%
SuspectSubCount	Possible Outcome?	Total cumulative crime count that suspect committed during 365-day follow-up in Kent	0%
SuspectHistoricCHITotal	Predictor	Suspect Historic CCHI Total of offending in Kent	0%
SuspectHistoricCount	Predictor	Suspect Historic crime count of offending in Kent	0%
SuspectPriorStalking	Predictor	Suspect shown as a suspect on a previously reported stalking in Kent	0%

SusFollow-upStalk		Suspect recorded as a suspect on a stalking report recorded during 365-day follow-up	0%
Suspect same as victim?		N/A	0%
Risk Grading	Predictor	Risk grading assigned by the initial attending officer at the stalking case	21.1%
Historic CHI Total	Predictor	Victim Historic CCHI Total of victimisations in Kent	0%
Historic Count Total	Predictor	Victim Historic crime count of victimisations in Kent	0%
Victim Prior Stalking	Predictor	Victim shown as a victim on a previously reported stalking in Kent	0%
Victim Follow-up Stalking		Victim recorded as a victim on a stalking report recorded during 365-day follow-up period	0%

## Appendix 5 – Random Forest Training Data Model

Random forest model based on TRAINING data

	Actual high	Actual low	Totals
Forecast high	88 true positive	215 false positive	303
Forecast low	60 false negative	1070 true negative	1130
Totals	148	1285	1433
Percent	59% (sensitivity)	83% (specificity)	
Total per cent of forecasts that are accurate in each box within 365 days			80%
Of those forecast to be high risk, per cent who were not (Cautious error)			70%
Of those forecast to be low risk, percent which actually were high risk (Dangerous error)			5%
Cautious error/Dangerous error ratio			1:3.5