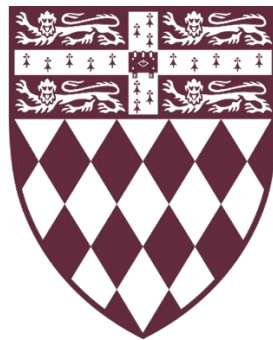Candidate No. Pol-2336

Daniel Melvin Thompson

Fitzwilliam College

'Moral Choice Architecture: a Randomised Controlled Trial of "Nudge" Interventions to Influence Stalking Re-offending'

Supervisor: Professor Barak Ariel

17,440 words

Submitted in part fulfilment of the requirements of the Master of Studies degree in Applied Criminology and Police Management

December 2024

Table of Contents

Lists of Figures and Tables

List of Figures:

List of Tables:

This page has been removed

Glossary of Terms

Term:	Definition:
Basic Command Unit (BCU)	A method of dividing London into local policing units with a standalone command structure. Basic Command Units can cover between 2 and 4 London boroughs.
Behavioural change programme	A blended series of interventions delivered by one or more agencies to discourage repeat offending
Breaching a Stalking Protection Order	Stalking offence whereby a court has imposed an Interim Stalking Protection Order or Full Stalking Protection Order, and the stalker has breached any condition.
Cambridge Crime Harm Index (CCHI)	System of scoring the harm caused by an offender based on the number of sentencing days assigned for a first offence with no aggravating features
CONNECT	From February 2024 onwards, the MPS used this system to record crimes, crime-related incidents, and other operational matters.
Crime Recording Information System (CRIS)	The system used by the MPS to record crimes and crime-related incidents until February 2024
Crime report flagging	A system of applying codes to crime reports for certain features of an offence to allow analysis of multiple reports
General Practitioner (GP)	Community-based doctors' surgeries operate and provide medical services in the United Kingdom.
His Majesty's Inspectorate of Constabulary and Fire and Rescue Services	Inspecting body for police services in England and Wales
Impact of Event Scale (IES)	PTSD scoring system based on a self-survey questionnaire, focusing on the impact of an event on an individual
London Metropolitan Police District	The Metropolitan Police Service polices the area of London, which comprises 32 boroughs but excludes the City of London.
Moral persuasion	Description of interventions with offenders designed to make a moral case to change behaviour
MPS	Metropolitan Police Service

National Stalking Consortium	A body of stalking charities, academics, and other experts advises on stalking. It is one of the bodies in the United Kingdom that raised a 'Super Complaint' for investigation.
Nudge text	A Short Messaging Service (SMS) text is sent to an individual to influence behaviour and choices.
Protection from Harassment Act (1997)	Statute which defines stalking offences in England and Wales
Re-offending prevalence	The percentage of offenders in a sample who commit a further offence
s2a Course of Conduct Stalking	The lower tier of stalking offence in the Protection from Harassment Act (1997) does not require proof of victim impact.
s4a Stalking - Fear of Violence	Stalking offence where the words or actions of a stalker are such that a reasonable person would fear that violence would be used against the victim
s4a Stalking - Serious Alarm or Distress	Stalking offence where a significant impact on the victim's wellbeing is in evidence
Stalker Typologies	A set of categories for stalking behaviours set out by Pathe and Mullen (1997)
Test message	A generic SMS text message was sent to eligible participants to test the accuracy of mobile phone numbers.
Wirefast	The system used by the Metropolitan Police Service to distribute SMS texts to a large audience.

Abstract

'Moral Choice Architecture: a randomised controlled trial of "nudge" interventions to influence

stalking re-offending'

Researchers:

Principal Researcher – Daniel Thompson, Metropolitan Police Service

1st Co-Principal Investigator – Professor Barak Ariel, Institute of Criminology, University of

Cambridge

2nd Co-Principal Investigator – Professor Yuval Feldman, Bar Ilan University

Trial Design:

Conducted as a parallel randomised controlled trial with a single treatment and control group.

Method:

Participants

This randomised controlled trial was conducted between 15th April and 15th August 2024 within the

Metropolitan Police District of London, United Kingdom. Participants were drawn from individuals

listed as suspects of stalking on crime reports between 31st March 2022 and 31st March 2024. No

exclusion criteria were applied for the sex or ethnicity of suspects. Exclusion criteria were used for age

to remove suspects who were younger than 18 at the time an allegation of stalking was made against

them. Allegations of domestic abuse and non-domestic abuse were included.

Further exclusion criteria were applied to the participant cohort to ensure only suspects with valid UK

mobile phone numbers listed were included. Further exclusion criteria were used to remove suspects

who were deceased, for whom prosecution was deemed not in the public interest, whose cases were

considered crime-related incidents, whose cases were transferred out of the Metropolitan Police, those which were statutory body investigations, those suspected entered into reports in error, and where allegations were proven to be false.

In addition, suspects who had committed more than one stalking offence within the inclusion period were removed to ensure all suspects analysed had been accused of a single offence.

Unit of Analysis

The study's unit of analysis is stalking suspects against whom an allegation was made between March 31st 2022 and March 31st 2024.

Intervention

Participants were randomly assigned to receive or not receive a single nudge text Short Messaging Service (SMS) text intervention to their mobile phone number listed on crime reports. The intervention was delivered on April 15th 2024, commencing a four-month post-intervention period from April 15th to August 15th 2024.

The principal researcher designed the intervention with advice from the first and second co-principal investigators to blend a deterrence and moral persuasive message. The former element was based on existing literature on the therapeutic treatment of stalking suspects and deterrence theory, highlighting the risks of imprisonment for repeat offending. The latter comprised a summary of the harms to stalking victims and associated third parties observed in the existing literature.

Objective

Specific Objective

To compare the effectiveness of a nudge text intervention in reducing the prevalence of re-offending and crime harm as defined by the Cambridge Crime Harm Index. To do so, compare treatment and control overall, male vs. female suspects, suspects with male vs. female victims, suspects by three age

categories, domestic abuse vs. non-domestic abuse cases and male vs. female suspects across two categories of time lapse between their qualifying stalking offence and date of intervention delivery.

Hypothesis

The combination of the threat of sanction (deterrence) and a realisation of the harm caused by offending behaviour would prompt suspects to re-offend at a lower rate than those who did not receive treatment.

Outcome

A manual search of Metropolitan Police Service crime indices was conducted by the primary researcher after the conclusion of the post-intervention period. A record was compiled of how many suspects committed one or multiple stalking offences in a binary format (0/1) to create a total of re-offending suspects in each cohort and measure the prevalence of re-offending suspects as a percentage of the cohort. In addition, the number of stalking offences committed in treatment and control were recorded against three stalking offences listed in the Protection from Harassment Act (1997) and a single stalking offence from the Stalking Protection Act (2019), Breach of a Stalking Protection Order. Corresponding Cambridge Crime Harm Index (CCHI) scores were recorded for treatment and control to facilitate comparison of crime harm between both groups.

Randomisation

Following the application of exclusion criteria, the first co-principal investigator at the Institute of Criminology, University of Cambridge, randomised the sample using the Microsoft Excel randomisation function. As reported in the ' Findings ' chapter, a balance table was completed to assess the satisfactory comparability of demographic and crime-specific factors between the treatment and control groups.

<div align="center">**Results:**</div>

Numbers Randomised and Analysed

The original sample comprised n = 4,443 suspects. The application of exclusion criteria reduced the sample to n = 1,673 (excluded n = 2,770). Randomisation resulted in the treatment comprising n = 846 and control n = 848.

Subsequent analysis identified n = 34 repeat suspects requiring exclusion. Randomising multiple counts of the same suspect offending removed n = 34 repeat suspects appearing multiple times across treatment and control, reducing the treatment group to n = 817 and the control group to n = 822.

To allow the 'Treatment as Delivered' analysis, suspects who did not receive the nudge text intervention were excluded (n = 293), resulting in a follow-up analysis of n = 524 in the treatment group. The control group of n = 822 did not vary in follow-up analysis.

Trial Status

The trial concluded at the end of the post-intervention period on the 15th of August, 2024.

Outcomes

Overall, no statistically significant crime prevalence or harm reductions were observed for any primary or secondary outcomes, although two came close to statistical significance.

Primary Research Outcome: ***Does the delivery of a nudge text intervention that combines moral persuasion with deterrence cause a reduction in stalker re-offending and Cambridge Crime Harm Index score compared to no intervention?***

Treatment did not achieve an overall reduction in crime prevalence. The treatment group (n = 524, M = .04, SD = .20) re-offended at a rate of 4.2% (n = 22) compared with control (n = 822, M = .04, SD = .18) who re-offending at a rate of 3.5% (n = 29), a backfire effect of 0.7% (p = .5378).

For crime, harm reduction treatment resulted in crime harm reduction, not at the level of statistical significance. The treatment group (n = 524, M = 2.56, SD = 21.16) returned a score of 1,341 compared with the control (n = 822, M = 3.09, SD = 25.52), where a score of 2,542 was observed (p = .6778) or a reduction of 47.25% (n = 1,201).

Secondary Research Outcome 1: *Does the delivery of a nudge text intervention that combines moral persuasion with deterrence cause a reduction in stalker re-offending and Cambridge Crime Harm Index score compared to no intervention, comparing male and female suspects?*

Treatment did not achieve reductions in crime harm prevalence for male suspects but did for female suspects, not at the level of statistical significance. In treatment, male suspects (n = 407, M = .04, SD = .21) re-offended at a rate of 4.4% (n = 18) compared with control (n = 627, M = .03, SD = .17) who re-offended at a rate of 2.9% (n = 18), a backfire effect of 1.5% (p = .2035). Female suspects in treatment (n = 83, M = .04, SD = .19) re-offended at a rate of 3.6% (n = 3) compared with control (n = 132, M = .07, SD = .25) re-offending at a rate of 6.8% (n = 9), a reduction effect of 3.2% (p = .2895).

Treatment achieved crime harm reductions for male and female suspects. For male suspects in treatment (n = 407, M = 3.21, SD = 23.96), a CCHI score of 1,306 was recorded compared with control (n = 627, M = 2.47, SD = 23.01) at 1,548 – reduction of 15.63% (n = 242) (p = .6223). For female suspects in treatment (n = 83, M = .36, SD = 1.88), a CCHI score of 30 was observed compared with control (n = 132, M = 6.82, SD = 38.45), where a score of 900 was observed, a reduction in harm of 96.67% (n = 900) (p = .0561).

Secondary Research Outcome 2: *Does the delivery of a nudge text intervention that combines moral persuasion with deterrence cause a reduction in stalker re-offending and Cambridge Crime Harm Index score compared to no intervention, comparing suspects with female victims to suspects with male victims?*

Treatment did not achieve crime prevalence reductions for suspects with female victims but did for those with male victims. In treatment suspects with female victims (n = 427, M = 0.05, SD = 0.21) re-offended at 4.7% (n = 20) compared with control (n = 690, M = .03, SD = .17) at 2.9% (n = 20) – backfire effect of 1.8% (p = .1394). Suspects with male victims in treatment (n = 87, M = .02, SD = .15) re-offended at 2.3% (n = 2) compared with control (n = 116, M = .07, SD = .25) at 6.9% (n = 8) – crime reduction effect of 4.6% (p = .1098).

Crime harm reductions were observed in both categories following treatment delivery. In treatment, suspects with female victims (n = 427, M = 3.09, SD = 23.40) returned a CCHI score of 1,321 compared with control (n = 690, M = 2.37, SD = 22.17) with a score of 1,637, a reduction of 19.30% (n = 316) (p = .6097). Suspects with male victims in treatment (n = 87, M = .23, SD = 1.51) returned a CCHI score of 20 compared with control (n = 116, M = 7.67, SD = 40.96) at 890, a reduction of 97.75% (n = 870) (p = .0528).

Secondary Research Outcome 3: *Does the delivery of a nudge text intervention that combines moral persuasion with deterrence cause a reduction in stalker re-offending and Cambridge Crime Harm Index score compared to no intervention, comparing suspects' age category?*

Treatment resulted in crime prevalence reductions for suspects aged 18-29 and 40 years or older but prompted a moderate backfire effect for those aged 30-39. In treatment suspects aged 18-29 (n = 135, M = .04, SD = .19) re-offended at a rate of 3.7% (n = 9) compared with control (n = 232, M = .04, SD = .19) re-offending at a rate of 3.9% (n = 9) (p = .9323). In treatment suspects aged 30-39 (n = 165, M = .05, SD = .23) re-offended at a rate of 5.5% (n = 9) compared with control (n = 318, M = .03, SD = .17) re-offending at a rate of 3.1% (n = 10) (p = .2550). Suspects aged 40 or older in treatment (n = 224, M = .04, SD = .19) re-offended at a rate of 3.6% (n = 8) compared with control (n = 272, M = .04, SD = .19) re-offending at a rate of 3.7% (n = 10) (p = .9504).

Crime harm reductions were observed for all age categories, in treatment suspected aged 18-29 years (n = 135, M = 2.71, SD = 22.83) returned a CCHI score of 366 compared with control (n = 232, M = 4.50,

SD = 29.97) with a score of 1,043 – a reduction of 64.91% (*n* = 677) (*p* = .5215). Suspects aged 30-39 in treatment (*n* = 165, *M* = 3.99, *SD* = 24.69) returned a score of 658 compared with control (*n* = 318, *M* = 3.37, *SD* = 28.28) with a score of 1,073 – a reduction of 38.68% (*n* = 415) (*p* = .8056). Suspects aged 40 years or older in treatment (*n* = 224, *M* = 1.42, *SD* = 16.90) returned a score of 317 compared with control (*n* = 272, *M* = 1.57, *SD* = 16.35) at 426 – a reduction of 25.59% (*n* = 109) (*p* = .9200).

Secondary Research Outcome 4: *Does the delivery of a nudge text intervention that combines moral persuasion with deterrence cause a reduction in stalker re-offending and Cambridge Crime Harm Index score compared to no intervention, comparing whether an offence was classified as domestic abuse or non-domestic abuse?*

No crime prevalence reduction benefit was observed for cases of domestic abuse but was for non-domestic abuse as a result of treatment. Domestic abuse cases in treatment (*n* = 410, *M* = .05, *SD* = .22) re-offended at a rate of 4.9% (*n* = 20) whilst control (*n* = 633, *M* = .04, *SD* = .19) re-offended at 3.8% (*n* = 24) (*p* = .4064). Non-domestic abuse cases in treatment (*n* = 114, *M* = .02, *SD* = .13) re-offended at 1.8% (*n* = 2) compared with control (*n* = 189, *M* = .03, *SD* = .16) at 2.6% (*n* = 5) (*p* = .6005).

Crime harm reductions were observed in both categories. For domestic abuse cases in treatment (*n* = 410, *M* = 3.04, *SD* = 23.54), a CCHI score of 1,247 is observed compared with control (*n* = 633, *M* = 3.16, *SD* = 25.35) at 2,003 – reduction of 37.74% (*n* = 756) (*p* = .9374). For non-domestic abuse cases in treatment (*n* = 114, *M* = .82, *SD* = 7.91), a score of 94 is observed compared with control (*n* = 189, *M* = 2.85, *SD* = 26.13) at 539 – a reduction of 82.56% (*n* = 445) (*p* = .3211).

Secondary Research Outcome 5: *Does the delivery of a nudge text intervention that combines moral persuasion with deterrence cause a reduction in stalker re-offending and Cambridge Crime Harm Index score compared to no intervention, comparing the time lapse between the commission of a first stalking offence and the delivery of the nudge text intervention across male and female suspect sex?*

Crime prevalence reductions were observed for female suspects irrespective of the age of their alleged offence at the time of the intervention. In contrast, for male suspects, reductions were seen at <180 days and a moderate backfire effect at 180 days or more. Male suspects whose alleged offence was 179 days or less old at the time of intervention in treatment (n = 32, M = .03, SD = .18) re-offended at a rate of 3.1% (n = 1) compared with control (n = 26, M = .08, SD = .27) re-offending at a rate of 7.7% (n = 2) (p = .4634). Female suspects in the same time category in treatment (n = 11, M = .00, SD = .00) re-offended at a rate of 0.0% (n = 0) compared with control (n = 22, M = .14, SD = .35) at 13.6% (n = 3) (p = .1055). Male suspects whose alleged offence was 180 days or older at the time of intervention in treatment (n = 329, M = .04, SD = .20) re-offended at a rate of 4.3% (n = 14) compared with control (n = 541, M = .03, SD = .16) at 2.6% (n = 14) (p = .2026). Female suspects in the same time category in treatment (n = 70, M = .04, SD = .20) re-offended at a rate of 2.9% (n = 2) compared with control (n = 106, M = .06, SD = .23) at 5.7% (n = 6) (p = .6795).

Crime harm reductions were observed in all but one category. For male suspects whose alleged offence was 179 days or less old at the time of intervention in treatment (n = 32, M = 0.31, SD = 1.77), a CCHI score of 10 was observed compared with control (n = 26, M = 10.27, SD = 49.39) at 267 – reduction of 96.25% (n = 257) (p = .3137). For female suspects in the same category (n = 11, M = 0.00, SD = 0.00), a score of 0 is observed compared with control (n = 22, M = 5.18, SD = 20.05) at 114 – reduction of 100% (n = 114) (p = .2009). For male suspects whose offence was 180 days or older in treatment (n = 329, M = 3.11, SD = 22.78), a score of 1,024 is observed compared with control (n = 541, M = 1.88, SD = 19.54) at 1,019 – an increase of 0.49% (n = 5) (p = .4163). For female suspects in the same category in treatment (n = 70, M = .43, SD = 2.04), we observe a CCHI score of 30 compared with control (n = 106, M = 7.42, SD = 41.97) at 786 – reduction of 96.18% (n = 756) (p = .0897).

Research and Policy Implications:

A series of recommendations are made for further research. Firstly, the sample size for the study may have been too small to observe statistically-significant results. Recommendations are made for a multi-force or national study. Further research should also identify which stalking typology a suspect fits with and consider a block-randomised approach to understand the most effective interventions depending on the stalking typology. Further research which separates deterrence from moral persuasive messaging is recommended to understand better which element had the most significant impact on reducing crime harm and re-offending rates – especially given the stark differences between male and female stalkers. Further research should also include a follow-up period of two years, given that approximately half of stalkers who re-offend will do so beyond the four-month post-intervention window used in this study. Recommendations are made for further research, employing personalised text messages to measure this impact on crime harm and prevalence reductions.

Several policy recommendations are made. Policymakers should only consider the use of nudge text interventions within 6 months after a first offence is alleged to have occurred, and in doing so, need not distinguish between male and female stalkers. Those seeking to deliver the intervention beyond 6 months should only provide the intervention to female stalkers or risk increasing stalking prevalence and harm. Stalking prevention strategies aimed at public education should utilise the workplace as a forum to reach the target audience, as stalking offending correlates closely with the working age of UK adults. The low rate of re-offending and the small number of suspects committing multiple stalking offences should alert policymakers to the need for effective policies that are robust and adaptable enough to accommodate an ever-changing population of stalking suspects and victims. For that small cohort of high-frequency stalkers, the numbers are sufficiently small to warrant more intensive offender management and victim safeguarding strategies.

Introduction

In the same years this study was designed and conducted (2022-4) the National Stalking Consortium launched a super-complaint about the police response to stalking. They called on policing across England and Wales to improve, amongst other things, how it manages stalking perpetrators (GOV.UK, 2022). The ability of policing and the criminal justice system to respond effectively to stalkers has become critical to maintaining the public's confidence to preserve their safety.

At its core, stalking perpetration is behavioural, and as such it is potentially amendable through effective warnings and messages. Short Messaging Service (SMS), commonly called "texts," to influence behaviour is an emerging area of police practice. Such interventions by law enforcement form part of offering suspects 'choice architecture', highlighting the negative legal consequences of offending with the intention that some offenders will make better-informed choices. Doing so is described in this study and elsewhere as "nudging" an offender away from noncompliance and toward compliant behaviours. This study further develops a concept of "moral choice architecture" within this line of inquiry: the theory behind this approach is that, in addition to deploying deterrence theory to highlight the legal implications of offending behaviour, the perpetrator is also presented with a moral argument not to offend.

This study was conducted as a randomised controlled trial by the University of Cambridge and Metropolitan Police Service in 2024 and sought to answer the following primary question: does the delivery of a nudge text intervention that combines moral persuasion with deterrence cause a reduction in stalker re-offending and Cambridge Crime Harm Index score compared to no intervention? The study then goes to analyse the results in multiple pre-determined subgroups, based on a matrix of offenders' and victims' genders, age brackets, domestic versus non-domestic abuse,

and time elapsed between the commission of a first stalking offence and the delivery of the nudge text.

This study will begin by setting out the theoretical context in which it is situated, explored in the 'Literature Review' chapter. The methodology employed to design the randomised controlled trial, construct and deliver the nudge text intervention, and retrieve follow-up results after a four-month post-intervention period will be discussed in the 'Methodology' chapter. The same chapter will explore the external validity of the trial. The findings chapter will set out findings to primary and secondary research questions. The discussion chapter will then consider the theoretical findings of the trial, limitations in study design, the findings' external validity, and the policy implications of trial results.

Overall, the study has been able to demonstrate a non-statistically significant crime prevalence and harm reduction effect in cases not older than 6 months from reporting, and – for female stalkers – beyond that time limitation. It has shown a backfire effect for male stalkers beyond 6 months. The absence of a *p* value <0.5 must ultimately alert us to the need for further research to confirm what this study can only suggest.

This study presents an argument that delivering a nudge text intervention to stalking suspects that blends deterrence and moral persuasion messaging may deliver crime prevalence and harm reduction benefits if delivered under certain conditions. It suggests, caveated with the need for further research, that such an intervention may have a greater effect on female than male stalkers. The concept that police may require a different approach for stalkers depending on sex is a complex one. In the context of evidence-based policing, this study only provides an initial insight. However, the gender differences in stalking intervention offer significant scope to develop this study's findings further.

Chapter I

Literature Review

<u>Overview</u>

This chapter begins by considering why stalking perpetration is an essential yet under-investigated area of research. It considers the risk of repeat offending and the need for more proactive police interventions. It will set out the relevance of deterrence theory and how that theory can inform the construction of a nudge text intervention to influence the likelihood of effective disruption of offending. Specific focus will be given to nudge interventions, reviewing existing and relevant learning, the role of personalisation of messages, and the relevance of a 'social norms' approach (Chivers & Barnes, 2018, p.9). Moral persuasion will be considered as well (Ariel, 2012). Finally, a review is provided of existing literature on the impact of stalking on victims, which informs the construction of this trial's nudge text as described in the subsequent chapter.

<u>The Risks of Stalking</u>

A review of the literature identifies several risks posed by stalking perpetrators, demonstrating some of the predictable features of stalking. First, Churcher and Nesca's (2013) meta-analysis of 27 studies of stalking risk identified, from a sample of 5,114 stalking perpetrators, that 35% (n = 1,059) of stalkers were violent and 29% (n = 604) of stalking victims were injured. The study identified that prior intimate relationship (i.e. domestic abuse) was a particularly strong predictor of violence.

Furthermore, a moderate effect size was identified for verbal threats as predictors of future physical violence (Chrucher & Nesca, 2013). This correlation between verbal threats and physical violence has been observed in other studies. A qualitative study of verbal threats as a predictor of violence asked questions of 187 female intimate partner stalking victims to identify common trends (Brewster, 2000).

Questions were asked about violent history, drug and alcohol misuse, frequency of phone calls, sending of letters, physical following by the stalker, and the victim's age and education. Three dependent variables were then analysed for each survey response: whether violence occurred, the number of violent incidents, and whether physical injury occurred. Using a linear and logistic regression model, the study identified a moderate statistically significant correlation between verbal threats and subsequent violence. The study further found that alcohol and drug abuse predicted physical injury during stalking (*Ibid*). Arguably, this study's findings should be approached cautiously – the case sample was 187 victims, making it difficult to detect an effect size for the dependent variables Brewster cites. Brewster's (2000) study deals with demographic factors, an analysis lacking in other papers; curiously, the analysis identified that ethnicity, age and education had no statistically significant effect on escalation to violence, whilst alcohol and drug misuse, threats of violence and marital status did (*Ibid*).

Developing the concept of stalking risk further, cases which lead to homicide may have distinguishing features, at least in the case of ex-intimate partners. A review of 374 homicides identified common stages perpetrators progress through - from pre-relationship, early relationship, relationship, triggers, escalation, changes in thinking/decision, planning, and homicide (Monckton-Smith, 2020). Critically, the study shows how a perpetrator can escalate rapidly and bypass stages to arrive at severe violence and homicide more quickly – for example, experiencing an acute reaction to the 'trigger' stage, bypassing any planning and carrying out the homicidal act.

In summary, it is apparent that stalking risk between dyads evolves dynamically, sometimes following a predictable pattern and sometimes escalating unpredictably, even to homicide. The level of that risk is informed by a complex interplay of factors, from the presence of threats and substance misuse to a suspect's antecedents and violence history. A victim's or suspect's demography has no statistically significant impact on the likelihood of offending or escalation to physical violence. This is relevant, as the present study would obtain no benefit from block-randomising into demographic characteristics,

and the crime systems from which data was drawn (set out in the methods chapter) do not allow isolation of cases to the above predictors of re-offending and escalation.

The Harm Caused by Stalking

Research consistently shows the damage caused by stalking behaviours. First, 53% terminated employment; 39% moved residence; 83% reported increased anxiety levels; 55% experienced intrusive recollections, nightmares, appetite disturbance and depressed mood; 24% reported suicidal ideation; and 37% met the diagnostic criteria for post-traumatic stress disorder (Pathe & Mullen, 1997). Storey et al. (2023) have also found that 59% of respondents (n = 119) "reported a clinically significant level of psycho-medical symptoms" (*Ibid*, p.796-797). Where studies talk about the prevalence of Post-Traumatic Stress Disorder (PTSD) symptoms in stalking victims, they often refer to the 'Impact of Event Scale' (IES), which uses a self-survey questionnaire with an overall score and sub-scores for 'intrusion', 'avoidance' and 'hyper-arousal'. Of a maximum score of 88, a total score of >33 indicates the presence of PTSD. Storey, Pina and Williams (2023) found a mean score of 39.7 (SD = 17) across their 119 respondents, with a score of 18.0 (SD = 7.9) for intrusion and 18.2 (SD = 8.6) for avoidance (*Ibid*). Looking specifically at intimate-partner stalking, Kamphuis et al. found the mean average for IES was 40.50 (SD = 16.24), with sub-scores for intrusion at 18.28 (SD = 7.56) and avoidance at 18.53 (SD = 8.48) (Kamphuis et al., 2003). The researchers attributed this to post-intimate stalking victims suffering, "prolonged exposure to a wide range of violent and non-violent stalking behaviours" (*Ibid*, p.154).

Of critical importance to the moral case for stalking cessation is the impact on children. Elklit et al. used a qualitative 'maternal diagnostic interview' for children in the 0-6 age range and age-appropriate questionnaires for children aged 7-19 (Elklit et al., 2019). Their findings are illustrative: 22% of 0-6-year-olds met PTSD criteria, increasing to 85% for 7-11-year-olds and 58% for 12-19-year-olds. Together, these various studies present a moral case upon which behavioural change for stalkers

could be formulated, highlighting the social consequences of stalking and the ethical imperative for change.

Typologies of Stalkers

A critical study in this space is Pathe and Mullen's (2009) work. Stalkers are described as "a lonely, inadequate, disturbed…group of human beings" with "modifiable characterological and social skills deficits" (p.251). The typology, which continues to inform practice to date, is a helpful disaggregation of stalkers into categories. As in this cohort, we expect to have in the present study's cohort 'predatory stalkers', intent on committing violent or sexual assaults; 'rejected stalkers' pursuing a lost relationship or seeking to inflict punishment for its end; 'resentful stalkers' seeking restitution for a perceived wrong; 'intimacy seekers' intent on forming a close, emotional relationship; and 'incompetent suitors' seeking short-term romantic encounters but unaware of social conventions or unable to process a victim's refusal. This tapestry of motivations, vulnerabilities and manifestations of stalking should alert any study to the diversity of this cohort: there is no 'typical' stalker.

Pathe and Mullen (2009) have also identified that treatments which focus on exposing stalkers to the harm they cause to victims are ineffective. The criticism lies in the fact that stalking treatment has borrowed from sex offender treatment in seeking to encourage victim empathy. However, while most stalkers are aware of "the costs and self-defeating nature of their behaviours", many treatments can only reinforce the negative consequences for them (*Ibid*, p.257). Still, many stalkers maintain a "capacity to deny, minimise and rationalise" their behaviours (Ibid, p.254). It could be that some stalkers do not believe they are stalking at all, so a treatment that focuses on harm to victims alone is unlikely to exhibit efficacy.

More recently, Wheatley, Henley and Farnham (2023) used a qualitative, interview-based methodology to understand stalkers' experiences and the relevance of deterrence theory. They found

that most offenders felt they possessed little knowledge of the legal consequences of stalking – imprisonment, providing a sobering moment (Wheatley et al., 2023). This is important, particularly when we think about the content of an effective message: stalkers ought to be notified not only that their behaviour is devastating but also criminal.

Preventing Stalking with Deterrence?

Mustaine and Tewksbury's (1999) utilised a routine activity theory approach to explicate female stalking victimisation. Stalkers are motivated by an intense, internalised fixation; their targets may be vulnerable by their routine, which stalkers may target, and police involvement as capable guardians may impact offending rates. The study found that demography had no statistically significant impact on predicting victimisation. Conversely, employment status, residential location, drink and drug misuse and whether or not victims employed self-protection measures were found to be predictive of stalking victimisation. Importantly, the lifestyle of victims can predict where stalking is more likely to take place – which then offers opportunities for prevention.

One important theory that is linked to prevention is deterrence. Effective deterrence theory materialises when the certainty, celerity, and severity are sufficiently elevated, all contributing to effective threats. When crime is a deliberate, calculated behaviour, law enforcement can capitalise on rational thinking to lead to prevention through warnings – i.e., threats (Ariel, 2012).

Within this line of inquiry, specific deterrence – the attempt to prevent an individual from recidivism through warnings – is particularly relevant (Braga et al., 2018). Here, direct communication with the offending population raises awareness of increased scrutiny of their offending, and a clear moral case for behavioural change reinforced with an offer of support may yield positive results. Law enforcement might create a sense of omnipresence by establishing a foothold in the offender's environment (physical or digital). Durlauf and Nagin (2011, p. 14) propose "a focus on a more effective

use of police to make the risks of crime clearer and the consequences of crime faster and more certain".

However, whether this approach is valid for preventing stalking is presently unclear. Where research exists on the police role in deterrence, it does not relate to stalking. Likewise, whilst research exists on the role of deterrence in stalking, it does not explicitly cover police efforts to deter, rather than punish, stalking perpetrators. This study, therefore, will seek to fill that gap by providing a randomised controlled trial of a law enforcement agency sending a deterrent text message to those accused of stalking.

Preventing Stalking through Social Norms

As the preceding section of this review showed, therapeutic interventions with stalkers focus on realising the harm they are causing. Whether supplementary approaches would more effectively reduce recidivism is presently unclear. One possible approach is the effect of 'moral persuasion' on prospective offenders. Ariel's (2012) study considered this approach, which compared deterrence with moral persuasion to influence tax evaders. Ariel highlights that taxpayers have a moral imperative based on social norms: many are "not assumed to search for ways to maximise their economic utility because they have moral and social obligations at stake, in their preference to comply" (*Ibid*, p.28). Paternoster & Simpson (1996) identified that people are "more likely to report an intention to commit the act [i.e. a crime] when it was common practice within the firm" (Ibid, p.568). Thus, group morality is worthy of consideration when attempting to deter prospective offenders. One interpretation of the moral persuasion is the underscoring of victims' suffering and that stalking is not an acceptable behaviour. As reviewed earlier, victims of stalking are indeed harmed, and underscoring this harm may focus the minds of offenders on impact.

Summary

The literature reviewed in this section has been selected to inform how an RCT should approach sending text interventions to stalking perpetrators. The research illustrates the complexity of stalking risks—escalating at an unpredictable rate although following a predictable pattern, informed by coefficients such as the presence of threats, substance misuse, suspect antecedents, and other factors. Considering deterrence theory, the absence of any study specifically on stalking is notable.

Though the literature suggests that moral persuasion—what this study conceives of as 'Moral Choice Architecture'—may have an effect, this review has uncovered that no impact evaluation looked at the use of moral choice architecture in the context of stalking. Pathe and Mullen (2009) posit the various typologies of stalkers and their motivations, highlighting further complexity. However, an intervention that is well-designed and considered at the pre-mortem stage will have to contend with this.

There are, however, consistent considerations for all stalkers when considering deterrence theory: the need to emphasise the certainty of apprehension whilst not disregarding severity and the role of electronic communication, such as SMS text, in extending the 'capable guardian' role of law enforcement. This review has uncovered no study specifically using nudge texts to deter stalkers. No study that looks specifically at blending moral persuasion with the threat of sanction has been identified. No study looked specifically at a nudge text intervention informing offenders of the harm they are causing to victims.

Overall, the literature reviewed in this chapter suggests that if the best practice identified is followed and 'treatment as delivered' (Cumberbatch & Barnes, 2018) is achieved, a nudge text intervention is a viable tactic to disrupt stalking perpetration. The approach taken by this study to achieve that end shall be explored in the next chapter.

Chapter II

Methods

Overview

This chapter will outline participant selection and exclusion criteria and the journey from the eligible population to the final list of participants for randomisation. The impact of exclusion criteria on the final sample selected for randomisation will be outlined. The study's setting, the London Metropolitan Police District, will be described as well as demographic factors unique to London which may influence external validity. Critical to this study is the design of the nudge text intervention and how this was achieved. The primary and secondary outcomes will be listed. Randomising the participant sample into control and experimental groups will be explained. Finally, the selected statistical methods used to compare those groups for the primary and secondary outcomes and the reasons for doing so will be set out.

Setting

This study was conducted in London, the capital city of England in the United Kingdom. London is situated in the south-east of England. Its centre comprises the City of London, excluded from this study as forming a separate police service area (City of London Police). It has a densely populated inner-city area. Its outer-city area, with a large residential population, comprises a mix of rural and urban settings. Policing services to the 32 boroughs of Greater London are provided by the Metropolitan Police Service across twelve Basic Command Units (BCU's), policing between two to four boroughs apiece. These BCU's provide frontline policing to communities. Greater London, the UK's largest metropolitan area, has a population of approximately nine million people (Statista, 2024).

External Validity

London as a large metropolitan area has distinct demographic characteristics which may influence external validity. In addition, whilst this study drew from all available stalkers, a temporal exclusion criteria was nevertheless applied which makes this study's cohort distinct – stalkers whose offence was reported between the 31st March 2022 and 31st March 2024. Such a sample may differ, for example, from stalkers reported between 2020-22, 2018-20 and so on. The time criteria was selected because it was felt inappropriate to contact a stalker whose case had not been live for more than two years. The question arises as to whether those suspects with allegations between those dates have distinct characteristics when comparing ethnicity, sex, age and whether or not a case was domestic abuse. As Table 1 demonstrates, however, stalking suspects in this trial showed little variance from those from 2013-23.

To assess the generalisability of findings set out in the following chapter, this study has compared the age distribution in the treatment and control groups reported in the preceding chapter with data showing the age distribution of London in 2022 and England and Wales from the 2021 census. As seen in Table 2 and Fig. 1, London has a younger population than England and Wales. The age category 20-24 years is comparable at 6.2% in London and 6% across England and Wales. For 25-29 years, however, a variance was observed at 8.4% for London and 6.5% for England and Wales. This variance holds up until 50-54 years and beyond when London's population aged 50 years and above becomes proportionally smaller than England and Wales. As seen in Table 3 and Fig. 2, the same age demographics seen in London apply in Manchester, England. Therefore, large urban centres such as London may attract a younger demographic, which must be considered by any policymaker seeking to transport these findings to a non-urban setting.

Furthermore, stalking suspects in the treatment and control groups disproportionately fall in the age categories 20-44, accounting for 71.48% and 76.29% of stalking suspects in the trial, respectively.

Table 1 - Stalking reports drawn from police systems between 31/12/13 and 31/12/23 and this trial's treatment and control groups

	Category:	Treatment	Control	Stalking Reports (31/12/13-31/12/23)	Variance with Treatment Group	Variance with Control Group
Suspect Ethnicity	**Afro-Caribbean**	10.0%	13.0%	13.86%	**3.9%**	**0.9%**
	Arabian/Egyptian	2.0%	3.0%	2.14%	**0.1%**	**-0.9%**
	Asian	9.0%	12.0%	10.58%	**1.6%**	**-1.4%**
	Dark European	4.0%	4.0%	4.37%	**0.4%**	**0.4%**
	Oriental	1.0%	1.0%	0.67%	**-0.3%**	**-0.3%**
	Unknown	1.0%	1.0%	1.88%	**0.9%**	**0.9%**
	White European	23.0%	18.0%	21.56%	**-1.4%**	**3.6%**
	Not Recorded	50.0%	48.0%	44.94%	**-5.1%**	**-3.1%**
Sex	**Male**	78.0%	77.0%	75.83%	**-2.2%**	**-1.2%**
	Female	16.0%	16.0%	12.86%	**-3.1%**	**-3.1%**
	Unknown	1.0%	1.0%	0.90%	**-0.1%**	**-0.1%**
	Not Recorded	5.0%	6.0%	10.40%	**5.4%**	**4.4%**
Domestic Abuse?	**Domestic Abuse**	77.7%	77.1%	70.5%	**-7.2%**	**-6.6%**
	Non-Domestic Abuse	22.3%	22.9%	29.5%	**7.2%**	**6.6%**
Suspect Age (Years)	**0-4 years**	0.00%	0.00%	6.7%	**6.7%**	**6.7%**
	5-9 years	0.00%	0.00%	0%	**0%**	**0%**
	10-14 years	0.00%	0.00%	0.3%	**0.3%**	**0.3%**
	15-19 years	2.08%	2.31%	3.1%	**1.1%**	**0.8%**
	20-24 years	11.38%	10.83%	9.8%	**-1.6%**	**-1.0%**
	25-29 years	14.69%	15.09%	14.9%	**0.2%**	**-0.2%**
	30-34 years	18.85%	20.32%	16.4%	**-2.5%**	**-3.9%**
	35-39 years	13.22%	18.25%	14.7%	**1.4%**	**-3.6%**
	40-44 years	13.34%	11.80%	11.0%	**-2.3%**	**-0.8%**
	45-49 years	8.20%	7.66%	8.0%	**-0.2%**	**0.4%**
	50-54 years	7.59%	6.20%	7.0%	**-0.6%**	**0.8%**
	55-59 years	6.00%	2.92%	3.9%	**-2.1%**	**1.0%**
	60-64 years	2.94%	2.92%	2.1%	**-0.8%**	**-0.8%**
	65-69 years	0.98%	0.97%	1.1%	**0.1%**	**0.2%**
	70-74 years	0.24%	0.36%	0.5%	**0.2%**	**0.1%**
	75-79 years	0.49%	0.24%	0.3%	**-0.2%**	**0.1%**
	80-84 years	0.00%	0.12%	0.1%	**0.1%**	**0.0%**
	85-90 years	0.00%	0.00%	0.0%	**0.0%**	**0.0%**

Table 2 - Distribution of age in London's 2020 Population (Statista, 2024), the 2021 England and Wales Census Data (Office for National Statistics, 2024) and the Treatment and Control groups in the trial

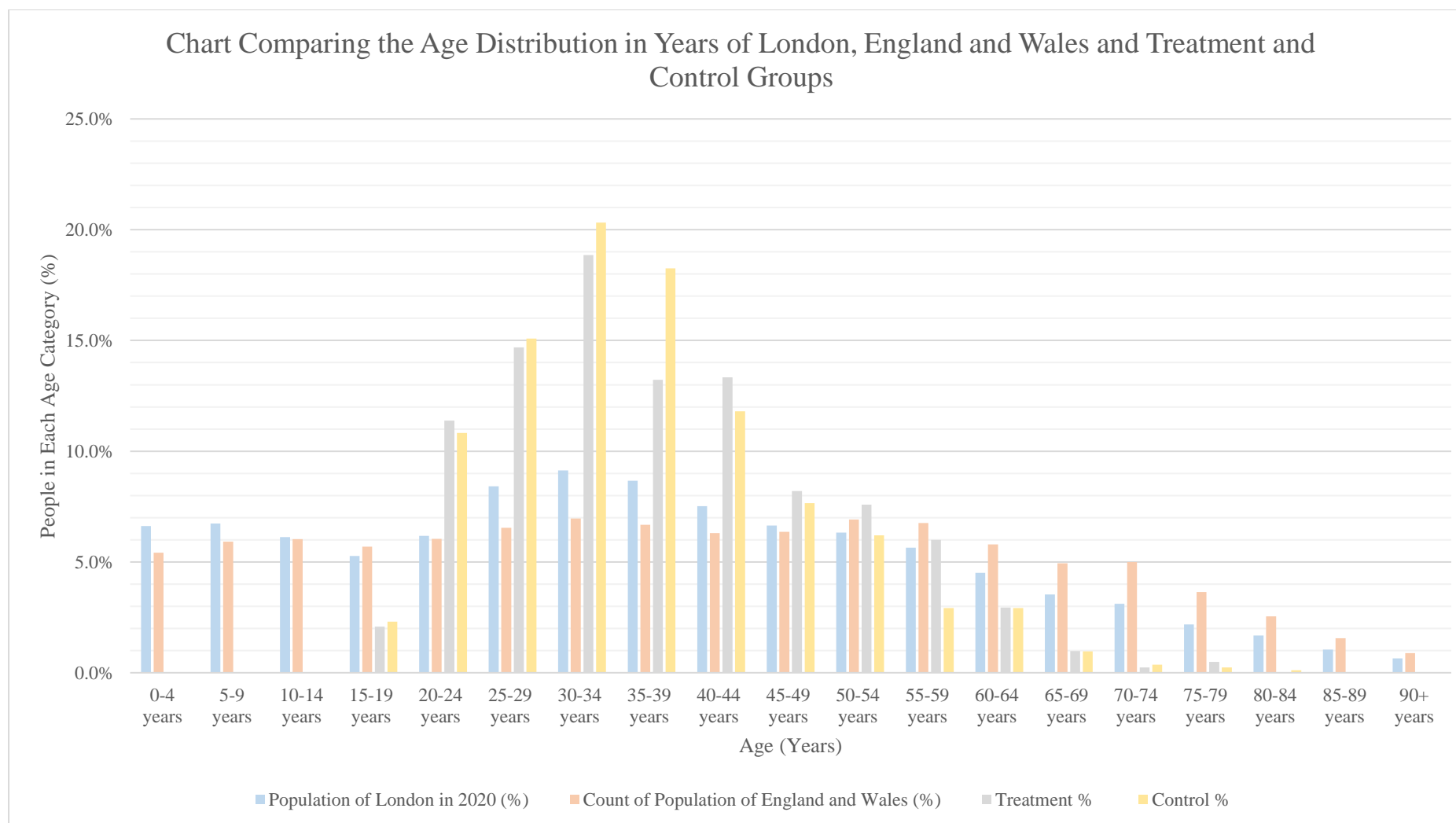|	Count of the Population of London in 2020 (thousands)	% of the Population of London in 2020	Count of Population of England and Wales (Census Data, 2021)	% of the population of England and Wales (Census Data, 2021)	Treatment Count of Participants	Treatment % of Participants	Control Count of Participants	Control % of Participants
0-4 years	595.8	6.6%	3232036	5.4%	0	0.0%	0	0.0%
5-9 years	606.33	6.7%	3524627	5.9%	0	0.0%	0	0.0%
10-14 years	550.75	6.1%	3596029	6.0%	0	0.0%	0	0.0%
15-19 years	474.46	5.3%	3394665	5.7%	17	2.08%	19	2.31%
20-24 years	556.59	6.2%	3602128	6.0%	93	11.38%	89	10.83%
25-29 years	757.85	8.4%	3901740	6.5%	120	14.69%	124	15.09%
30-34 years	822.08	9.1%	4148800	7.0%	154	18.85%	167	20.32%
35-39 years	779.93	8.7%	3981617	6.7%	108	13.22%	150	18.25%
40-44 years	677.46	7.5%	3755757	6.3%	109	13.34%	97	11.80%
45-49 years	598.54	6.6%	3788721	6.4%	67	8.20%	63	7.66%
50-54 years	569.94	6.3%	4123432	6.9%	62	7.59%	51	6.20%
55-59 years	508.72	5.7%	4029043	6.8%	49	6.00%	24	2.92%
60-64 years	405.58	4.5%	3455604	5.8%	24	2.94%	24	2.92%
65-69 years	318.14	3.5%	2945137	4.9%	8	0.98%	8	0.97%
70-74 years	280.43	3.1%	2977984	5.0%	2	0.24%	3	0.36%
75-79 years	196.42	2.2%	2170271	3.6%	4	0.49%	2	0.24%
80-84 years	150.98	1.7%	1515077	2.5%	0	0.00%	1	0.12%
85-89 years	93.8	1.0%	925346	1.6%	0	0.00%	0	0.0%
90+ years	58.67	0.7%	529532	0.9%	0	0.00%	0	0.0%
Total	9002.47		59597546					

Figure 1 - The distribution of age in London's 2020 Population (Statista, 2024), the 2021 England and Wales Census Data (Office for National Statistics, 2024) and the Treatment and Control groups in the trial

Table 3 – Comparing the age distribution of London in 2022 (Statista, 2024), England and Wales in 2021 (2024) and Manchester, England in 2021 (Manchester City Council, 2021)

	The population of London in 2020 (thousands)	%	Count of Population of England and Wales (Census Data, 2021)	%	Population of Manchester, England (Census Data, 2021)	%
0-4 years	595.8	6.6%	3232036	5.4%	34,378	6.2%
5-9 years	606.33	6.7%	3524627	5.9%	36,561	6.6%
10-14 years	550.75	6.1%	3596029	6.0%	36,344	6.6%
15-19 years	474.46	5.3%	3394665	5.7%	42,418	7.7%
20-24 years	556.59	6.2%	3602128	6.0%	61,887	11.2%
25-29 years	757.85	8.4%	3901740	6.5%	52,583	9.5%
30-34 years	822.08	9.1%	4148800	7.0%	47,999	8.7%
35-39 years	779.93	8.7%	3981617	6.7%	42,050	7.6%
40-44 years	677.46	7.5%	3755757	6.3%	35,778	6.5%
45-49 years	598.54	6.6%	3788721	6.4%	31,067	5.6%
50-54 years	569.94	6.3%	4123432	6.9%	30,370	5.5%
55-59 years	508.72	5.7%	4029043	6.8%	26,693	4.8%
60-64 years	405.58	4.5%	3455604	5.8%	21,642	3.9%
65-69 years	318.14	3.5%	2945137	4.9%	16,482	3.0%
70-74 years	280.43	3.1%	2977984	5.0%	13,817	2.5%
75-79 years	196.42	2.2%	2170271	3.6%	9,206	1.7%
80-84 years	150.98	1.7%	1515077	2.5%	6,689	1.2%
85-89 years	93.8	1.0%	925346	1.6%	5973	1.1%
90+ years	58.67	0.7%	529532	0.9%		
Total	**9002.47**		**59597546**		**551937**	

Comparing the Age Distribution in Years of London in 2020, Manchester England in 2021 and England and Wales in 2021
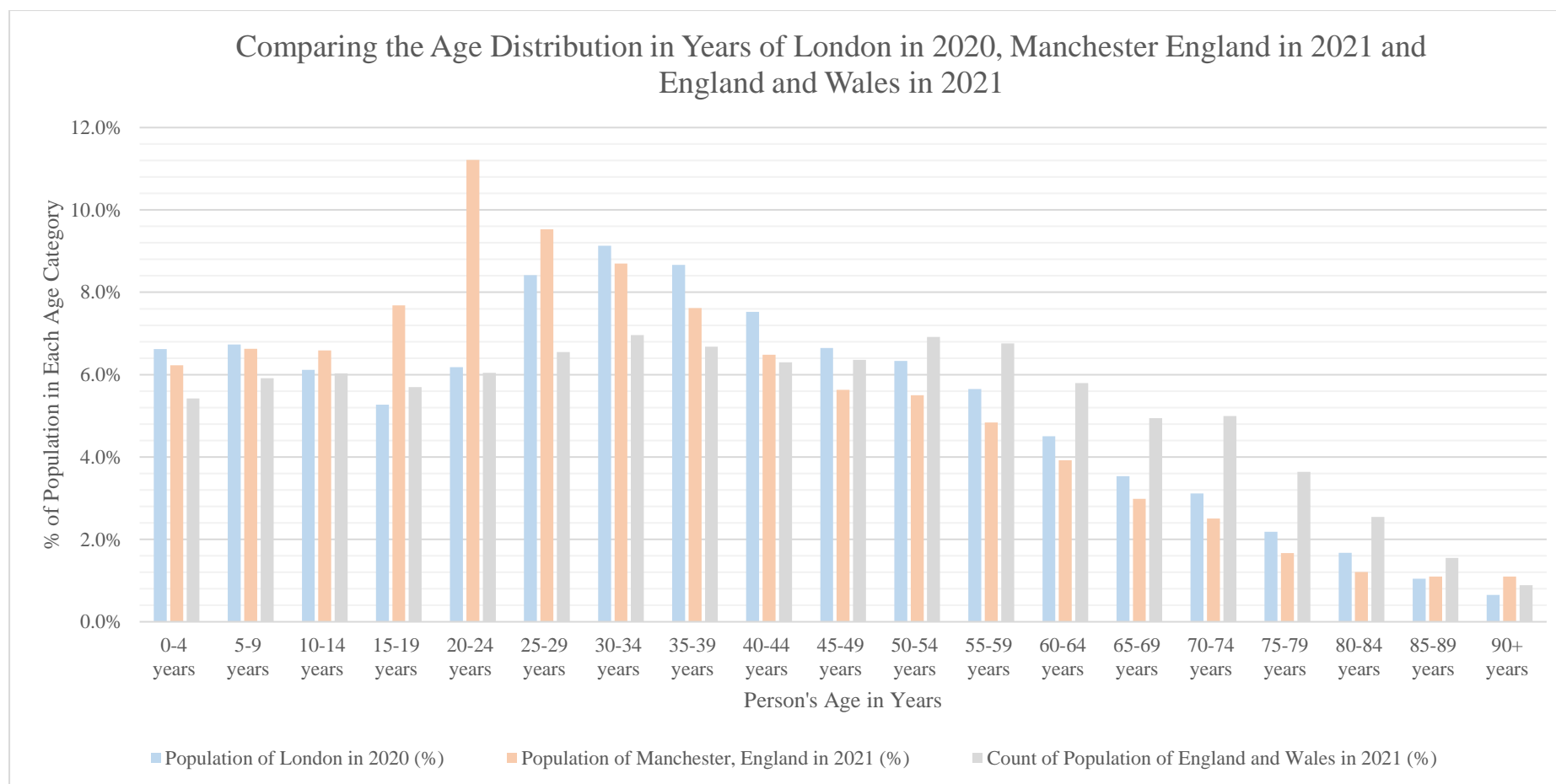
Figure 2 – Comparing the age distribution in years of London in 2021 (Statista, 2024), Manchester, England in the 2021 Census data (Manchester City Council, 2024) and England and Wales in the 2021 Census Data (Office for National Statistics, 2024)

2021 Census Data for England and Wales, which includes the London region as a category, has been obtained, and a comparison of London and England and Wales was conducted. The results are shown in Table 4. They demonstrate that across England and Wales, the proportion of the population self-identifying as 'White' or 'White British' accounted for 78.1% of the population, whereas in London, the proportion was 45.3% and 32.8% lower, respectively. Whilst the 2021 Census Data used different ethnicity categories commonly used in policing and used to report ethnicity in this trial, those other than 'White' and 'White British' can similarly be categorised under Black and Minority Ethnic (BAME). Therefore, England and Wales's BAME population is 21.9%, whereas in London, it is 54.7%. These unique demographic characteristics of London should be considered when determining the transportability of this trial's findings to other populations.

Table 4 – Showing the population's ethnicity according to 2021 Census Data in London and England and Wales (GOV.UK, 2024).

Ethnicity	England and Wales (Count)	England and Wales (%)	London (Count)	London (%)
Any Other Asian Background	972,783	0.8%	401,250	2.3%
Any Other Black Background	297,781	0.2%	145,915	0.8%
Any Other Ethnic Background	923,776	0.8%	416,980	2.4%
Any Other Mixed/Multiple Ethnic Background	467,116	0.4%	170,691	1.0%
Any Other White Background	3,667,993	3.1%	1,290,838	7.3%
Arab	331,856	0.3%	139,792	0.8%
Asian	5,515,455	4.6%	1,817,647	10.3%
Bangladeshi	644,900	0.5%	322,052	1.8%
Black	2,409,283	2.0%	1,188,369	6.8%
Black African	1,488,387	1.2%	697,054	4.0%
Black Caribbean	623,115	0.5%	345,400	2.0%
Chinese	445,646	0.4%	147,523	0.8%
Gypsy Or Irish Traveller	67,757	0.1%	7,029	0.0%
Indian	1,864,304	1.6%	656,269	3.7%
Mixed	1,717,977	1.4%	505,776	2.9%
Mixed White And Asian	488,228	0.4%	125,188	0.7%
Mixed White And Black African	249,593	0.2%	77,340	0.4%
Mixed White And Black Caribbean	513,040	0.4%	132,557	0.8%
Other	1,255,632	1.1%	556,772	3.2%
Pakistani	1,587,822	1.3%	290,553	1.7%
Roma	100,964	0.1%	37,690	0.2%
White	48,699,231	40.9%	4,731,172	26.9%
White British	44,355,044	37.2%	3,239,280	18.4%
White Irish	507,473	0.4%	156,335	0.9%

Census data from 2021 shows that sex distribution did not alter significantly in the London region when looking at England and Wales. This suggests that the findings in terms of the suspect's sex are transportable to regions outside of London.

Table 5 – 2021 Census data showing the sex breakdown of males and females in England and Wales compared with the London region (Office for National Statistics, 2024)

	England and Wales (2021 Census Data)	(%)	London Region (2021 Census Data)	(%)
Male	29,177,343	49.0%	4,531,622	51.5%
Female	30,420,202	51.0%	4,268,106	48.5%

<u>Trial Design</u>

The trial design employed was a parallel group randomised controlled trial, with two parallel groups – control and experimental allocated at a 1:1 ratio.

<u>Participants</u>

The unit of randomisation was individuals who had been named as suspects on a crime report recorded by the MPS between 31st March 2022 and 31st March 2024 as an offence under the following legislation:

i)   'Course of Conduct Stalking' contrary to s2a Protection from Harassment Act (1997);

ii)   'Stalking Involving Serious Alarm and Distress' contrary to s4a Protection from Harassment Act (1997);

iii)   'Stalking Involving Fear of Violence' contrary to s4a Protection from Harassment Act (1997); and,

iv)   'Breaching a Stalking Protection Order', contrary to s8 of the Stalking Protection Act (2019).

When the original dataset was obtained, criminal offences in the MPS were recorded on the 'Crime Recording Information System' (CRIS) platform. This platform allows officers to classify the alleged

offence and record details such as victim, suspect, and witness information. For follow-up data drawn through manual search, once the follow-up period concluded on the 15th of August 2024, a change in MPS software had taken effect, and crimes were recorded on the newly adopted 'CONNECT' system. We used the 'Integrated Intelligence Platform' (IIP) to conduct manual re-offending searching, which could search the CONNECT system.

Data was obtained from CRIS for 31/03/2022-31/03/2024 for suspects named on any crime report classified under the four offences set out above and their victims. In addition, data were obtained for 2013-23 to allow a broader comparison when assessing the external validity of the dataset.

Inclusion/Exclusion Criteria

All suspects in treatment and control groups had committed a single stalking offence to qualify for random assignment and follow-up analysis allocation. Those repeat suspects in the sample were excluded because they have had a dissimilar baseline engagement with law enforcement to those with a single qualifying offence. From that point, a cleaning process then occurred, ensuring the data was suitable for transit outside the MPS to the University of Cambridge. The following fields were deleted and a 'sanitised' version created: suspect forename, surname, date of birth, address, the second half of postcode, grid reference; victim forename, surname, date of birth, address, the second half of postcode and grid reference. Two unique reference numbers were applied to assist with the analysis. Firstly, a 'MPS/XX' reference number was used for unique crime reports, allowing these to be traced through the data as a single entity. Secondly, all victims and suspects were provided with a URN to enable an analysis of repetition in the data sample and facilitate the use of Microsoft Excel's VLOOKUP function to identify suspects with victims by their sex, domestic abuse and other categories.

Several exclusion criteria were applied to the 2022-24 dataset before randomisation. The reasons for doing so are set out in Table 1. From an original sample of 4,443 offenders drawn from MPS systems for 31/3/22-31/3/24, applying the above exclusion criteria reduced the sample to 1,673 individuals named in a stalking allegation – an exclusion of 2,770 suspects.

Table 6 – Suspect Exclusion Criteria Applied and Rationale

Exclusion Criteria Applied	Rationale
Cases other than those listed under s2a and s4a of the Protection from Harassment Act (1997) (stalking) and Breaching a Stalking Protection Order contrary to s8 of the Stalking Protection Act (2019).	Whilst other offences, such as Breaching a Restraining Order, could arguably fall under the umbrella term 'stalking', they are not universally driven by stalking behaviours. These exclusion criteria were necessary to ensure that the cohort included in the sample was engaging in stalking behaviours.
Cases falling outside 31st March 2022 and 31st March 2024.	The intervention was planned for April 2024, and it was necessary to have recent data as most stalkers re-offend shortly after their index offence, necessitating the 31st of March 2024 as a cut-off date. Two years of data were obtained because it was deemed unethical to prompt recall of offending amongst stalkers whose offences were more than two years old and had a lower risk of re-offending.
Cases where the suspect was eliminated from enquiries under the heading 'Crime Related Incident'.	In cases where this outcome is listed, an offence has not been identified, excluding the presence of stalking behaviours.
Cases where the suspect was eliminated from enquiries under the heading 'Duplicate Crime Report'.	The same incident may sometimes prompt two separate reports as a recording error. In this case, including these cases may have resulted in offenders receiving two texts.
Cases where the suspect was eliminated from enquiries under the heading 'Entered in Error'.	Again, a suspect's name may be added to the wrong report due to recording errors, or their details may be entered incorrectly. In such cases, including the individual in this experiment would be unethical.
Cases where the suspect was eliminated from enquiries under the heading 'False Crime Report'.	A high threshold is applied to a decision to record a criminal allegation as false. In such circumstances, including individuals in the experiment would be unethical.

Cases where the suspect was eliminated from enquiries under the heading 'Not in the Public Interest.'	In such cases, officers have deemed it contrary to the public good for a criminal investigation to proceed. In such cases, intervening with a nudge text may prompt unpredictable outcomes and would be unethical for this study.
Cases where the suspect was eliminated from enquiries under the heading 'Suspect Deceased'.	A nudge text would not affect re-offending in such cases.
Crimes where the suspect was eliminated from enquiries under the heading 'Transferred to Another Force'.	Whilst stalking can span counties and even international boundaries, this RCT has been focused explicitly on offending within the London environment. A geographical limitation had to be applied, and as such cases would involve the offender residing outside of London, it was felt inappropriate to include such cases in this study.
Crimes where the suspect was eliminated from enquiries under the heading 'Statutory Body Investigation'.	This study is focused on interventions that police services can make to impact stalking re-offending. It was felt inappropriate to include any case where the police were not the investigating body.
Cases where no suspect mobile phone number was listed.	For no other reason than a nudge text could not be sent for these cases.
Cases where the suspect mobile number was <11 or >11 digits (excluding the +44 dialling code).	With 11 digits recorded, verifying the correct number sequence and preventing messages from going to the wrong parties is possible.
In cases where the suspect mobile number field listed a landline number, for example, beginning with '020'.	This would result in the nudge text not sending.
Cases where an international dialling code/phone number was listed as the suspect mobile number.	While it is accepted that offenders in London may use international codes and vice versa, UK mobile numbers, applying this exclusion criterion outside the UK, felt it appropriate to minimise the risk of influencing the behaviours of individuals outside of UK jurisdiction.

Sample size

As noted, the final sample comprised 1,673 individuals named in a stalking allegation. This is the entire population of eligible stalkers in London during the study period. As shown in Fig. 3 below, a test with one treatment and one control group of this size, assuming 80% power and .05 alpha level, is expected to detect statistically significant differences at d = 0.137, which is considered a small effect size under Cohen's (1980) criteria. Note that this group is unique, with idiosyncrasies that make it impossible to compare coherently with participant cohorts in other nudge intervention RCT's or RCT's more broadly. They may differ in innumerable ways from participants in trials considered in the literature review– for example, securing court attendance through SMS text messages. Therefore, this study's findings provide a benchmark for the sample size required in similar future studies to achieve a statistically significant effect.
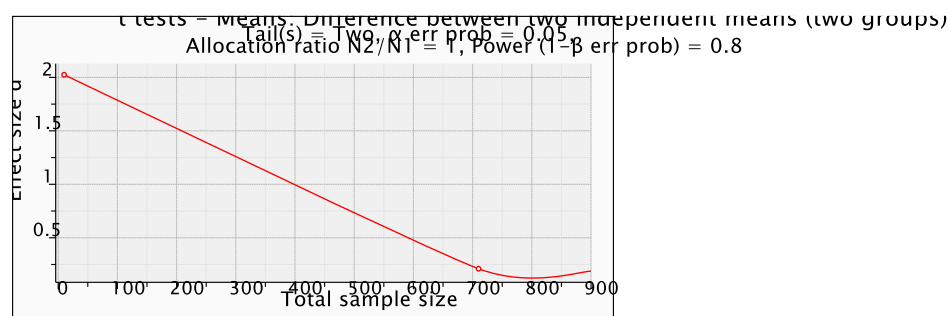

Figure 3 – Calculated Effect Size Based on a Sample of 1,673 participants

The University of Cambridge, Institute of Criminology, conducted randomisation, separating this process from the principal researcher and ensuring its integrity. Randomisation was completed using the Microsoft Excel randomisation function. Subsequent analysis was done on the following metrics to ensure a balance between experimental and control groups—which was indeed achieved for the significant prognosis variables: ethnicity; suspect sex; victim sex; whether or not a case was domestic abuse; the age category of the suspect; and the category of stalking offence alleged against them.

Interventions

As noted, the intervention in this study comprised of a text message sent to offenders' verified phone numbers. The opening line of the intervention reads, "You are receiving this message following your recent stalking allegation reported to the police". The intention is to demonstrate police awareness of the subject's behaviours and create an element of personalisation – "your recent stalking allegation".

The following lines read, "Stalking is an extremely harmful behaviour. We take stalking very seriously, especially when multiple offences are reported. You can end up in prison for repeated stalking offences". These lines introduced the concept of victim harm, which subsequent lines made more explicit. Reference is made to stalking as an organisational priority for the MPS, and the risk of imprisonment is highlighted. It should be noted here that an earlier iteration of the message recorded the risk of imprisonment, specifically, "maximum sentence of ten years' custody in the most serious cases". In the event, it was felt that doing so unnecessarily limited the perceived risk of imprisonment by limiting it numerically and indeed could be said to be misleading as most stalking cases would be unlikely to achieve this higher threshold without a series of aggravations or extensive antecedents – outlining that was clearly beyond the purview of a nudge text intended to be quick and effective.

Moreover, as noted in the literature review, Durlauf and Nagin (2011) have urged law enforcement to move away from a severity-based approach when communicating sanctions and reinforce the element of certainty. Two elements achieve this – organisational prioritisation of stalking and the *general* risk of imprisonment, as opposed to the length of sentence likely to be imposed in repeat stalking cases.

Blended with this core deterrence is a further paragraph which looks at victim impacts – the 'moral persuasion' element of the nudge text. The lines read, "Stalking can damage lives. Victims of stalking have been known to terminate employment or move home because of being stalked. We know many stalking victims have to get medical help for anxiety, nightmares, recurring flashbacks, and thoughts of self-harm as a result of being stalked. Many victims and their children exhibit signs of post-traumatic stress symptoms too". The theoretical basis of these points is covered extensively in the 'Literature Review' chapter and will not be reiterated here. This line is critical to the overall impact the message is intended to convey. Therapeutic treatment of stalkers focuses on helping offenders understand the harm their behaviour is causing.

It is worth considering this element across the broad spectrum of stalkers who inevitably would be represented in the eligible population for this study. MPS crime reports do not classify reports by the typology of stalker – intimacy seeker, incompetent suitor, resentful, rejected or predatory. Therefore, altering the content of the messages by these groups was beyond the scope of this study, as was an analysis of more than 1,700 reports to apply the typology based on a reading of the crime reports. Future studies, however, may benefit from considering whether all stalkers would be influenced to change their behaviour by understanding victim impact. Pathe and Mullen (2009) explore the concept of deconstructing the delusions some stalkers offend under and doing so in a therapeutic setting, with the structure and support offered by clinical psychology or other forms of professional help required. These lines sought to prompt a recognition that behaviour is harmful and – as the next section of the text will show – encourage them to seek professional help.

The final section of the nudge text flags the need for support. It reads, "If you are struggling to control your contact with someone against their wishes, help is available. You can speak to your GP, who can direct you to proper support. Many people have been able to get the appropriate help and avoided returning to stalking behaviours". This study found no charity in England and Wales supporting stalking perpetrators. Behavioural change programmes, often described as 'perpetrator interventions', exist in the context of domestic abuse but are accessed through multi-agency referral mechanisms and are often isolated to the highest-risk cases. Such programmes do not offer a mechanism for individuals struggling with fixated stalking behaviours to reach out for support. As such, the appropriate mechanism in England and Wales is to speak to a General Practitioner (GP) who can then signpost offenders to appropriate support.

Having finalised the design of the nudge text, the mobile phone numbers for the experimental group (n = 817) were isolated in a separate CSV file and sent to the MPS digital team with a copy of the final nudge text. Before this, the principal researcher did a final check of numbers, removing any negative space between numbers so that all comprised a continuous eleven-digit number and, where the UK dialling code '+44' or other variants were found, replacing this with a '07'. Ten digits were identified, and the number began with a '7' and a '0', placed before the number to ensure it could be read by the mass-text system, 'WireFast'. 'WireFast' was then used to upload the text message content and individual numbers ready for sending.

The text message was delivered at 1137 a.m. on Monday, 15th April 2024. This date was selected deliberately. As mentioned, the data used ran from 31/03/2022 to 31/03/2024, meaning some allegations in the sample were very recent. A two-week post-allegation period allowed investigating officers to conduct initial risk management. This was balanced with the need to deliver the intervention in the initial six-month window, where it was deemed that re-offending was most likely to occur. A period of two and a half weeks followed up to the 2nd May 2024 before requesting a 'send

report' from WireFast to allow all text messages trying to send to go through – for example, if a suspect did not turn their phone on immediately.

Treatment fidelity

A high level of treatment fidelity is critical for reliable nudge text RCT findings. If the nudge texts are successfully sent to most treatment participants, we can show causality between the text and resultant behaviour. If it sends to a low proportion, we cannot. Cumberbatch and Barnes' study (2018) of nudge texts encouraging victim and witness attendance at court showed no statistically significant difference between treatment and control groups. The latter achieved 24% court attendance rates; the former 22% against a base rate attendance rate of 26%. As treatment fidelity was 84%, their findings were reliable.

Conversely, Chivers and Barnes' 2018 nudge text RCT looking at defendant attendance at court, which likewise found no statistically significant difference between treatment and control overall or through subgroup moderator analysis, achieved low treatment fidelity with 62% of numbers uncontactable or invalid (Chivers & Barnes, 2018). However, when researchers analysed findings for valid numbers in isolation, the study found a promising but non-significant boost in court attendance rates, concluding, "personalised SMS reminders, activating social norms and deterrence, delivered at a timely moment may increase the attendance rates of defendants" (Ibid, p.7). Thus, high treatment fidelity can return valid findings for no, small or large effect sizes. This study adopts the Metropolitan Police Service 'Wirefast' system, which allows an analysis of which text messages were not sent, remained stuck in transit or were successfully delivered.

A report was returned to the research team identifying the number of messages successfully delivered – or 'treatment fidelity', as shown in Table 7. Out of n = 817 messages attempted, n = 524 were delivered successfully, and n = 293 were not. Those which were sent, 32.5% (n = 266) were shown as

'Unable to Deliver', which indicates a correct destination number but, for example, the suspect's phone was switched off. Only 0.12% (n = 1) was shown as 'Invalid destination address', suggesting a high rate of recording accuracy for suspect phone numbers in the treatment group. For 6 participants, unknown reasons for no delivery were recorded. Finally, an error by the principal researcher resulted in 2.4% (n = 20) of suspects in treatment being allocated without a valid mobile phone number recorded. This means that treatment fidelity was achieved for 64.1% (n = 524) of participants in the treatment group.

Table 7 – Treatment Fidelity Rates in the Treatment Group

Outcome:	No.	(%)
Delivered Successfully	524	64.1%
Failed 1 - Unable to deliver	266	32.5%
Allocated to treatment without mobile number	20	2.4%
Delivery report not available	6	0.73%
Failed 2 - Invalid destination address	1	0.12%
Total	**817**	**100%**

Outcomes

Following the post-intervention period on the 15th of August 2024, a manual search of Metropolitan Police crime indices was conducted for all suspects in treatment and control, covering the period 15th April – 15th August 2024. Two outcome measures were searched for – prevalence and harm. Firstly, if a suspect had committed any of the stalking offences used in this study, they would receive a binary score (1 for re-offending, 0 for not re-offending) to determine if further stalking had occurred. This figure allows a total for treatment and control for re-offending prevalence in each group.

Secondly, for each suspect who re-offended, the number of offences for each stalking category (course of conduct, serious alarm and distress, fear of violence, breaching a Stalking Protection Order) was totalled. This allowed a total Cambridge Crime Harm Index (CCHI) score to be assigned to each participant, as seen in Table 8. The CCHI was used to assign a harm score to each offence type. The CCHI operates by assigning a score to each offence based on the number of imprisonment days a

suspect could expect to receive with no previous convictions or offending aggravations. Notably, as no specific harm score was assigned to the 'Breach of Stalking Protection Order', and CCHI guidance directs researchers to assign a similar offence score, the score of '5' for the 'Breach of Restraining Order' was assigned for this offence.

Table 8 – CCHI scores assigned to each stalking offence

Stalking Offence Type	CCHI Score
S4a Serious Alarm and Distress	252
S4a Fear of Violence	84
S2a Course of Conduct Stalking	10
Breach of a Stalking Protection Order	5

Statistical Methods

Interpretation of findings for follow-up analysis was achieved through a *t*-test. The *t*-test is "a parametric test to compare the means of two samples which can be related…or they can be unrelated samples" (Marshall and Jonker, 2011, p.e4). In this study, we are concerned with comparing the mean number of suspects re-offending (prevalence rates) and the mean level of crime harm between treatment and control and understanding whether the observed difference is at the level of statistical significance – i.e. can be extrapolated beyond this study's sample. Ariel et al. (2022) recommend this approach for field RCT's in policing.

For re-offending prevalence, the mean and standard deviation of the number of suspects re-offending were calculated for treatment and control across the primary and all secondary outcomes. The sample size for each primary and secondary outcome across treatment and control was then observed, generating the t and p values. The same was replicated for crime harm scores, but the mean crime harm score was used instead.

As standard, the level of statistical significance was set at $p \leq .05$. This allowed a determination to be made of whether any of the differences between treatment and control were at the level of statistical significance.

A similar statistical approach was used for the subgroup analyses, observing the treatment effect in groupings based on gender, age brackets, and victim-offender relationships.

Chapter III

Findings

<u>Participant Flow</u>

As shown in the participants' flowchart, a sample of 1,673 suspects was randomised into treatment and control conditions. Of those suspects, 1,639 were unique offenders, and 34 were repeat offenders. Their inclusion in the randomised sample resulted in the following distribution. For the 30 suspects with 2 offences in the sample, 8 were duplicated in treatment, 5 were duplicated in control, and 17 occurred once in both treatment and control. For the 4 suspects with 3 offences, 3 were duplicated in treatment and occurred once in control, and 1 was duplicated in control and occurred once in treatment. No suspects with 3 offences appeared three times in treatment or control only. Therefore, although there were only 34 repeat suspects in the randomised data, overall, they represented a count of 26 suspects in the control group and 29 suspects in the treatment group – 55 in total – because 13 appeared only within either treatment (n = 8) or control (n = 5) whereas 21 suspects appeared in both groups.

A further decision was taken to remove participants in the treatment group who did not receive the intervention text from the follow-up analysis (n = 293). Whilst this study could have approached findings from an 'Intention to Treat' methodology – to explore the actual process of sending nudge texts, which would inevitably include an error rate – the risks inherent in stalking necessitate this study to present findings on the effectiveness of the nudge text, with findings not skewed by the inclusion of suspects' re-offending behaviours which weren't influenced by the intervention itself.
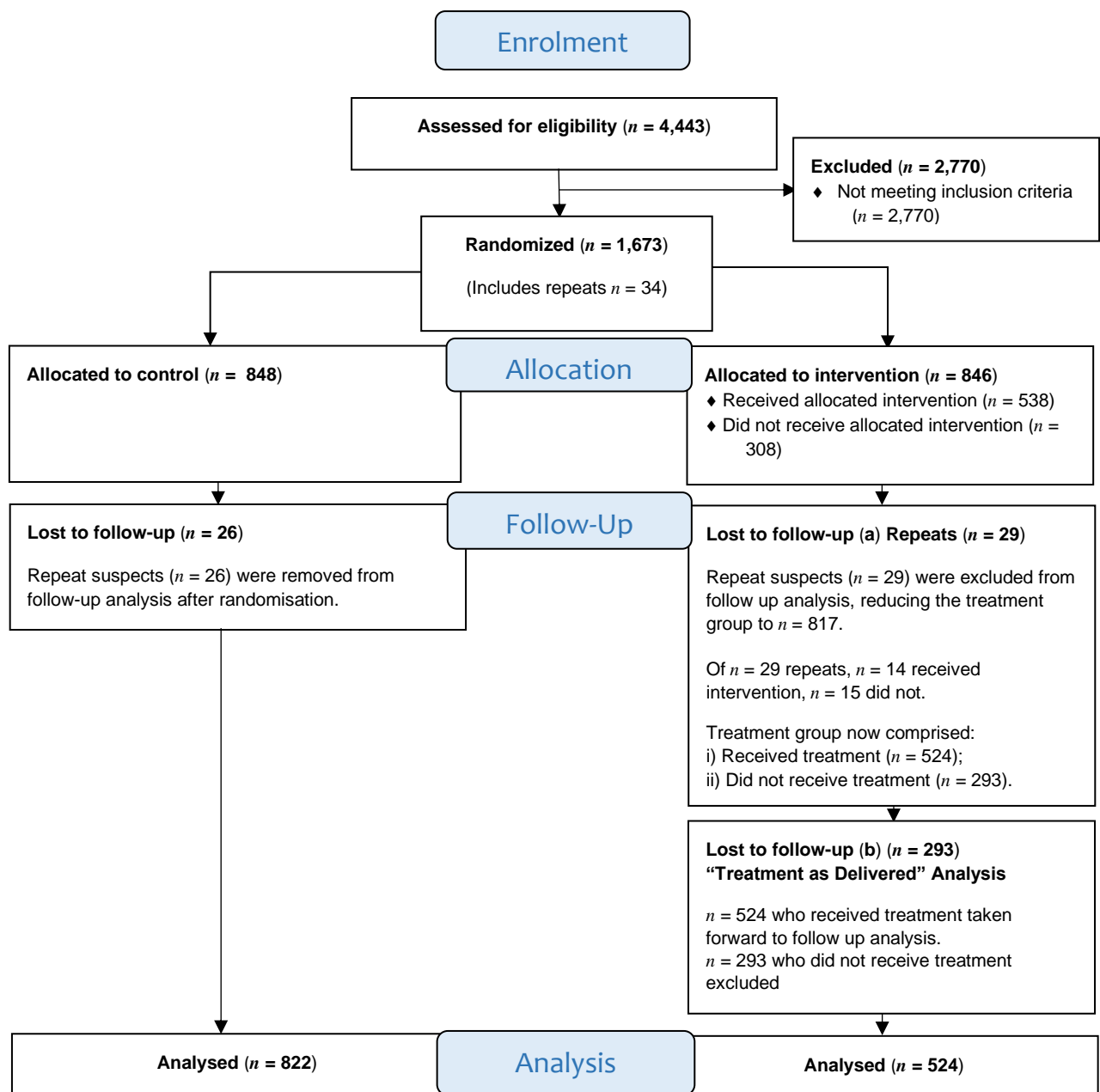
Figure 4 – CONSORT 2010 Flow Diagram showing participant progression through the trial (CONSORT, 2024)[1]

---

[1] Original analysis for the study included both 'Intention to Treat' and 'Treatment as Delivered' findings, with a comparison of effect sizes. Following discussion at the Thesis Oversight Committee, it was decided that only 'Treatment as Delivered' analysis would be presented in the study, to ensure findings represented the efficacy of the intervention specifically. It is accepted that in real-world conditions, law enforcement agencies implementing this policy would experience text message failure rates similar to those seen in this experiment.

Period of Recruitment, Follow-up Period and Rationale

Looking at the sample of suspects before exclusion criteria being applied, there were 4,443 in total. Of those, 3.8% (n = 173) were repeat suspects – i.e. received more than one stalking allegation between 31/3/22 and 31/3/24. This is important as it teaches us about the likelihood of repeated offending in this crime type.

Table 9 – The number of suspects in the pre-exclusion criteria sample by how many stalking allegations they received from 31/3/22-31/3/24

No. of Stalking Offences	No. of Suspects	(%)
1	4,270	96.11%
2	138	3.11%
3	25	0.56%
4	5	0.11%
5	4	0.09%
6	0	0.00%
7	1	0.02%

The time between the first and second allegation is a typical measurement for all suspects—whether they committed 2 or 7 offences—and is used to estimate the length of time between offending we observe in the cohort. This measurement is also a common value for all repeat suspects.
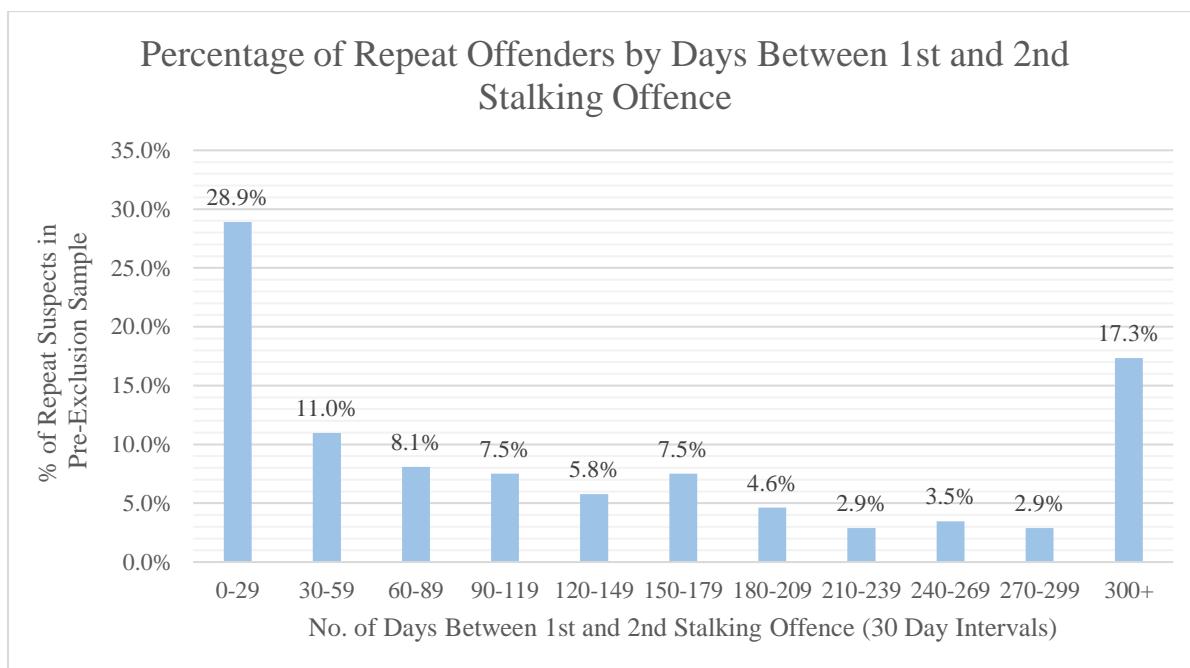
Percentage of Repeat Offenders by Days Between 1st and 2nd Stalking Offence

Figure 5 — Time lapse between first and second stalking offence

As shown in Fig. 5, the mean average days to commit a second offence was 238, ranging from 0-7,531 days. There are three outliers where historic offences were reported within inclusion dates: 2,539 days, 4,731 days, and 7,531 days. If we remove those suspects from analysis, we find a mean average days to commit a second offence of 155 days with a range of 0-925 days. Overall, we observe that the most significant proportion of suspects at 28.9% re-offend within 0-29 days—the remainder re-offend within 2 years, with a small number of outliers re-offending thereafter.

These re-offending characteristics provide a rationale for recruiting suspects between 31/3/22 and 31/3/24. At the same time, they suggest that a follow-up period of 4 months (15/4/24-15/8/24) may have been too short, which is explored further in the discussion chapter.

Baseline Characteristics of the Sample

Ethnicity.

As shown in Table 10 and Fig. 6, in the treatment group, 'White European' represented 23.5% (n = 192) compared with 18.4% (n = 151) in control. All BAME ethnicity categories accounted for 25.6% (n = 209) in treatment compared with 31.6% (n = 260) in control. Suspects whose ethnicity was actively recorded by officers as 'Unknown' represented 1.2% (n = 10) in treatment compared with 1.3% (n = 11) in control. Finally, suspects for whom the ethnicity field was not completed or 'Not Recorded' accounted for 49.7% (n = 406) in treatment and 48.5% (n = 399) in control, accounting for nearly half the sample and representing a significant gap in demographic data for this trial.

Gender.

In the treatment group, males comprised 78.3% (n = 640), females 15.5% (n = 127) and 'Unknown' 0.7% (n = 6). Suspects whose sex was not recorded comprised 5.4% (n = 44) of the treatment group. Looking at the control group, males comprised 76.3% (n = 627), females 16.1% (n = 132) and 'Unknown' 1.1% (n = 9). Suspects in the control group whose sex was not recorded comprised 6% (n = 54) of the sample.

It is further possible to set out the sex of the suspect and victim with four variables: male suspect-female victim, male suspect-male victim, female suspect-male victim, and female suspect-female victim. Male suspect-female victims comprised 71.5% (n = 584) in treatment compared with control at 70.9% (n = 583). Male suspect-male victims comprised 5.1% (n = 42) in treatment compared with 3.8% (n = 31) in control. Female suspect-male victims comprised 10.0% (n = 82) in treatment compared with 9.6% (n = 79) in control. Female suspect-female victims comprised 5.3% (n = 43) in treatment compared with 6.3% (n = 52) in control. Cases where either suspect or victim sex was not recorded comprised 8.1% (n = 66) in treatment and 9.4% (n = 77) in control.

Table 10 – Baseline characteristics of the treatment and control groups:

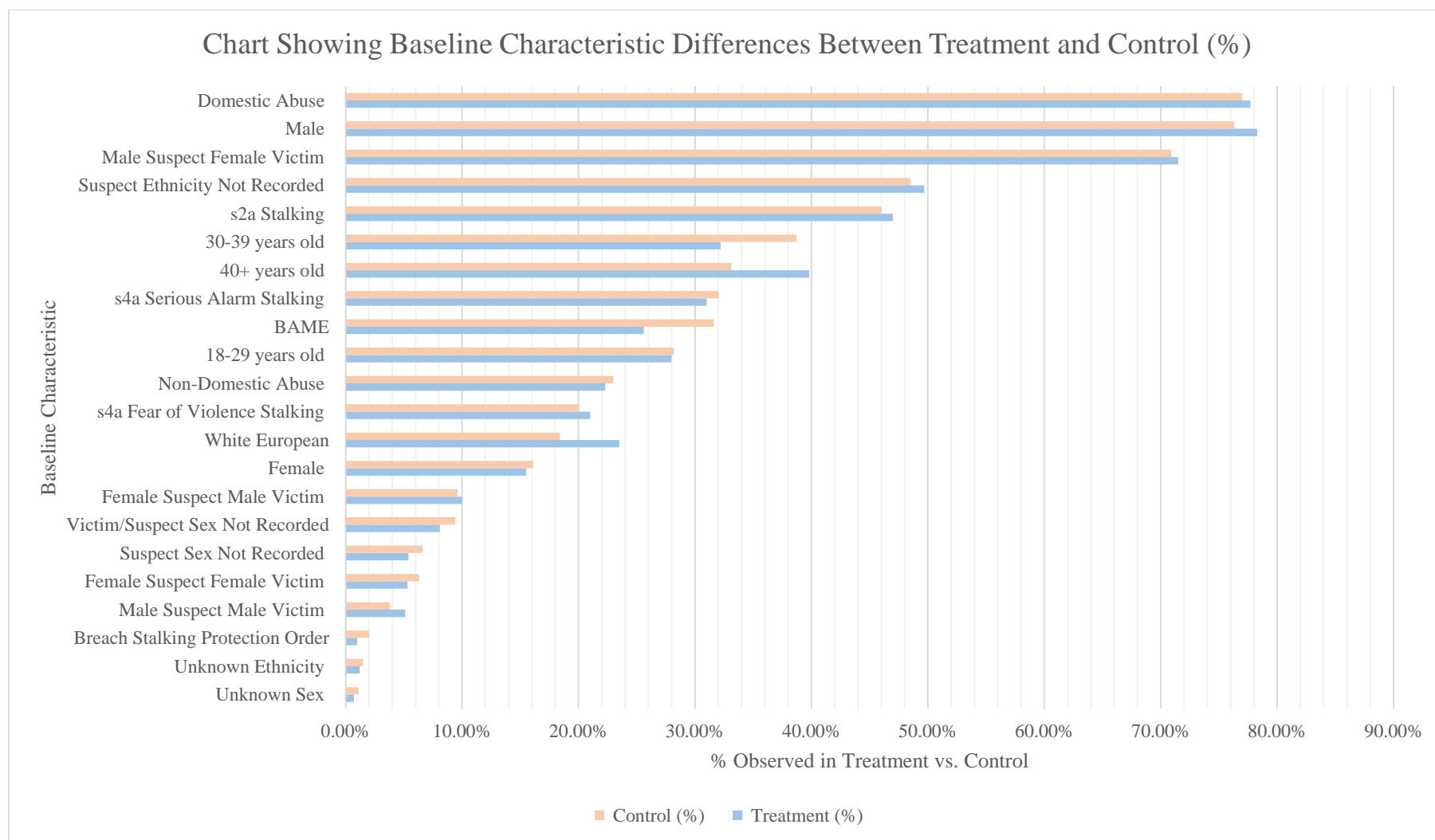|	Baseline Characteristic:	Treatment Group		Control Group	
		Suspects (*n*)	(%)	Suspects (*n*)	(%)
Ethnicity	White European	192	23.5%	151	18.4%
	BAME	209	25.6%	260	31.6%
	Unknown	10	1.2%	12	1.5%
	Not Recorded	406	49.7%	399	48.5%
Sex	Male	640	78.3%	627	76.3%
	Female	127	15.5%	132	16.1%
	Unknown	6	0.7%	9	1.1%
	Not Recorded	44	5.4%	54	6.6%
Victim-Suspect Sex	Male Suspect Female Victim	584	71.5%	583	70.9%
	Male Suspect Male Victim	42	5.1%	31	3.8%
	Female Suspect Male Victim	82	10.0%	79	9.6%
	Female Suspect Female Victim	43	5.3%	52	6.3%
	Victim/Suspect Sex Not Recorded	66	8.1%	77	9.4%
DA?	Domestic Abuse	635	77.7%	633	77.0%
	Non-Domestic Abuse	182	22.3%	189	23.0%
Suspect Age	18-29 years old	229	28.0%	232	28.2%
	30-39 years old	263	32.2%	318	38.7%
	40+ years old	325	39.8%	272	33.1%
Qualifying Offence	s2a Stalking	380	47.0%	379	46.0%
	s4a Serious Alarm Stalking	256	31.0%	267	32.0%
	s4a Fear of Violence Stalking	174	21.0%	163	20.0%
	Breach Stalking Protection Order	7	1.0%	13	2.0%

Figure 6 – Comparing ratios of baseline characteristics observed in the treatment and control groups

Age.

Age was divided into three categories. Suspects aged 18-29 comprised 28.0% (*n* = 229) in treatment compared with 28.2% (*n* = 232) in control. Suspects aged 30-39 comprised 32.2% (*n* = 263) in treatment compared with 38.7% (*n* = 318) in control. Suspects aged 40 or older comprised 39.8% (*n* = 325) in treatment compared with 33.1% (*n* = 272) in control. As seen in Fig. 7, dividing suspect ages into intervals of 5 years provides a more detailed comparison of treatment and control whilst also demonstrating the age curve for suspects of stalking.
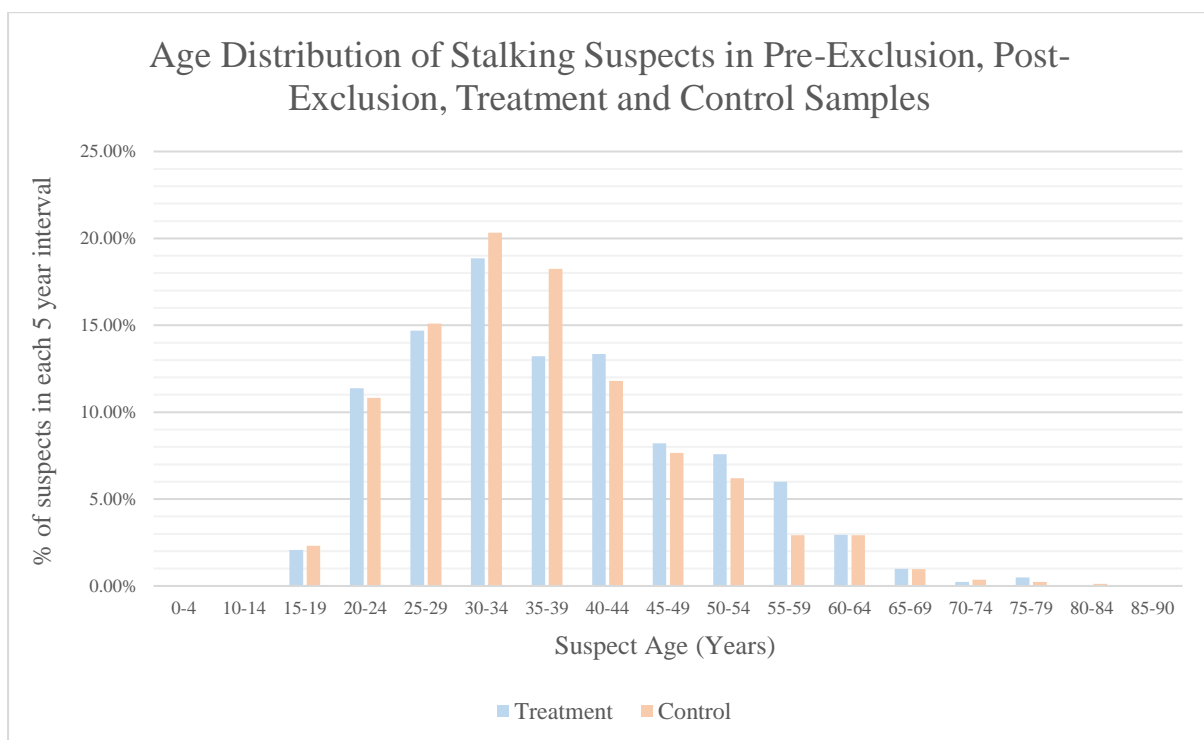

Figure 7 – The distribution of suspect ages in 5-year intervals: treatment vs. control groups

Relationship between Victim and Offender.

Cases of domestic abuse stalking comprised the greatest proportion of the sample in both treatment and control. Domestic abuse cases were represented at 77.7% (*n* = 635) in treatment and 77.0% (*n* =

633) in control. Non-domestic abuse cases were represented at 22.3% (n = 182) in treatment and 23.0% (n = 189) in control.

As set out in the Consort participant flowchart above, suspects who committed more than one offence were excluded from the sample. Of those remaining, a single stalking offence qualified them for inclusion in the trial. No difference of >1% in the proportion of offences was observed between treatment and control. Section 2a Course of Conduct stalking was represented at 47% (n = 380) in treatment and 46% (n = 379) in control. Section 4a Serious Alarm stalking was defined at 31% (n = 256) in treatment and 32% (n = 267) in control. Section 4a Fear of Violence stalking was represented at 21% (n = 174) in treatment and 20% (n = 163) in control. Breach of a Stalking Protection Order was represented at 1% (n = 7) in treatment and 2% (n = 13) in control.

Primary and Secondary Outcomes and Estimations

Table 11 – Outcomes: sample size, n of suspects who re-offended, re-offending rate, mean number of re-offending suspects, standard deviation, *t* score and *p-value* for each primary and secondary outcome

Outcome Type (Primary/Secondary)	Outcome Description	Treatment			Control			Mean no. of suspects re-offending		Standard Deviation		Outcomes	
		Treatment Sample (*n*)	Re-offending (*n*)	Re-offending (%)	Control Sample (*n*)	Re-offending (*n*)	Re-offending (%)	Treatment	Control	Treatment	Control	*t* score	*p-value*
Primary	All Suspects	524	22	4.2%	822	29	3.5%	0.04	0.04	0.20	0.18	0.6163	0.5378
Secondary	Male Suspects	407	18	4.4%	627	18	2.9%	0.04	0.03	0.21	0.17	1.2727	0.2035
	Female Suspects	83	3	3.6%	132	9	6.8%	0.04	0.07	0.19	0.25	-1.0621	0.2895
Secondary	Suspects with Female Victims	427	20	4.7%	690	20	2.9%	0.05	0.03	0.21	0.17	1.4793	0.1394
	Suspects with Male Victims	87	2	2.3%	116	8	6.9%	0.02	0.07	0.15	0.25	-1.6060	0.1098
Secondary	Domestic Abuse Suspects	410	20	4.9%	633	24	3.8%	0.05	0.04	0.22	0.19	0.8307	0.4064
	Non-Domestic Abuse Suspects	114	2	1.8%	189	5	2.6%	0.02	0.03	0.13	0.16	-0.5244	0.6005
Secondary	Suspects (18-29 years old)	135	5	3.7%	232	9	3.9%	0.04	0.04	0.19	0.19	-0.0850	0.9323
	Suspects (30-39 years old)	165	9	5.5%	318	10	3.1%	0.05	0.03	0.23	0.17	1.1405	0.2550
	Suspects (40+ years)	224	8	3.6%	272	10	3.7%	0.04	0.04	0.19	0.19	-0.0622	0.9504
Secondary	Male suspects re-offending up to 179 days from the first offence to intervention	32	1	3.1%	26	2	7.7%	0.03	0.08	0.18	0.27	-0.7393	0.4634
	Female suspects re-offending up to 179 days from first offence to intervention	11	0	0.0%	22	3	13.6%	0.00	0.14	0.00	0.35	-1.2773	0.1055
	Male suspects re-offending 180+ days from first offence to intervention	329	14	4.3%	541	14	2.6%	0.04	0.03	0.20	0.16	1.2756	0.2026
	Female suspects re-offending 180+ days from first offence to intervention	70	2	2.9%	106	6	5.7%	0.04	0.06	0.20	0.23	-0.4139	0.6795

Table 12 - Sample size, cumulative Cambridge Crime Harm Index (CCHI), variance in CCHI in Treatment compared with Control, mean CCHI score, standard deviation, *t* score and *p-value* for each primary and secondary outcome

| Outcome Type (Primary/Secondary)	Outcome Description	Treatment Sample (*n*)	Total CCHI Score	Control Sample (*n*)	Total CCHI Score	Reduction in CCHI Score Treatment to Control	% Reduction	Mean CCHI Score		Standard Deviation		Outcomes	
								Treatment	Control	Treatment	Control	*t* score	*p-value*
Primary	All Suspects	524	1341	822	2542	-1201	-47.25%	2.56	3.09	21.16	25.52	-0.4156	0.6778
	Male Suspects	407	1306	627	1548	-242	-15.63%	3.21	2.47	23.96	23.01	0.4927	0.6223
Secondary	Female Suspects	83	30	132	900	-870	-96.67%	0.36	6.82	1.88	38.45	-1.9255	**0.0561**
	Suspects with Female Victims	427	1321	690	1637	-316	-19.30%	3.09	2.37	23.40	22.17	0.5107	0.6097
Secondary	Suspects with Male Victims	87	20	116	890	-870	-97.75%	0.23	7.67	1.51	40.96	-1.9553	**0.0528**
	Domestic Abuse Suspects	410	1247	633	2003	-756	-37.74%	3.04	3.16	23.54	25.35	-0.0786	0.9374
Secondary	Non-Domestic Abuse Suspects	114	94	189	539	-445	-82.56%	0.82	2.85	7.91	26.13	-0.9939	0.3211
	Suspects (18-29 years old)	135	366	232	1043	-677	-64.91%	2.71	4.50	22.83	29.97	-0.6418	0.5215
Secondary	Suspects (30-39 years old)	165	658	318	1073	-415	-38.68%	3.99	3.37	24.69	28.28	0.2463	0.8056
	Suspects (40+ years)	224	317	272	426	-109	-25.59%	1.42	1.57	16.90	16.35	-0.1005	0.9200
	Male suspects re-offending up to 179 days from the first offence to intervention	32	10	26	267	-257	-96.25%	0.31	10.27	1.77	49.39	-1.0274	0.3137
	Female suspects re-offending up to 179 days from first offence to intervention	11	0	22	114	-114	-100.00%	0.00	5.18	0.00	20.05	0.8502	0.2009
Secondary	Male suspects re-offending 180+ days from first offence to intervention	329	1024	541	1019	5	0.49%	3.11	1.88	22.78	19.54	0.8133	0.4163
	Female suspects re-offending 180+ days from first offence to intervention	70	30	106	786	-756	-96.18%	0.43	7.42	2.04	41.97	-1.7107	**0.0897**
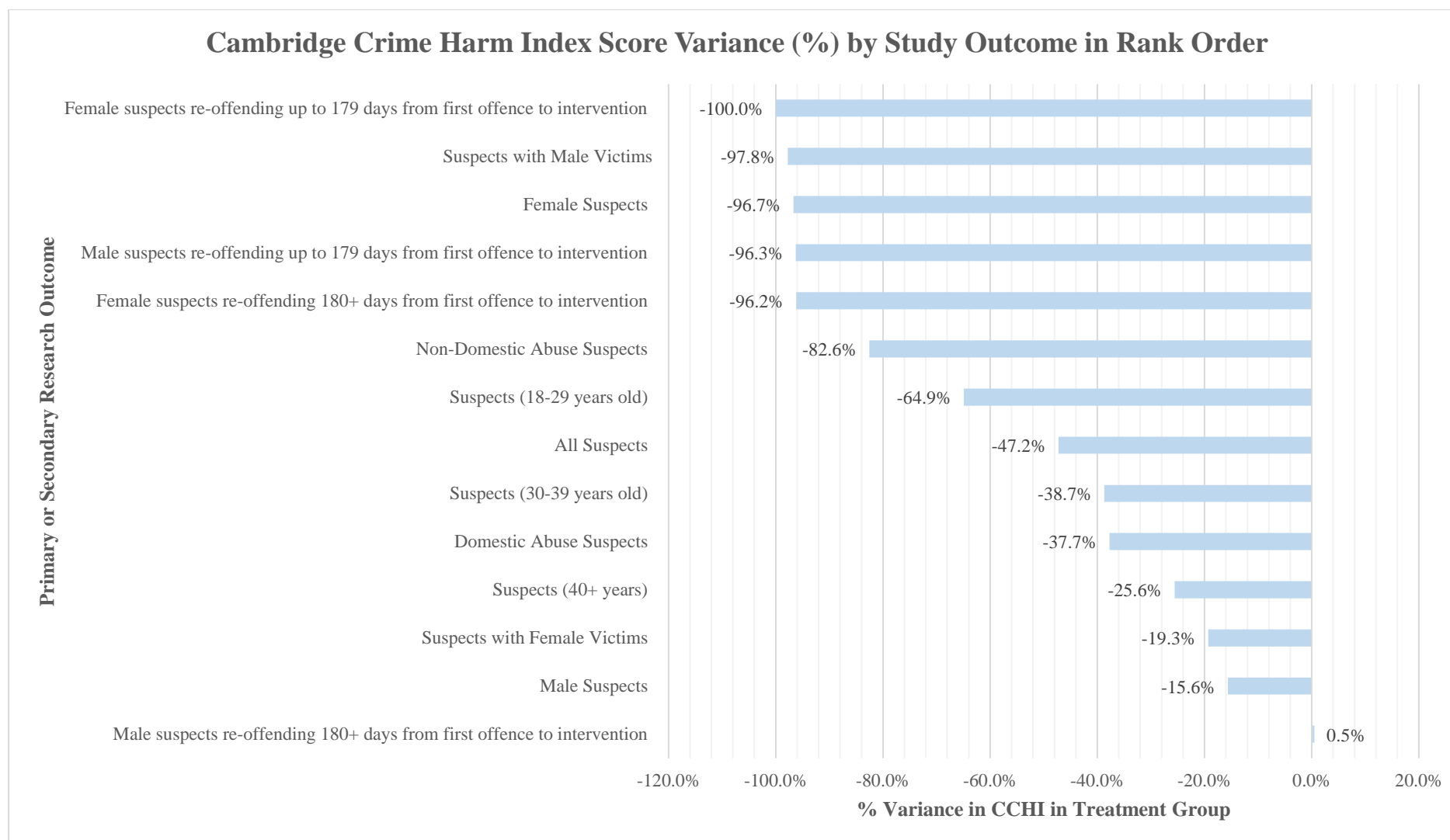
Figure 8 - Showing CCHI score variance observed as a result of treatment application for 'Treatment as Delivered' analysis

Primary and Secondary Outcomes:

Primary Research Question - Does the delivery of a nudge text intervention that combines moral persuasion with deterrence cause a reduction in stalker re-offending and Cambridge Crime Harm Index score compared to no intervention?

Prevalence of Re-offending:

No statistically significant difference between treatment and control groups was observed in re-offending prevalence. The treatment group (n = 524, M = 0.04, SD = 0.20) re-offended at a rate of 4.2% (n = 22) whereas the control group (n = 822, M = 0.04, SD = 0.18) re-offended at a rate of 3.5% (n = 29), p = .5378.
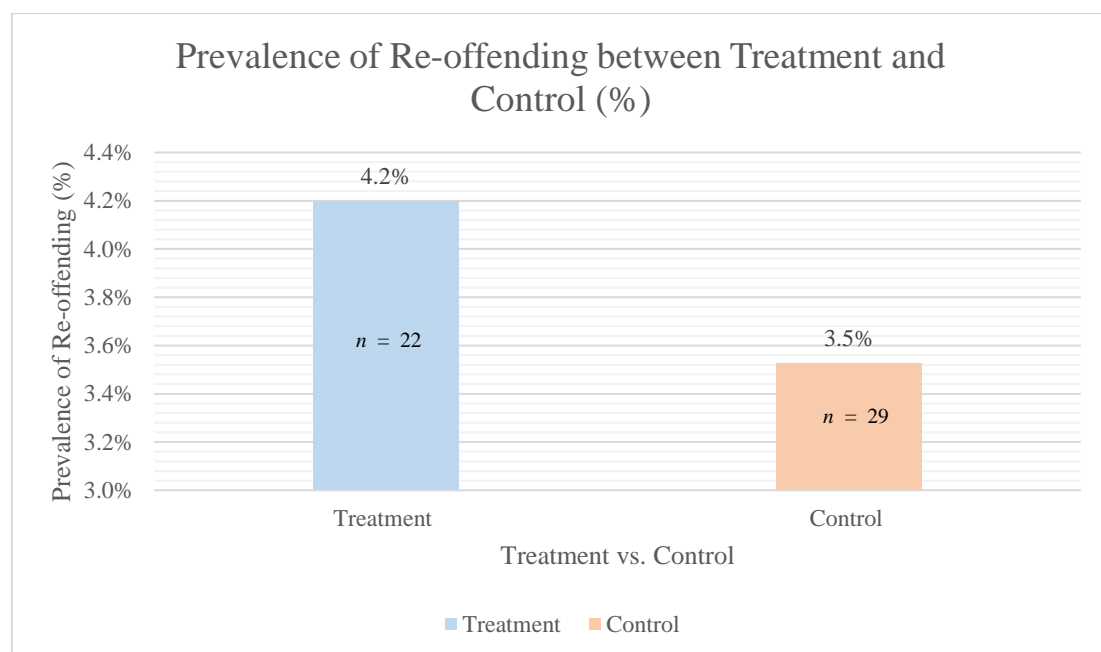
Figure 9 – Prevalence of re-offending in treatment and control groups

Cambridge Crime Harm Index Scores:

No statistically significant difference was observed between the treatment and control groups for the variance of the CCHI score. For the treatment group (n = 524, M = 2.56, SD = 21.16), a CCHI score of 1,341 was observed compared with the control (n = 822, M = 3.09, SD = 25.52), which returned a CCHI

score of 2,542, p = .6778. When applying the treatment, this represented a non-statistically significant CCHI score reduction of 47.25% (n = 1,201).

Secondary Research Question 1 — Does delivering a nudge text intervention that combines moral persuasion with deterrence reduce stalker re-offending and Cambridge Crime Harm Index score compared to no intervention, <u>comparing male and female suspects</u>?

Prevalence of Re-offending:

No statistically significant difference was observed in re-offending prevalence when comparing male and female suspects. The treatment group saw male suspects (n = 407, M = 0.04, SD = 0.21) re-offend at a rate of 4.4% (n = 18), whereas the control group saw male suspects (n = 627, M = 0.03, SD = 0.17) re-offend at a rate of 2.9% (n = 18), p = .2035. The treatment group saw female suspects (n = 83, M = 0.04, SD = 0.19) re-offend at a rate of 3.6% (n = 3) compared with the control group which saw female suspects (n = 132, M = 0.07, SD = 0.25) re-offend at a rate of 6.8% (n = 9), p = .2895, suggesting a non-statistically significant crime reduction effect for female suspects who got the treatment.
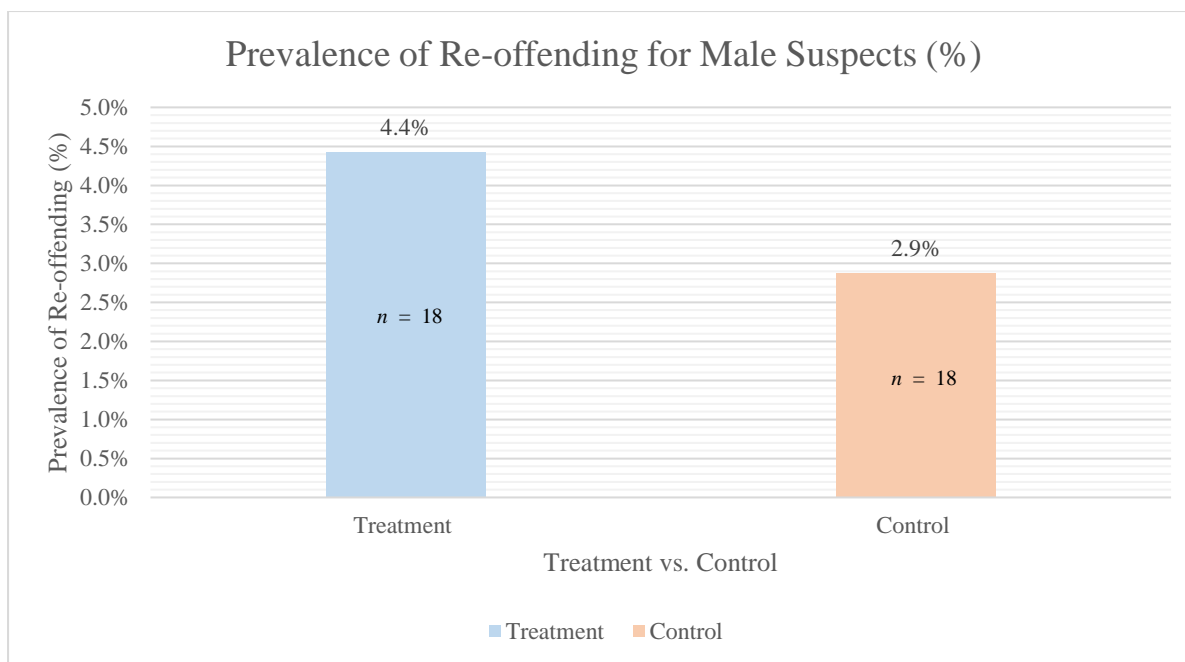
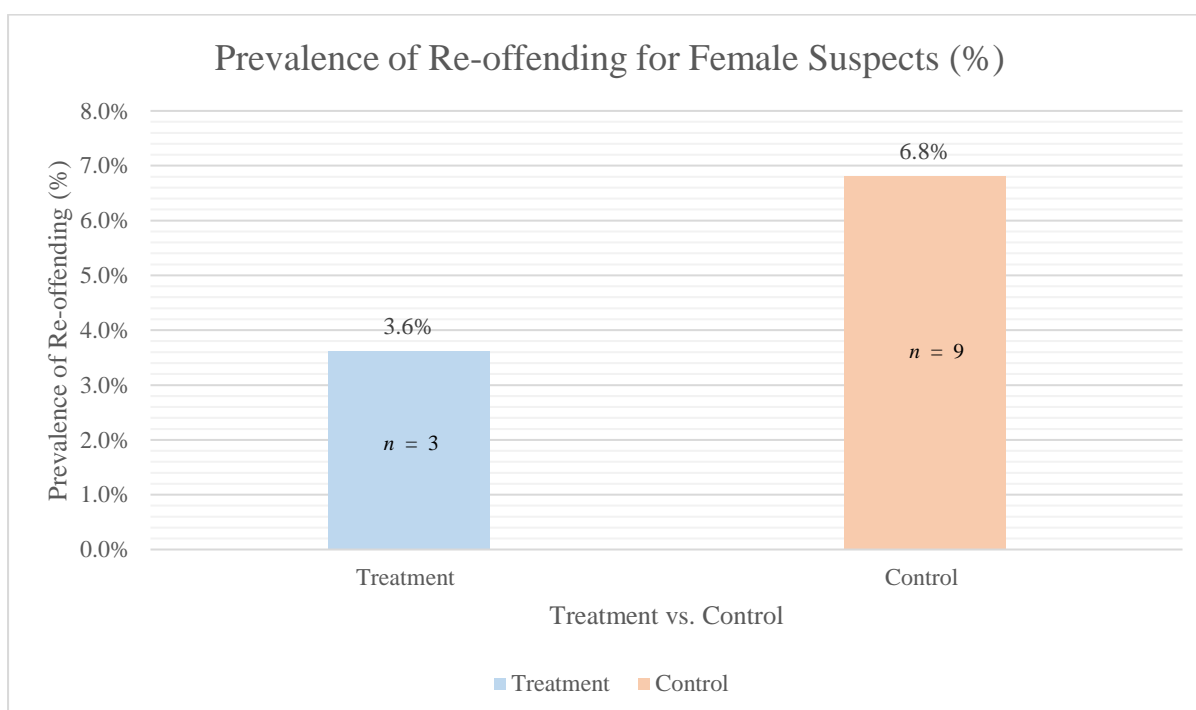**Figure 10– Prevalence of re-offending for male suspects between treatment and control**

Figure 11 - Prevalence of re-offending for female suspects between treatment and control

Cambridge Crime Harm Index Score:

No statistically significant variance in CCHI score was observed between treatment and control groups

when comparing sex. For male suspects in the treatment group (n = 407, M = 3.21, SD = 23.96), a CCHI

score of 1,306 was observed compared with the control group (n = 627, M = 2.47, SD = 23.01), where a CCHI score of 1,548 was observed, p = .6223. This represented a non-statistically significant reduction in the CCHI score of 15.63% (n = 242). For female suspects in the treatment group (n = 83, M = 0.36, SD = 1.88), a CCHI score of 30 was observed, compared with the control group (n = 132, M = 6.82, SD = 38.45), where a CCHI score of 900 was followed, p = .0561. This value came close to statistical significance but ultimately is a non-statistically significant reduction of the CCHI score of 96.67% (n = 900).

Secondary Research Question 2 — Does the delivery of a nudge text intervention that combines moral persuasion with deterrence reduce stalker re-offending and Cambridge Crime Harm Index score compared to no intervention, comparing suspects with <u>female victims to suspects with male victims?</u>

Prevalence of Re-offending:

Suspects with female victims in the treatment group (n = 427, M = 0.05, SD = 0.21) re-offended at a rate of 4.7% (n = 20) compared with control (n = 690, M = 0.03, SD = 0.17) where suspects re-offended at a rate of 2.9% (n = 20), p = .1394. Suspects with male victims in the treatment group (n = 87, M = 0.02, SD = 0.15) re-offended at a rate of 2.3% (n = 2), whereas in the control group (n = 116, M = 0.07, SD = 0.25) suspects re-offended at a rate of 6.9% (n = 8), p = .1098. Again, we see a backfire effect for suspects with female victims and a crime reduction effect for those with male victims, but not at the level of statistical significance when isolating those who received the intervention.
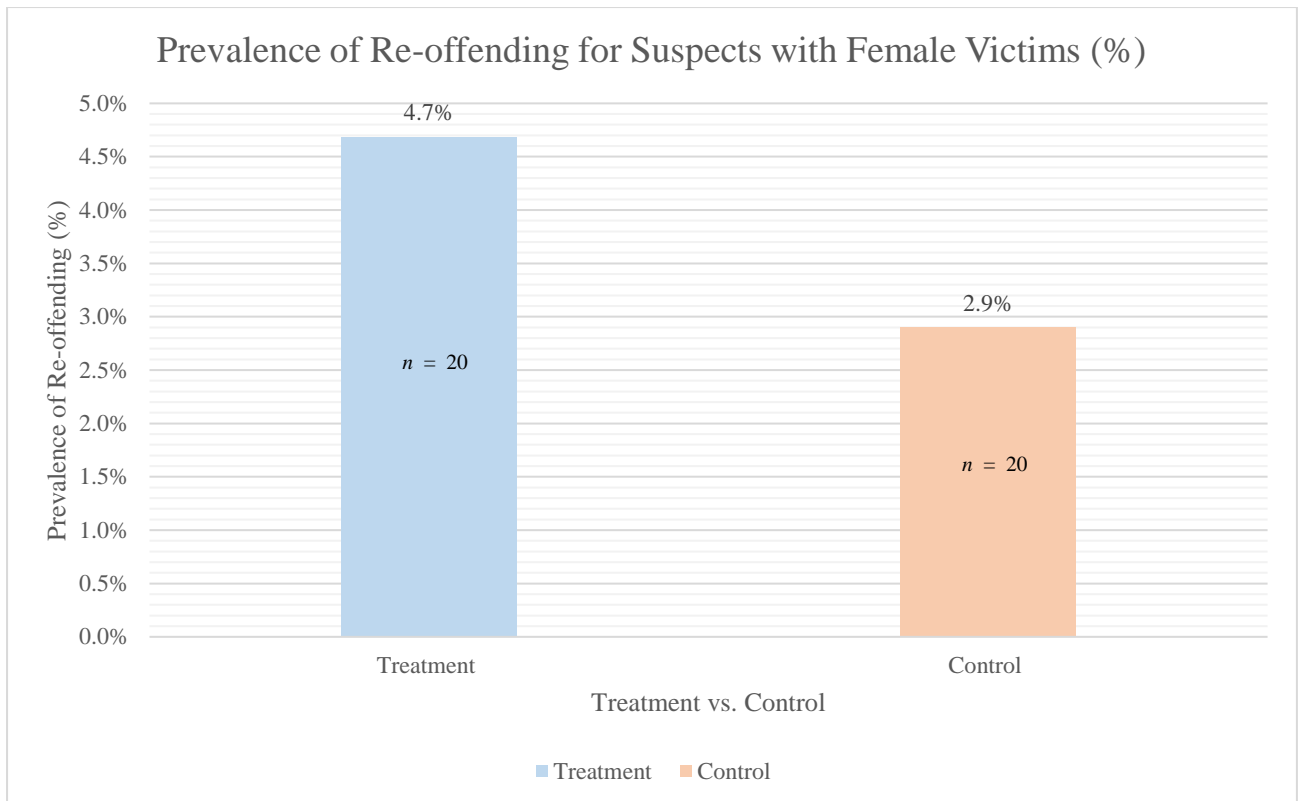
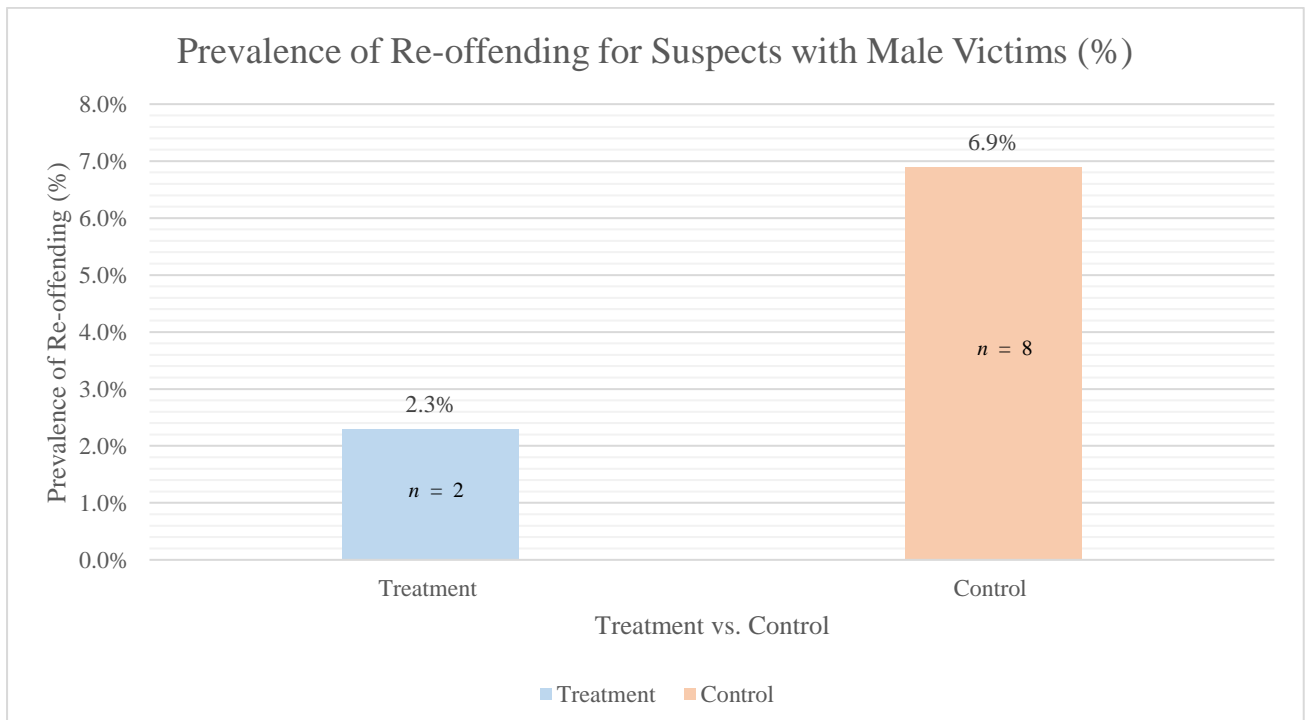**Figure 12 – Showing the prevalence of re-offending for suspects with female victims, comparing treatment and control**

Figure 13 – Showing the prevalence of re-offending for suspects with male victims comparing treatment and control

Cambridge Crime Harm Index Score:

For suspects with female victims in the treatment group (*n* = 427, *M* = 3.09, *SD* = 23.40), we observe a CCHI score of 1,321 compared with control (n = 690, M = 2.37, SD = 22.17), where we observe a CCHI score of 1,637, p = .6097. We, therefore, observe a CCHI reduction of 19.3% (n = 316), not at the level of statistical significance. For suspects with male victims in the treatment group (*n* = 87, *M* = 0.23, *SD* = 1.51), we observe a CCHI score of 20 compared with the control (n = 116, M = 7.67, SD = 40.96), where we observe a CCHI score of 890, p = .0528. Again, we observe a reduction in CCHI of 97.75% (*n* = 870) with the application of treatment, which is close to, but not at, the level of statistical significance.

Secondary Research Question 3 — Does the delivery of a nudge text intervention that combines moral persuasion with deterrence reduce stalker re-offending and Cambridge Crime Harm Index score compared to no intervention, comparing suspects' <u>age categories</u>?

Prevalence of Re-offending:

Suspects in the treatment group aged 18-29 (*n* = 135, *M* = 0.04, *SD* = 0.19) re-offended at a rate of 3.7% (*n* = 5) compared with the control (*n* = 232, *M* = 0.04, *SD* = 0.19) who re-offended at a rate of 3.9% (*n* = 9), p = .9323. Suspects aged 30-39 in the treatment (*n* = 165, *M* = 0.05, *SD* = 0.23) re-offended at a rate of 5.5% (*n* = 9) compared with control (*n* = 318, *M* = 0.03, *SD* = 0.17) who re-offended at a rate of 3.1% (*n* = 10), p = .2550. Suspects aged 40 and above in the treatment group (*n* = 224, *M* = 0.04, *SD* = 0.19) re-offended at a rate of 3.6% (*n* = 8) compared with control (*n* = 272, *M* = 0.04, *SD* = 0.19) who re-offended at a rate of 3.7% (*n* = 10), p = .9504. Overall, we observe a marginal crime reduction effect following treatment application for suspects 18-29 years old or 40 or more years old and a small backfire effect for suspects aged 30-39 years old. No result is at the level of statistical significance.
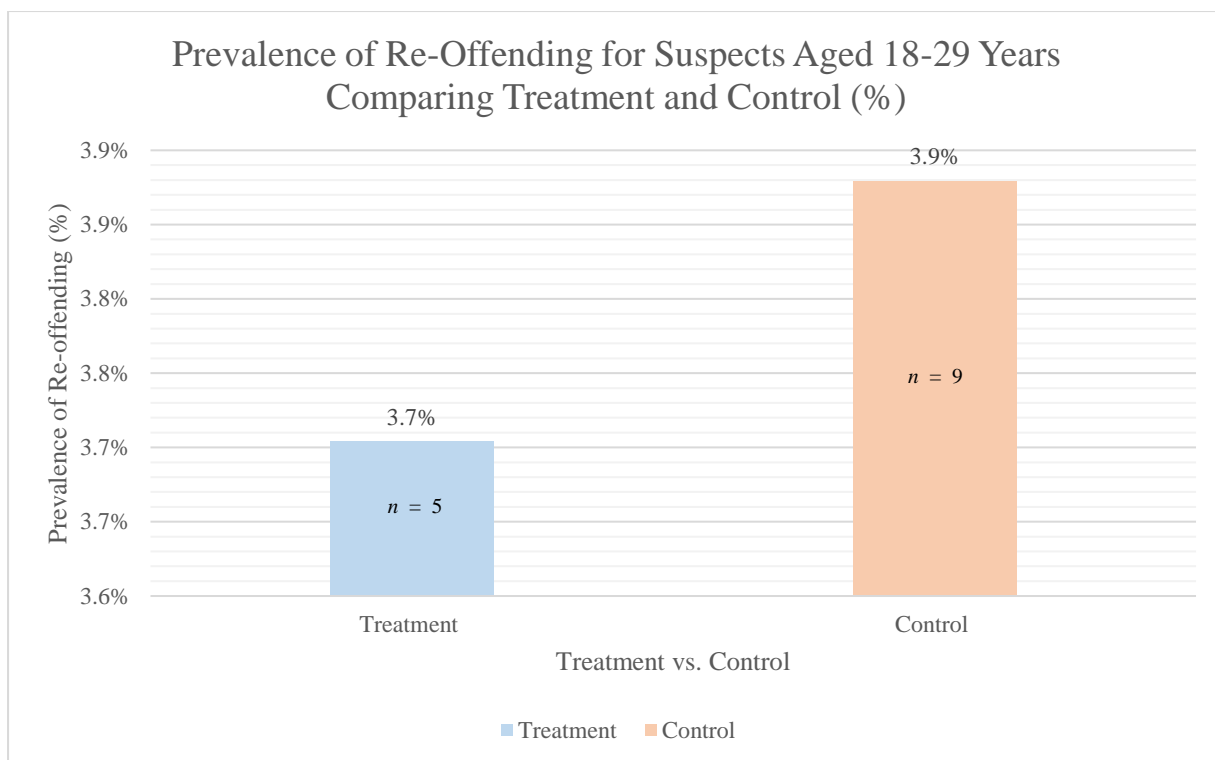
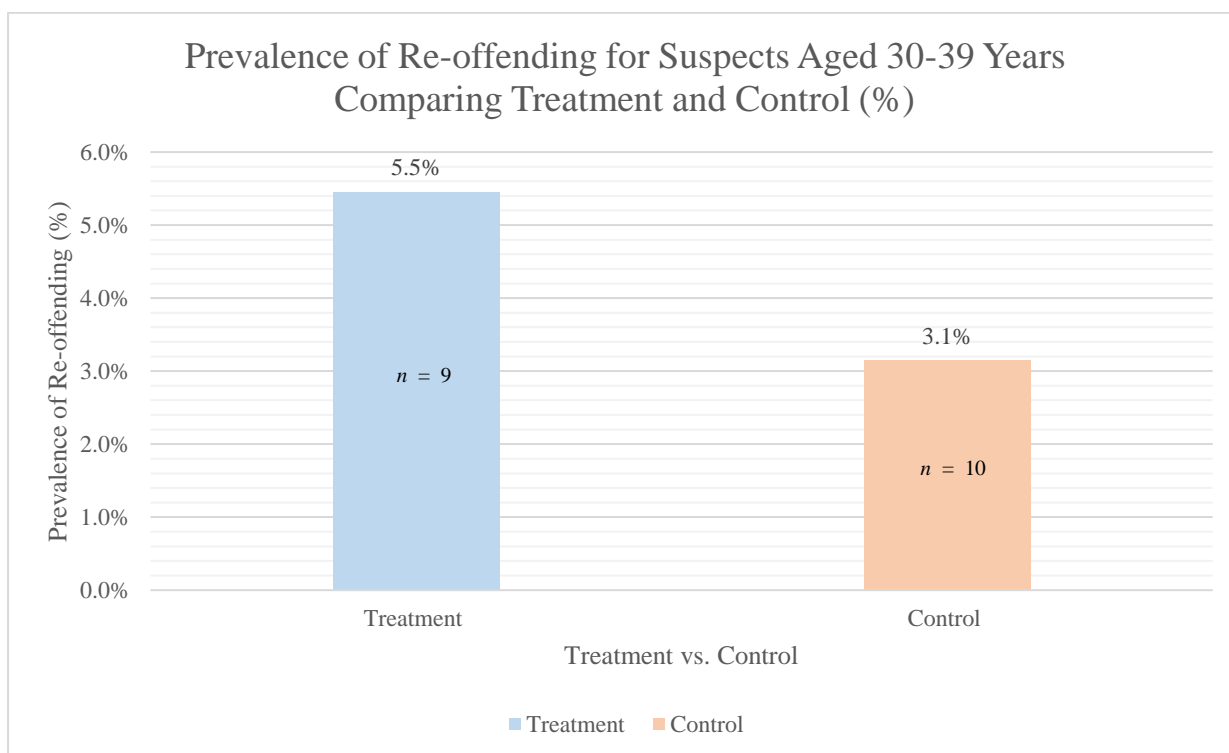**Figure 14 – Prevalence of re-offending for suspects aged 18-29 years, comparing treatment and control**

Figure 15 – Prevalence of re-offending for suspects aged 30-39 years, comparing treatment and control
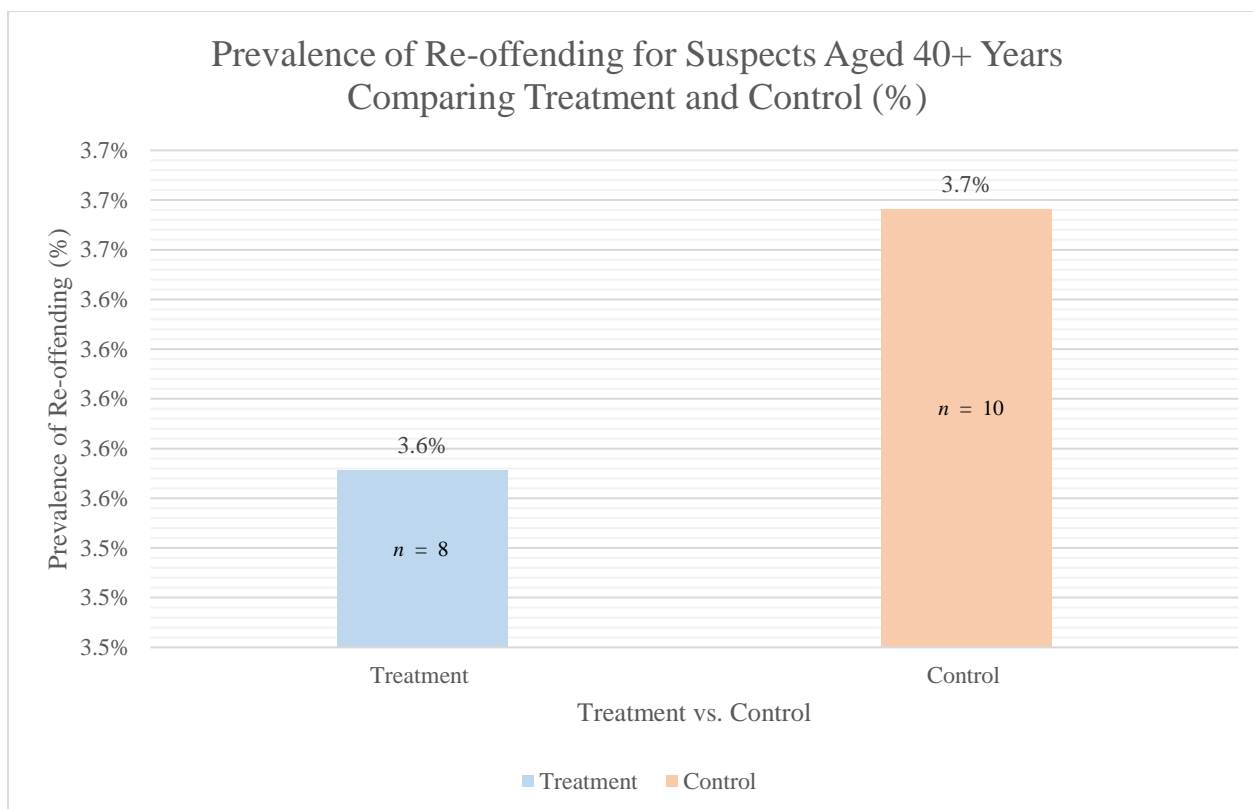
Figure 16 – Prevalence of re-offending for suspects aged 40+ years, comparing treatment and control

Cambridge Crime Harm Index Score:

For suspects aged 18-29 years old in the treatment group (n = 135, M = 2.71, SD = 22.83) we observe a CCHI score of 366 compared with control (n = 232, M = 4.50, SD = 29.97) where we observe a CCHI score of 1,043, p = .5215. This represents a crime harm reduction of 64.91% (n = 677) not at the level of statistical significance. For suspects aged 30-39 years old in the treatment group (n = 165, M = 3.99, SD = 24.69) we observe a CCHI score of 658 compared with control (n = 318, M = 3.37, SD = 28.28) where we observe a score of 1,073, p = .8056. We therefore observe a crime harm reduction of 38.68% (n = 415) not at the level of statistical significance. For suspects aged 40 or above in the treatment group (n = 224, M = 1.42, SD = 16.90) we observe a CCHI score of 317 compared with control (n = 272, M = 1.57, SD = 16.35) where we observe a CCHI score of 426, p = .9200. This represents a reduction in crime harm of 25.59% (n = 109) not at the level of statistical significance.

Secondary Research Question 4 – Does the delivery of a nudge text intervention that combines moral persuasion with deterrence cause a reduction in stalker re-offending and Cambridge Crime Harm Index score compared to no intervention, comparing whether an offence was classified as domestic abuse or non-domestic abuse?

Prevalence of Re-offending:

Domestic abuse suspects in the treatment group (n = 410, M = 0.05, SD = 0.22) re-offended at a rate of 4.9% (n = 20) compared with control (n = 633, M = 0.04, SD = 0.19) who re-offended at a rate of 3.8% (n = 24), p = .4064. Non-domestic abuse suspects in treatment (n = 114, M = 0.02, SD = 0.13) re-offended at a rate of 1.8% (n = 2) compared with control (n = 189, M = 0.03, SD = 0.16) who re-offended at a rate of 2.6% (n = 5), p = .6005.
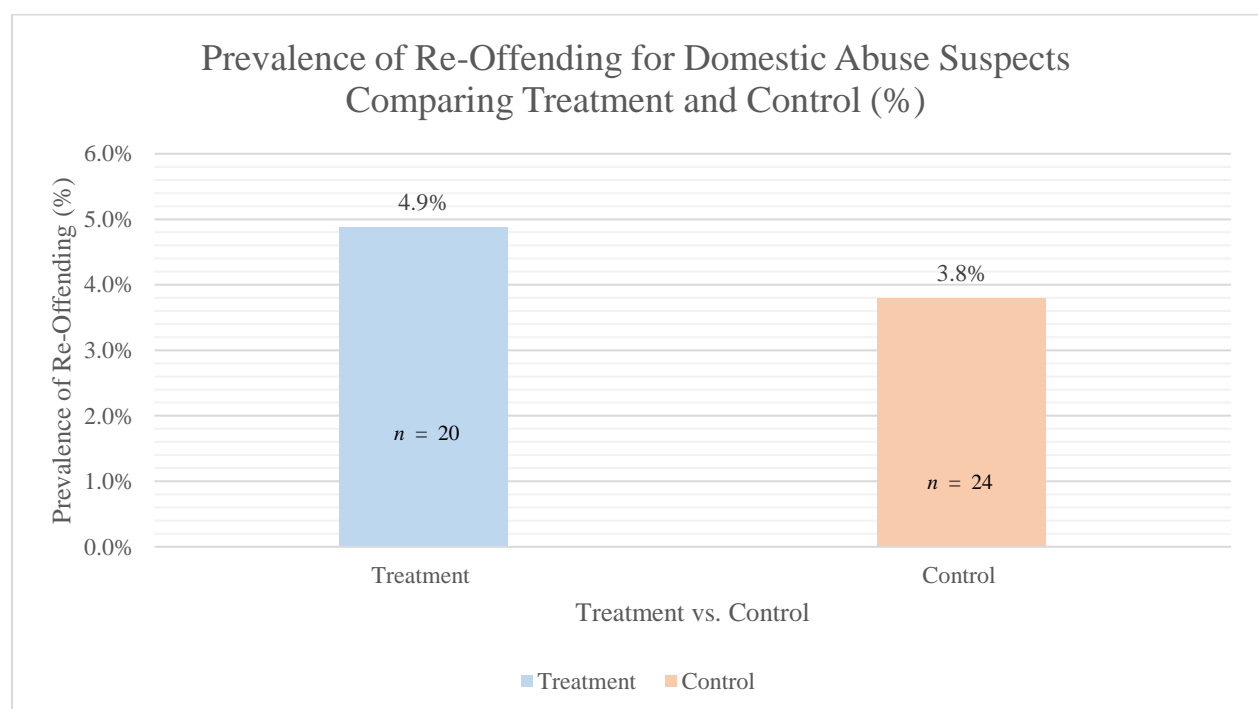

Figure 17 – Showing the prevalence of re-offending for domestic abuse stalking suspects comparing treatment and control groups
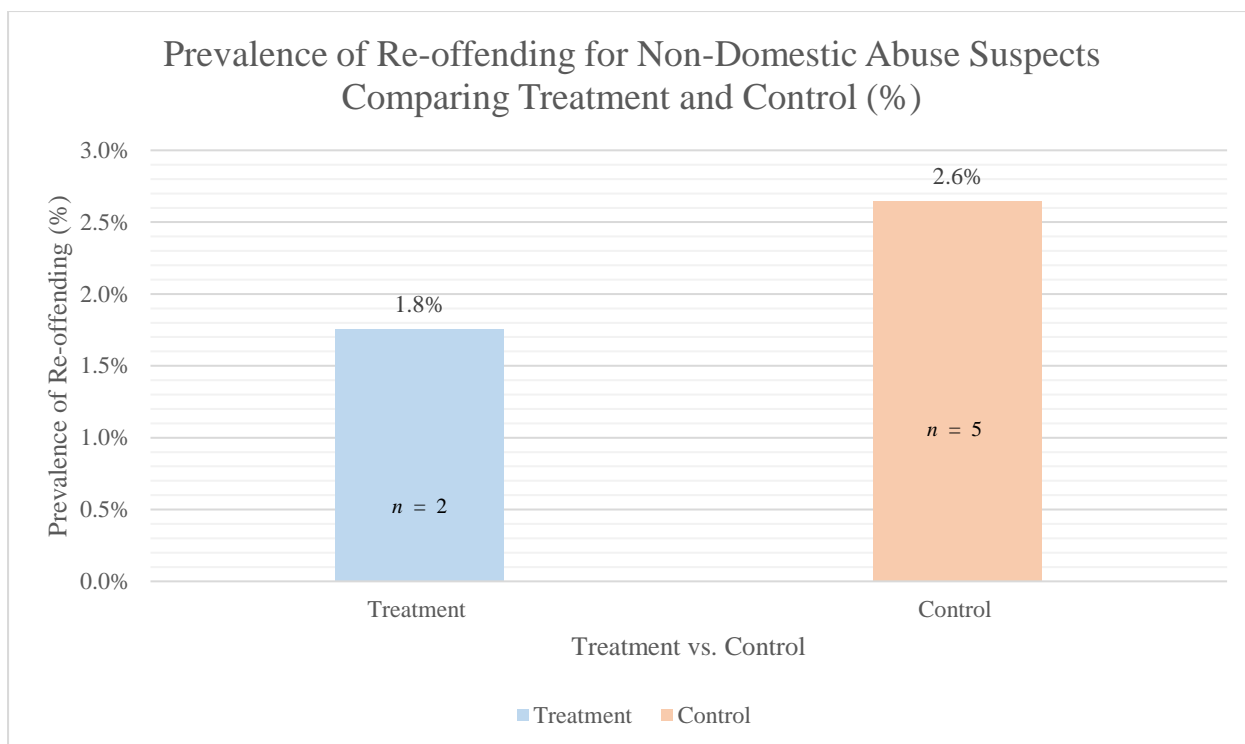
Figure 18 – Showing the prevalence of re-offending for non-domestic abuse stalking suspects comparing treatment and control groups

Cambridge Crime Harm Index Score:

Domestic abuse suspects in treatment (*n* = 410, *M* = 3.04, *SD* = 23.54) returned a CCHI score of 1,247 compared with control (*n* = 633, *M* = 3.16, *SD* = 25.35) who returned a CCHI score of 2,003, *p* = .9374. This represents a reduction of 37.74% (*n* = 756) points not at the level of statistical significance. For non-domestic suspects in treatment (*n* = 114, *M* = 0.82, *SD* = 7.91), we observe a CCHI score of 94 compared with control (n = 189, M = 2.85, SD = 26.13), where we observe a CCHI score of 539, *p* = .3211. This represents a reduction in crime harm of 82.56% (*n* = 445), not at the level of statistical significance.

Secondary Research Question 5 - Does the delivery of a nudge text intervention that combines moral persuasion with deterrence cause a reduction in stalker re-offending and Cambridge Crime Harm Index score compared to no intervention, <u>comparing the time lapse between the commission</u>

of a first stalking offence and the delivery of the nudge text intervention across male and female suspect sex?

Prevalence of Re-offending:

Male suspects whose first offence was up to 179 days prior to intervention in the treatment group (n = 32, M = 0.03, SD = 0.18) re-offended at a rate of 3.1% (n = 1) compared with control (n = 26, M = 0.08, SD = 0.27) who re-offended at a rate of 7.7% (n = 2), p = .4634. Female suspects whose first offence was up to 179 days prior to intervention in the treatment group (n = 11, M = 0.00, SD = 0.00) re-offended at a rate of 0.0% (n = 0) compared with control (n = 22, M = 0.14, SD = 0.35) who re-offended at a rate of 13.6% (n = 3), p = .1055. Male suspects whose first offence was 180 days or more prior to intervention in the treatment group (n = 329, M = 0.04, SD = 0.20) re-offended at a rate of 4.3% (n = 14) compared with control (n = 541, M = 0.03, SD = 0.16) who re-offended at a rate of 2.6% (n = 14), p = .2026. Female suspects whose first offence was 180 days or more before intervention in treatment (n = 70, M = 0.04, SD = 0.20) re-offended at a rate of 2.9% (n = 2) compared with control (n = 106, M = 0.06, SD = 0.23) where suspects re-offended at a rate of 5.7% (n = 6), p = .6795.
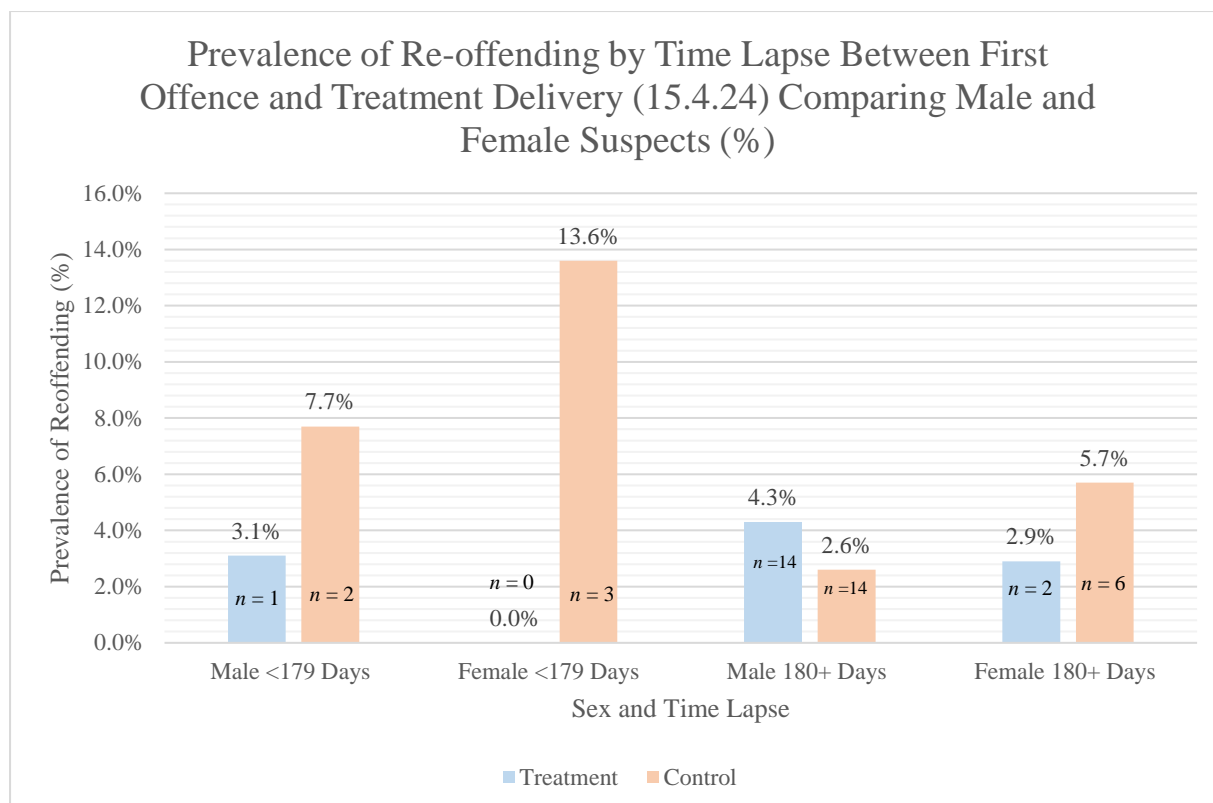
Figure 19 – Showing prevalence of re-offending (%) by time lapse between first offence and treatment delivery on 15.4.24, comparing male and female suspects

Cambridge Crime Harm Index Score:

For male suspects whose first offence was up to 179 days before intervention in the treatment group (*n* = 32, *M* = 0.31, *SD* = 1.77), a CCHI score of 10 is observed compared with control (*n* = 26, *M* = 10.27, *SD* = 49.39) where a CCHI score of 267 is observed, p = .3137. This represents a reduction in crime harm of 96.25% (*n* = 257), not at the level of statistical significance. For female suspects whose first offence was up to 179 days before intervention in the treatment group (*n* = 11, *M* = .00, *SD* = .00), we observe a CCHI score of 0 compared with control (*n* = 22, *M* = 5.18, *SD* = 20.05) where we observe a CCHI score of 114, p = .2009. We, therefore, observe a reduction in crime harm of 100% (*n* = 114). For male suspects whose first offence was 180 days or more prior to intervention in the treatment group (*n* = 329, *M* = 3.11, *SD* = 22.78), we observe a CCHI score of 1,024 compared with control (*n* = 541, *M* = 1.88, *SD* = 19.54) where we observe a CCHI score of 1,019, p = .4163. We, therefore, observe a marginal increase in crime harm of 0.49% (*n* = 5), not at the level of statistical significance. For female

suspects whose first offence was 180 days or more prior to intervention in the treatment group (n = 70, M = 0.43, SD = 2.04), we observe a CCHI score of 30 compared with control (n = 106, M = 7.42, SD = 41.97) where we observe a CCHI score of 786, p = .0897. We, therefore, observe a reduction in crime harm of 96.18% (n = 756), not at the level of statistical significance.

Summary

Firstly, no result was statistically significant and suitable to extrapolate beyond the trial's sample. Secondly, harm reduction is observed in all categories other than males whose offence was 180 days or more before intervention. Overall, a small prevalence backfire effect is observed when treatment is delivered.

If we accept that crime harm reduction is an almost universal benefit of this intervention, when does this combine with prevalence reductions? We observe this most noticeably for female stalkers generally and those stalkers with male victims (many of whom are female stalkers). Moreover, we observe prevalence reductions for both sexes of stalkers when an intervention is delivered within approximately a six-month period after an offence occurs.

If we revisit the point set out in the methods chapter that 28.9% of stalkers re-offend within the first month and further develop this by allowing two weeks to elapse prior to an intervention, we identify a two-week window where the intervention may most successfully prevent re-offending. Moreover, the intervention may be delivered at any point up to six months, with prevalence and harm reductions probable. We return to these results in the discussion chapter.

Chapter IV

Discussion

Overview

In September 2024, one month after the post-treatment period for this trial concluded, His Majesty's Inspectorate of Constabulary and Fire and Rescue Services (HMICFRS) published their findings following an investigation into the National Stalking Consortium's super-complaint about the police response to stalking in England and Wales. The report exhorts policing to take "appropriate action to stop offending" (HMICFRS, 2024, p.5). The imperative for law enforcement to understand which actions could "stop offending", the dearth of robust scientific evidence to offer solutions and the overwhelming complexity of stalker behaviour and motivations set out in this study come together in a confluence of challenges for modern policing.

Any study seeking to influence the re-offending of stalkers must grapple with what effective behavioural change looks like for stalkers. Is the locus of their fixated behaviour internalised or external? Do they wish to satisfy their compulsion for contact, in which the victim is collateral damage, or do they wish to inflict harm on their victims actively? Are stalkers empathetic and capable of understanding the harm they have caused? This study has shown that policing can begin to turn the dial on re-offending rates and that, in some unique contexts, a blended moral persuasive and deterrent message may influence behaviours. However, the absence of statistically significant results means further research is required to understand this picture fully.

Taking this view, this chapter will begin by assessing the theoretical implications of the findings, returning to the theoretical context in which this study is situated as set out in the 'Literature Review' chapter. It will then set out limitations in study design and considerations for how future research

could be framed to best effect. The chapter will proceed to consider the external validity of the findings, looking specifically at age, ethnicity and sex. Finally, the chapter will approach this trial's findings from the perspective of operational policing and criminal justice policymaking and assess the implications for the response of that sector to stalking perpetration.

Summary of Findings

If nudge text interventions are sent to suspects of both sexes before 180 days (approximately six months), crime prevalence and harm reduction benefits are observed. The same remains true for female stalking suspects regardless of when the intervention is delivered. Furthermore, 28.9% of stalkers who will re-offend do so within the first 29 days after their first offence, demanding early intervention to reduce the risk of re-offending. This would indicate that a policy where a nudge text intervention is delivered to stalking suspects, ideally within 29 days but no later than 6 months after their first offence, may yield benefits.

Theoretical Implications

The construction of the nudge text message used in this trial was based on deterrence theory and existing studies on using moral persuasion to influence offending behaviour. In the case of deterrence theory, Durlauf and Nagin (2011) called for law enforcement to move to a model explaining the consequences of offending to offenders. Wheatley et al. demonstrated that some stalkers may not be aware of the legal implications of their behaviour. Finally, Braga et al.'s study of the 'Operation Ceasefire' experiment in the 1990's recommended direct communication with the offending population and an offer of support to yield crime reduction benefits (Braga et al. 2018). These elements have informed this construction of the nudge text intervention and as such this trial, in part, is an experiment of these recommendations in a real-world setting. Its findings are directly relevant to advancing knowledge about whether informing offenders of the legal consequences of their behaviour and signposting support can create a crime reduction benefit.

A missing feature from the nudge text construction was personalisation. Braga and Weisburd's systematic review of deterrence highlighted the need for personalised deterrence messaging (Braga and Weisburd, 2018). Likewise, Chivers and Barnes' 2018 study of nudge text interventions to increase the court attendance rate of defendants extolled the importance of personalised text messages (Chivers and Barnes, 2018). Although the nudge text intervention used in this study is highly emotive and engineered to prompt personal reflection on offending behaviour, this trial is nonetheless an experiment of sending non-personalised text messages to offenders.

In addition to being a trial which explores the impact of deterrence on stalking suspects, this is also more broadly a trial of the impact of deterrence in general terms. This study has been unable to report on the extent to which trial participants were either known to law enforcement through their stalking alone or known for other criminality, for example, robbery, drug supply, violent crime or sexual violence. It is plausible that some of the participants would have far broader offending and that a direct text message from the police could have influenced that behaviour. The limitations explored elsewhere in this study arising from manually reviewing and recording follow-up results from police indices have prevented a broader review of offending behaviour. However, future study designs should consider retrieving this broader re-offending dataset to report on the impact.

That the behaviour of female stalkers can be influenced so dramatically by a nudge text intervention is arguably the key finding of this trial. Female stalking suspects accounted for 15.6% of suspects ($n = 256$) in the sample, with 62% ($n = 161$) alleged to stalk male victims and 38% ($n = 95$) female victims. Brooks et al.'s 2021 work on female-perpetrated stalking highlights that this area is significantly under-researched and that female stalkers present similar levels of risk to male stalkers but engage in nuanced behaviours (Brooks, Petherick, Kannan, Stapleton and Davidson, 2021). Meloy et al.'s 2011 study, 'The Female Stalker', which looked at 143 female stalkers, found behaviours which were less violent and threatening than male stalkers, with a greater prevalence of violence and threats present for prior-intimate partners (Meloy, Mohandie and Green, 2011). Moreover, the study observed

recidivism rates at 50% (*Ibid*). It is possible that compliance with the law is higher for women than men and that this holds for stalking suspects. Torgler and Valev's 2006 study, 'Women and Illegal Activities' used World Values Survey and European Values Survey data to test the difference between male and female attitudes to corruption and tax evasion, finding, "Women are significantly less likely to agree that corruption and cheating on taxes can be justified" (Torgler and Valev, 2006, p.22).

However, this study has reinforced the prevailing view that stalking is a gendered crime and one where male perpetrators of stalking against female victims predominate, comprising 71.2% (n = 1167) of cases sent for random assignment. Male victims of both female and male suspects were observed in the sample, at 9.8% (n = 161) and 4.5% (n = 73) respectively. We must approach data on male victimisation rates with caution due to the issue of under-reporting. Data did show a tendency for fewer male suspect-female victim relationships in dyads in non-domestic abuse cases, with male suspect-female victims comprising 76.3% (n = 358) for domestic abuse and 69.1% (n = 809) for non-domestic abuse.

Nonetheless, the data shows us that other women may predate on male or female victims, the latter at lower rates. These insights may prove helpful in commissioning victim support services and targeting educational, crime prevention campaigns and other public messaging. The nuance of female-perpetrated stalking set out in this chapter calls for stalking to hold a standalone strategic focus, accepting significant overlaps and inter-dependencies with crime disruption efforts in the domestic abuse and violence against women and girls arenas. Failing to provide stalking a standalone focus leaves a significant area of risk unresolved.

Research Implications

Only 64.1% (n = 524) received the intervention from a cohort of 817, demonstrating the inaccuracy of mobile phone numbers on police indices. An obvious policy implication is to confirm the accuracy of phone numbers with suspects when interacting with police. Improved treatment fidelity might have been achieved by sending a generic 'test' message to all participants in treatment and control to assess

mobile number accuracy. Doing so would not correlate with a real-world implementation of this approach, where confirmatory pre-messages would not be sent, could raise suspicion in suspects without any benefits and arguably diminish the effectiveness of the subsequent nudge text.

The use of stalking suspects, as opposed to those charged or convicted with stalking, requires reflection. Not everyone alleged to be a stalker in an initial report will have stalked. However, not everyone for whom no further action is taken before charge, nor those acquitted at court when charged, will *not* have engaged in stalking behaviours. Furthermore, as set out in the preceding chapter, most stalkers will re-offend shortly after their first offence. Progressing to charge, judicial disposal and sentence may take months to years, during which most stalkers have already re-offended, and the opportunity to intervene has passed. Adopting "suspect" as opposed to those charged (commonly referred to as an "accused" or a "defendant" in the criminal justice system) requires an acceptance of the above issues, but it is nonetheless the only viable route this study identified to reach the stalking population at greatest risk of re-offending.

A further reflection for this study has been the need for personalisation, which is extant in the research literature reviewed in Chapter I. The sample size for this study's RCT, n = 857 in the treatment group, and the mass transit software used to send the SMS did not allow personalisation. However, it is worth considering here why personalisation may yield improved treatment. Braga and Weisburd's 2018 systematic review of deterrence references that deterrence advertises "law enforcement strategy and the personalised nature of its application" (Braga & Weisburd, 2018, p.210). At the core of deterrence theory, it speaks directly to a prospective offender personally – 'What are the implications for me?' They further reference the procedural justice element of deterrence strategies, namely that how deterrence lands with an individual – favourably or unfavourably – may influence their willingness to engage (*Ibid*). Cumberbatch and Barnes' 2018 study of nudge texts to influence victim and witness court attendance attributes significant behavioural change to the personalisation of the text message (Cumberbatch & Barnes, 2018).

Earlier experiments confirm the need for personalisation: "The evidence supports a clear hypothesis that personalised SMS reminders… increase the attendance rates of defendants" (Chivers & Barnes, 2018, p.7). Personalisation goes beyond the name. Haynes, Green, Gallagher, John and Torgeson's 2013 study of nudge texts to influence delinquent fines outlines the offender's name and the specific satisfactory outstanding (Haynes et al., 2013). Compared to a baseline fine repayment of £4.46 in the control group, the experimental group who were named saw an £8.41 payment increase, and the experimental group whose specific fine was highlighted saw a rise of £6.07. When the name and specific amount were combined, the study saw an increased fine payment of £7.28 (*Ibid*). Accepting, then, that personalisation was outside the scope of this study, it is clear future study design must incorporate sufficient resource to personalise messages to create greater effect sizes.

A further reflection for future research design relates to the wording of the intervention. Objection could be raised to the suitability of the phrase "recent stalking allegation" as it could prompt a view that a new or "more recent" allegation of stalking has been made than that of which the offender was aware. Since the nudge intervention was sent, reflection suggests that a better compromise may have been to write words to the effect of, "You are receiving this text message because, over the last two years to March 2024, an allegation of stalking was made against you". Though seemingly a small point, the need for precise and clear wording in these messages speaks to the unique way stalkers fixate on words, situations and may react in a way inconsistent with society at large. Put simply, ambiguity should be avoided, and for future study design the wording should be reviewed by both a behavioural scientist and a clinical psychologist working with stalkers.

Future study design should also consider the population such an experiment is capable of reaching. The application of exclusion criteria, which included the removal of suspects with invalid or no mobile phone numbers, resulted in 62.3% (n = 2,770) of suspects being eliminated prior to randomisation. In addition, the failure rate of text message delivery at 35.9% (n = 293) further limited the sample of available participants. Should police services wish to communicate with an offender population

through SMS, their application of suitability criteria, the lack of mobile phone numbers, and the ability of remaining numbers to receive a text could result in them only reaching 31.1% (n = 1,380) of the target population. The need to improve the accuracy of recording suspect phone numbers and supplement any suspect contact strategy with methods other than SMS (e.g. emails, home visits) is clear.

Future research design should also consider the viability of block randomisation, with adjusted treatments across several groups. In this study we have been concerned with a blended deterrent and moral persuasive message. This design, whilst an important first step, cannot identify which of those elements had the greatest effect. A valid future design would be a treatment group with a blended message, a second with a deterrent message, a third with a moral persuasive message compared with a control group. Further block randomisation could be considered, depending on available sample size – for example, we could consider constructing different messages for male and female stalking suspects, given the results reported in this trial. Of course, we must return here to the sample size required to detect an effect. Clearly, were this study repeated in London, this study has shown there would be an insufficient sample size to support this design.

This issue of sample size is a key consideration for future study design. The test set out in Fig. 3 demonstrates that for a sample of 1,673, with a power of 80% and .05 alpha level, detection of statistically-significant differences could be expected at d = 0.137. The critical question is whether there were no statistically-significant differences in the results of this study, and therefore we have an accurate assessment of the transportability of these findings or whether – given the low rate of reported re-offending in the stalking population, we simply required a larger sample. London's stalking population, however, is broadly equal to that of the remainder of England and Wales. As such, a national study may only serve to double the sample size. Furthermore, researchers would be required to contend with the urban-rural disparity explored in the 'Methods' chapter. If further RCT's into

stalking interventions are to proceed, researchers must contend with a small population size and build the evidence base of the sample size required to detect statistically-significant effects in this area.

Developing this point, future studies should also consider why this research only came close to detecting statistically-significant effect sizes for crime harm reduction and not prevalence. Arguably, the numbers of suspects likely to re-offend was too small to expect such effect sizes, as explored above. However, the changes in crime harm were greater. Future studies must also consider the suitability of the CCHI to assessing stalking harm. When one considers that the 'Serious Alarm or Distress' offence carries a greater CCHI score than 'Fear of Violence', it is reasonable – accepting this reflects the seriousness ascribed by courts in sentencing – to question whether this accurately reflects the real harm. Is a victim of stalking in fear that violence will be used not also suffering serious alarm or distress? Looking at the score of '5' assigned to Breach of a Stalking Protection Order, if a victim has been sufficiently concerned to support a Stalking Protection Order application to protect themselves, are they not likewise suffering serious alarm or distress? A key reflection is whether sentencing, and consequent CCHI scores, reflect the actual crime harm victims suffer. Returning to the sheer complexity of this offence explored in the literature review, we may further question whether stalking harm is adequately captured in a numeric value. More complex assessments of stalking risk are available and used in the sector. A common theme in this study has been the benefit to be gained from co-designing and executing a study with a clinical psychologist experienced in working with stalkers. Doing so in future studies may yield a more accurate evaluation of harm.

<u>Policy Implications</u>

As referenced elsewhere in this study, texting a stalking suspect the day after an allegation is made might prompt risk escalation before police can effectively intervene. If a system of delivering nudge texts were to be automated, specific controls would be required – for example, a 'buffer zone' of two weeks or an automated text once a suspect is released on bail, charged or no further action is taken.

Furthermore, this trial has identified the age curve for stalking offences in London, with suspects falling primarily between the ages of 20-44. Early adulthood into middle age, therefore, appears to be a peak time when stalking behaviours are demonstrated in the offending population. The mean average age difference between stalking suspects and their victims in the treatment group was only 3 years, with a range of 46 years younger to 46 years older. As seen in Fig. 20, suspects in their twenties, thirties and forties are more likely to stalk victims broadly the same age. Therefore, for public education campaigns seeking to raise awareness of stalking amongst current and future victims, the age demographic 20-44 years old should be a focus, and media best-suited to reaching those audiences preferred. 2021 Census data for England and Wales shows that the average age for men and women to enter the workforce is 23 years old, and the average retirement age is 66 (Office for National Statistics, 2024). The ages where stalking perpetration and victimisation predominate, therefore, correlate with the ages at which adults in England and Wales are likely to be in employment. Policymakers may, therefore, consider whether clear guidance and advice are provided to employers on stalking and harassment policies to protect their workforce, as well as using workplaces as an opportunity to deliver preventative messaging to the most vulnerable subset of the population. Likewise public messaging campaigns in public spaces timed to correlate with when most workers may be travelling to and from the workplace may yield benefits, as well as potentially reaching a younger school age audience at the same time.
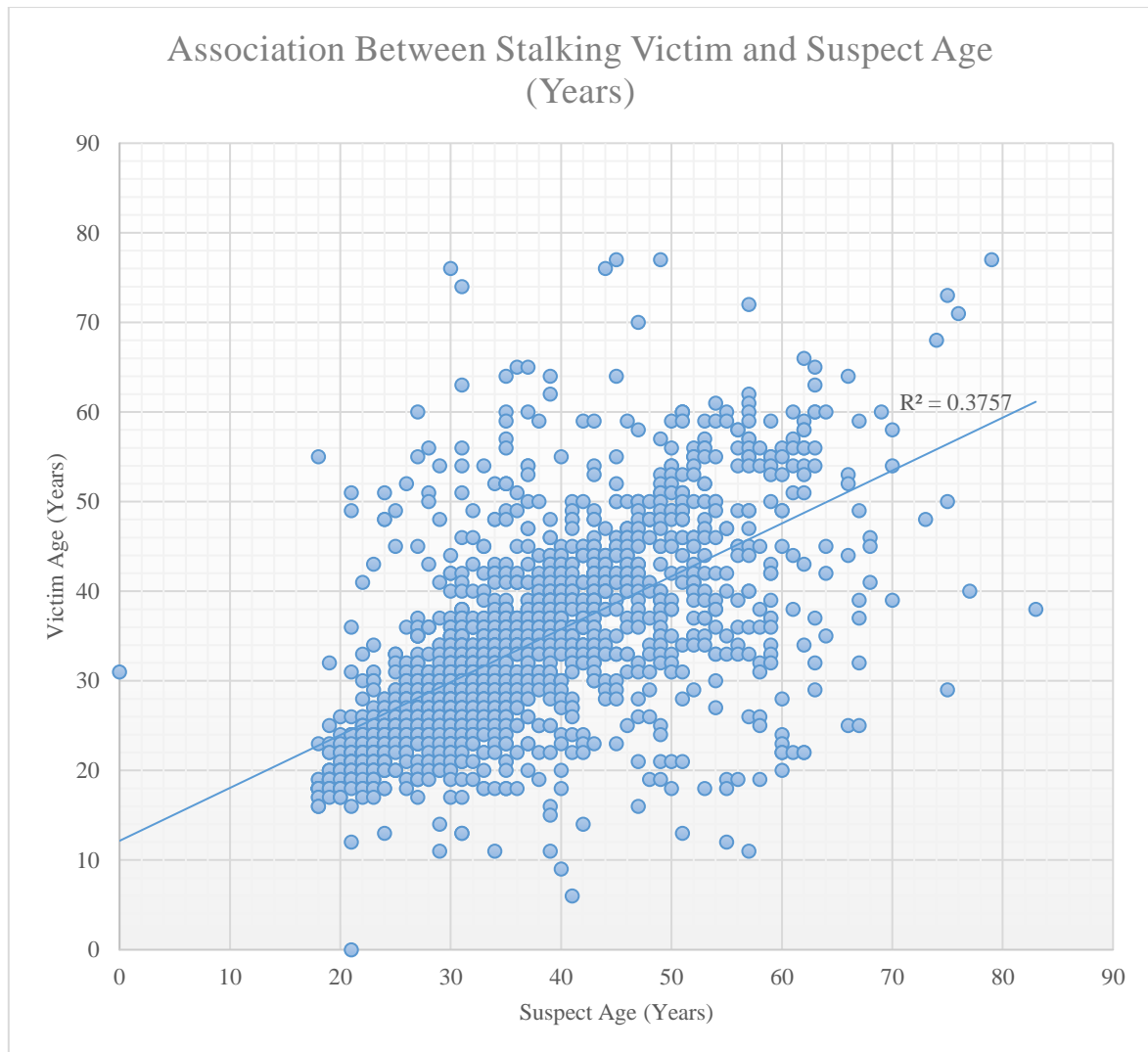
Figure 20 – Showing the correlation between victim and suspect age in years in the sample sent for randomisation

One may expect to observe a far higher re-offending rate in this cohort. In reality, between April 2022 and April 2024, 96.1% (*n* = 4,270) committed only one offence. Most stalking caseload volume reported in the Metropolitan Police Service area in this study's period, 96.1%, was still attributable to individual stalking suspects. From a policy perspective, therefore, whilst law enforcement may infer from these findings that there is a small cohort worthy of focused intervention, disrupting stalking overall requires a volume crime approach, tackling each case and stalking dyad as they are reported. Furthermore, returning to the subject matter discussed in the literature Review chapter, we know that re-offending is not predictive of harm escalation or homicide. Harm prediction lies in detail – in the

presence of violent threats, exploration of a suspect's adverse childhood experiences, alcohol or substance abuse. Identifying the risk of future harm and escalation and accepting it is vulnerable to human error and inconsistency lies in detailed individual risk assessments by reporting and investigating officers.

Further Limitations

Churcher and Nesca's (2013) study of stalking perpetration identified a statistically significant effect for prior intimate relationships in stalking and the progression to violence and a moderate effect size for the presence of verbal threats and escalation to violence. It should be noted that the methodology employed in this trial was not constructed to detect these effects. Firstly, re-offending was measured only as far as a suspect committed a further stalking offence. The absence of any flagging system on crime reports to indicate the presence of verbal threats and the fact that further offending under any violent crime category (common assault, actual bodily harm, grievous bodily harm, etc.) were not included in the follow-up analysis are limitations in this study's findings.

Returning to Pathe and Mullen's (1997) work setting out the typologies of stalkers as 'Rejected', 'Resentful', 'Predatory', 'Intimacy Seekers' and 'Incompetent Suitors' (Pathe and Mullen, 1997), it is worthy of discussion that this study could not isolate those cases in the findings, nor compare their occurrence across the treatment and control groups. However, it would be helpful to develop a theory to understand how stalkers falling into each of these commonly-used categories responded to a nudge text intervention. A more extensive study, for example, would draw participants from multiple force areas or even nationally, thereby obtaining a more comprehensive data sample that may benefit from allocating eligible participants into these categories and then block-randomising with tailored interventions.

Summary

The prevalence and harm reduction effects pointed out in this study, when weighed against the tremendous harms caused by stalking, are an important step forward in the research on stalking. Further research is required to understand the impact of nudge text interventions on preventing the escalation of stalking to the threat and commission of violence. Although this study has provided findings that female stalking suspects' behaviour is more readily affected by deterrent and moral persuasion messaging, further research is required to understand why. Randomised controlled trials utilising a more extended follow-up period – this study has proposed two years – may be better placed to detect any residual deterrence effect. The theory of using moral persuasion to impact offending behaviour has been further developed to show that such messaging can create a crime reduction benefit in stalking cases if delivered within 180 days. Further research is required to understand how the personalisation of nudge interventions may impact this trial's findings.

In addition, the chapter has set out a number of study design limitations and recommendations for future trials. It has explored what the policy implications of the findings may be. Looking across all findings reported in this trial, however, two conclusions stand out: prevalence and harm may be reduced if interventions are delivered within six months of a first offence; and female stalkers' re-offending and harm may be reduced with this treatment irrespective of the time lapse between first offence and intervention.

Conclusion

This study has demonstrated that re-offending risks are at their highest only a month after a first offence occurs. The clear imperative is to develop robust evidence of quick, adaptable and effective interventions to minimise harm, and to ensure these can be delivered at scale. If that position is reached, policing can begin to offer behavioural change which goes beyond apprehension, investigation, and supporting prosecution. Whether another study will repeat the same results is not clear, given the non-significant results produced by this experiment. Moreover, nudge texts are one of many hitherto untested interventions. Evidence based policing should look broadly to build a suite of interventions through which the harms of stalking to society can be reduced.

This study provides three critical additions to the literature on this topic. Firstly, the findings indicate that a blanket approach to nudge interventions for stalking suspects does not work and may actually increase the prevalence of stalking, albeit whilst reducing harm. Therefore, criminal justice agencies and policymakers should avoid any blanket policy to deliver this intervention.

Secondly, a nudge intervention in this context is time-sensitive, which affirms the prevailing view that early stalking intervention is most effective. This study has indicated that 6 months may be an ideal backstop after which a suspect is no longer eligible for the intervention.

Thirdly, the study has demonstrated an unexpected finding, namely that female stalkers and stalkers who target male victims are more readily influenced than male stalkers or those targeting female victims. Beyond 6 months, for male stalkers only, we observe a backfiring effect. This study has pointed to research on female compliance with the law, which may go some way in explaining the results. However, it has also demonstrated that further research on female stalkers is required broadly across criminology and other associated fields.

All findings must, of course, be caveated with the fact that, whilst some results related to female stalkers and male victims came close to the level of statistical significance, none returned a p-value of <.05. Extrapolating the results beyond the sample studied is not possible, and further research is required. It is suggested in this study that the unexpectedly low prevalence of re-offending in stalkers, consistently <5%, means a more extensive case sample is required. The challenge is that this study was conducted in the Metropolitan Police District and therefore engaged with the largest number of stalking suspects nationally in any one force. If these low re-offending rates are not unique to this sample and not unique to London, then we may infer that a national study would be far better placed to confirm whether reductions at the level of statistical significance can be observed.

Regarding the primary research question, a nudge text intervention used in this study may increase prevalence but reduce crime harm. Moving on to secondary questions, the intervention may increase the prevalence of re-offending for males and reduce the prevalence for female suspects, but reduce crime harm in both. It may increase the prevalence of re-offending for suspects with female victims but decrease the same for those with male victims and decrease crime harm in both. It may reduce re-offending prevalence for suspects aged 18-29 or more than 40 years old, increase prevalence for those aged 30-39, and reduce crime harm in all categories. It may increase the prevalence of re-offending in domestic abuse and decrease the prevalence in non-domestic abuse cases but may reduce crime harm in all categories. Finally, it may reduce re-offending prevalence for female suspects, however old the allegation is, but for male suspects, prevalence reductions may only be observed for cases not older than 179 days – this study has offered 6 months as a useful measure of time.

In all cases except one crime harm will be reduced – male suspects whose alleged offence was 180 days old or older at the time of the intervention, where crime harm was observed to increase at 0.5%, not at the level of statistical significance. It is for this reason that the overall policy recommendation of this study is that a nudge text intervention is delivered to stalking suspects of any sex within 6 months of the date on which the victim alleges the stalking ended. Anything older than that may

increase prevalence and a small increase in crime harm levels for male suspects and their largely female victims.

It would be optimistic for any policing agency or policymaker to assume stalking could be eradicated as a phenomenon. It will continue to materialise in current and future generations. The aspiration of the criminal justice system must be, to whatever extent it is possible, to reduce the scale of stalking and – where it does happen – do everything possible to prevent recurrence. However, stalking is an extemporising of internal challenges for stalkers. Victim safety planning may yield only temporary respite if we cannot influence the stalker's perspective – like mopping up water in a flooded building instead of turning off the water supply. The stakes could not be higher and research must proceed cautiously but at pace. Such research should test interventions through randomised controlled trials to build reliable evidence. By doing so, the discipline can work towards a systematic review of the interventions that work to disrupt stalking offences. Such a systematic review, in the hands of policy makers responsible for tackling stalking, would add academic rigour and consistency to how society safeguards itself against one of its most prevalent and harmful crimes.

References

Almeida, S.R., Lourenço, J.S., Dessart, F.J. and Ciriolo, E. (2016) 'Insights from behavioural sciences to prevent and combat violence against women.' *JRC Science for Policy Report, EUR 28235EN. DOI*, *10*, p.412325.

Ariel, B. (2012). 'Deterrence and moral persuasion effects on corporate tax compliance: findings from a randomized controlled trial.' *Criminology*, 50(1), p.27-69.

Ariel, B., Bland, M. and Sutherland, A. (2022) 'Experimental designs' *Sage*.

Barry, B., Morales, L. and Carthy, A. (2022) 'An Evidence Review of Behavioural Economics in the Justice Sector.'

Berk, R. (2017) 'An impact assessment of machine learning risk forecasts on parole board decisions and recidivism.' *Journal of Experimental Criminology*, 13, p.193–216.

Berk, R.A., Sorenson, S.B. and Barnes, G. (2016) 'Forecasting domestic violence: A machine learning approach to help inform arraignment decisions.' *Journal of Empirical Legal Studies*, *13*(1), p.94-115.

Berk, R.A., Sorenson, S.B. and Barnes, G. (2016) 'Forecasting domestic violence: A machine learning approach to help inform arraignment decisions.' *Journal of Empirical Legal Studies*, *13*(1), p.94-115.

Brady, P.Q. and Nobles, M.R. (2017) 'The dark figure of stalking: Examining law enforcement response.' *Journal of interpersonal violence*, *32*(20), p.3149-3173.

Braga, A.A., Weisburd, D. and Turchan, B. (2018) 'Focused deterrence strategies and crime control: An updated systematic review and meta-analysis of the empirical evidence.' *Criminology & Public Policy*, *17*(1), p.205-250.

Brewster, M.P., 2000. 'Stalking by former intimates: Verbal threats and other predictors of physical violence.' *Violence and Victims*, *15*(1), p.41.

Brooks, N., Petherick, W., Kannan, A., Stapleton, P. and Davidson, S. (2021) 'Understanding female-perpetrated stalking.' *Journal of threat assessment and management*, *8*(3), p.65.

Chivers, B., & Barnes, G. (2018). 'Sorry, wrong number: Tracking court attendance targeting through testing a "nudge" text.' *Cambridge Journal of Evidence-Based Policing*, 2, p.4-34.

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)* p.100-112.

Cumberbatch, J.R. and Barnes, G.C. (2018) 'This nudge was not enough: a randomised trial of text message reminders of court dates to victims and witnesses.' *Cambridge Journal of Evidence-Based Policing*, *2*, p.35-51.

D'Ignazio, C., Val, H.S., Fumega, S., Suresh, H. and Cruxên, I. (2020) 'Feminicide & machine learning: detecting gender-based violence to strengthen civil sector activism.'

Data Magazine (2023), 'Understand the levels of measurement for age in data analysis'. Retrieved 1536 on 03 April 2024 from Understanding the Levels of Measurement for Age in Data Analysis – Data Magazine.

Durlauf, S.N. and Nagin, D.S. (2011) 'Imprisonment and crime: Can both be reduced?' *Criminology & Public Policy*, *10*(1), p.13-54.

Elklit, A., Vangsgaard, L.A.G., Olsen, A.S.W. and Ali, S.A. (2019) 'Post-traumatic stress disorder (PTSD) symptoms in secondary stalked children of Danish stalking survivors—a pilot study.' *International journal of environmental research and public health*, *16*(5), p.725.

Eterovic-Soric, B., Choo, K.K.R., Ashman, H. and Mubarak, S. (2017) 'Stalking the stalkers–detecting and deterring stalking behaviours using technology: A review.' *Computers & security*, *70*, p.278-289.

Findley, M.G., Kikuta, K. and Denly, M. (2021) 'External validity' *Annual Review of Political Science*, *24*, p.365-393.

GOV.UK (2022), 'Correspondence – Super-complaint on the police response to stalking'. Retrieved 2024, 12th September from Super-complaint on the police response to stalking - GOV.UK (www.gov.uk).

GOV.UK (2024), 'Regional Ethnic Diversity – 2021 Census Data'. Retrieved 2024, 12th September from Regional ethnic diversity - GOV.UK Ethnicity facts and figures (ethnicity-facts-figures.service.gov.uk)

Haynes, L.C, Green, D.P., Gallagher, R., John P., and Torgerson, D.J. (2013) 'Collection of delinquent fines: an adaptive randomized trial to assess the effectiveness of alternative text messages.' *Journal of Policy Analysis and Management*, p.718–730.

His Majesty's Inspectorate of Constabulary and Fire and Rescue Services (2024), 'The Police Response to Stalking'. Retrieved 2024, 28th September from The police response to stalking: Report on the super-complaint made by the Suzy Lamplugh Trust on behalf of National Stalking Consortium

HM Government (2022), 'Regional ethnic diversity'. Retrieved 2033, 17th May 2024 from Regional ethnic diversity - GOV.UK Ethnicity facts and figures (ethnicity-facts-figures.service.gov.uk).

Hopewell, S., Clarke, M., Moher, D., Wager, E., Middleton, P., Altman, D.G., Schulz, K.F. and Consort Group (2008) 'CONSORT for reporting randomized controlled trials in journal and conference abstracts: explanation and elaboration.' *PLoS medicine*, *5*(1), p.e20.

Kamphuis, J.H., Emmelkamp, P.M. and Bartak, A. (2003) 'Individual differences in post-traumatic stress following post-intimate stalking: Stalking severity and psychosocial variables.' *British Journal of Clinical Psychology*, *42*(2), p.145-156.

London City Hall (2024), 'London: 20 facts about London's culture'. Retrieved 2037, 17th May 2024 from 20 facts about London's culture | London City Hall.

Manchester City Council (2024), 'Census 2021 – Demographic Summary'. Retrieved 2024, 12th September from Census 2021 - Demography Summary | Census 2021 - Demography Summary | Manchester City Council

Marshall, G. and Jonker, L. (2010) 'An introduction to descriptive statistics: A review and practical guide.' *Radiography*, *16*(4), p.e1-e7.

Marshall, G. and Jonker, L. (2011) 'An introduction to inferential statistics: A review and practical guide.' *Radiography*, *17*(1), p.e1-e6.

Meloy, J.R., Mohandie, K. and Green, M. (2011) 'The female stalker.' *Behavioral Sciences & the Law*, *29*(2), p.240-254.

Michie, S., Van Stralen, M.M. and West, R., 2011. 'The behaviour change wheel: a new method for characterising and designing behaviour change interventions.' *Implementation science*, *6*, p.1-12.

Moher, D., Hopewell, S., Schulz, K.F., Montori, V., Gotzche, P.C., Deveraux, P.J., Elbourne, D., Egger, M. and Altman, D.G. (2010) 'CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials'. *BMJ, 340*.

Monckton-Smith, J. (2020) 'Intimate partner femicide: Using Foucauldian analysis to track an eight stage progression to homicide' *Violence against women, 26(11)*, p.1267-1285.

MPS Data and Insights (2024), 'Stalking & Harassment Dashboard'. Retrieved 1704, 28 March 2023 from Workbook: PD1256 Stalking & Harassment Dashboard.

Mullen, P., Pathe, M., and Purcell, R. (2009), 'Stalkers and their Victims', *Cambridge University Press*.

Mustaine, E.E. and Tewksbury, R. (1999) 'A routine activity theory explanation for women's stalking victimizations.' *Violence Against Women*, *5*(1), p.43-62.

Nagin, D.S., (2013) 'Deterrence in the twenty-first century.' *Crime and justice*, *42*(1), pp.199-263.

Office for National Statistics (2024), 'Demography and migration data, England and Wales: Census 2021'. Retrieved 2024, 12th September from Demography and migration data, England and Wales - Office for National Statistics (ons.gov.uk)

Office for National Statistics (2024), 'Milestones: journeying through a modern life'. Retrieved 2024, 12th September from Milestones: journeying through modern life - Office for National Statistics (ons.gov.uk)

Office for National Statistics (2024). '2021 Census Data – Sex'. Retrieved 2024, 12th September from Sex - Office for National Statistics (ons.gov.uk)

Paternoster, R. and Simpson, S. (1996) 'Sanction threats and appeals to morality: Testing a rational choice model of corporate crime.' *Law and Society Review*, p.549-583.

Pathe, M. and Mullen, P.E. (1997) 'The impact of stalkers on their victims.' *The British Journal of Psychiatry*, *170*(1), p.12-17.

Schulz, K.F., Altman, D.G., Moher, D. (2010) 'CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.' *BMC Med* 8, 18.

Statista (2024), 'Population of London in 2020, by age group'. Retrieved 20424, 12th September from

London population 2020 by age group | Statista

Statista (2024), 'Population of London in 2022, by borough'. Retrieved 2027, 17th May 2024 from

London population by borough 2022 | Statista.

Storey, J.E., Pina, A. & Williams, C. (2023). 'The impact of stalking and its predictors: Characterising the

needs of stalking victims.' *Journal of Interpersonal Violence*.

Towards Data Science (2024), 'Parametric Tests – the t-test'. Retrieved 1747 on 04 April 2024 from

Parametric Tests — the t-test. An intro to parametric tests and a deep… | by Shubhangi Hora

| Towards Data Science.

Travaini, G. V., Pacchioni, F., Bellumore, S., Bosia, M. & De Micco, F. (2022). 'Machine learning and

criminal justice: A systematic review of advanced methodology for recidivism risk prediction.'

Int. J. Environ. Res. Public Health 19, p.10594.

University of Cambridge (2024), 'The Cambridge Crime Harm Index (CCHI)'. Retrieved 1544, 15th July

2024 from The Cambridge Crime Harm Index (CCHI) | Institute of Criminology.

Verrey, J., Ariel, B., Harinam, V. and Dillon, L. (2023). 'Using machine learning to forecast domestic

homicide via police data and super learning.' *Scientific reports*, *13*(1), p.22932.

Wheatley, R., Henley, S. & Farnham, F. (2023). 'Stalking: Issues of deterrence "When I was stalking, I

was so dedicated to it. Nothing would stop me. It was my focus".' *Journal of Criminal

Psychology*, *13*(2), p.136–153.

Youngs, D., Ioannou, M. and Straszewicz, A. (2013). 'Distinguishing stalker modus operandi: An

exploration of the Mullen, Pathé, Purcell, and Stuart (1999) typology in a law-enforcement

sample.' *Journal of Forensic Psychiatry & Psychology*, *24*(3), p.319–336.

Appendices

Appendix A – Crim-PORT

Crim-PORT 1.0:

Criminological Protocol for Operating Randomized Trials

@ 2009 by Lawrence W. Sherman and Heather Strang

INSTRUCTIONS: Please use this form to enter information directly into the WORD document as the protocol for your registration on the Cambridge Criminology Registry of Experiments in Policing Strategy and Tactics (REX-POST) or the Registry of Experiments in Correctional Strategy and Tactics (REX-COST).

CONTENTS:

1. Name and Hypotheses

2. Organizational Framework

3. Unit of Analysis

4. Eligibility Criteria

5. Pipeline: Recruitment or Extraction of Cases

6. Timing

7. Random Assignment

8. Treatment and Comparison Elements

9. Measuring and Managing Treatments

10. Measuring Outcomes

11. Analysis Plan

12. Due Date and Dissemination Plan

1. Name and Hypotheses

A. **Name of Experiment**

Moral Choice Architecture: A randomised controlled trial of "Nudge" Interventions to Reduce Stalking Re-offending

B. Principal Investigator   (*n*ame) Daniel M. Thompson

                                (Employer) Metropolitan Police Service

C. 1st Co-Principal Investigator (*n*ame) Professor Barak Ariel

                                (Employer) University of Cambridge, Institute of Criminology

D. 2nd Co-Principal Investigator (*n*ame) Professor Yuval Feldman

                                (Employer) Bar Ilan University

E. <u>**General Hypothesis**</u>: Delivering a nudge intervention that blends moral persuasion and focused deterrence to an experimental group of stalkers reduces stalkers' re-offending and crime harm as defined by the Cambridge Crime Harm Index (CCHI) score compared to no intervention.

F. <u>**Specific Hypotheses:**</u>

1. List all variations of treatment delivery to be tested.
   A. all treatment cases will receive the exact nudge text

2. List all variations of outcome measures to be tested.
   A. Primary Outcome – the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI.
   B. Secondary Outcome 1 – the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI, comparing male and female suspects.
   C. Secondary Outcome 2 - the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI, comparing suspects with female victims to those with male victims.
   D. Secondary Outcome 3 - the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI, comparing suspects by age category.
   E. Secondary Outcome 4 - the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI, comparing domestic abuse to non-domestic abuse suspects.

F. Secondary Outcome 5 - the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI, comparing male to female suspects and the time lapse between their first offence and the nudge text intervention delivery.

3. List all subgroups to be tested for all varieties of outcome measures.
   A. Male and female suspects
   B. Suspects with male victims and suspects with female victims
   C. Suspects aged 18-29 years, 30-39 years and 40 years or older
   D. Suspects offending in a domestic abuse context compared with those in a non-domestic abuse context (domestic relationship defined as familial/intimate or ex-intimate partner context).
   E. Male suspects whose first offence in the sample was less than 180 days prior to intervention compared with those whose offence was 180 days or more prior. Female suspects whose first offence in the sample was less than 180 days prior to intervention compared with those whose offence was 180 days or more prior.

2. Organizational Framework: Check only one from a, b, c, or d

A. ~~**In-house** delivery of treatments, data collection and analysis~~

B. **Dual Partnership:** Operating agency delivers treatments with independent research organization providing random assignment, data collection, analysis__

   Name of Operating Agency **Metropolitan Police Service**

   Name of Research Organisation **University of Cambridge, Department of Criminology**

C. ~~**Multi-Agency Partnership:** Operating agencies delivers treatments with independent research organization providing random assignment, data collection, analysis~~

   ~~Name of Operating Agency 1_____~~

   ~~Name of Operating Agency 2_____~~

   ~~Name of Operating Agency 3_____~~

   ~~Name of Research Organization_____~~

D. **Other Framework** (describe in detail).

3. Unit of Analysis

Check only one

__A. People (describe role: offenders, victims, etc.) Stalking offenders.

~~__B. Places (describe category: school, corner, face-block, etc)_____~~

~~__C. Situations (describe: police-citizen encounters, fights, etc.)_____~~

~~__D. Other (describe)_____~~

4. Eligibility Criteria

    A. Criteria Required
        a. Has committed a stalking offence 24 months before the randomised controlled trial, with only offences reported on or after 31/03/22 being eligible for inclusion.
        b. Offender was at least 18 years old on the date of their index offence.
        c. c. It has a valid UK phone number, denoted by 11 digits, and a "+44" or "07" code, which together indicate a likelihood the number was recorded correctly and the offender is likely to be in the UK.

    B. Criteria for Exclusion
        a. The most recent offence was reported more than 12 months prior to the date of the RCT, reported before 31/03/22.
        b. Offender was under 18 on the date of their index offence.
        c. Suspects where there is a landline phone number recorded, the phone number is fewer or more than 11 digits, and the phone number begins with a code indicating it is not a UK number.

5. Pipeline: Recruitment or Extraction of Cases (answer all questions)

    A. Where will cases come from?
        a. Crime data drawn from Metropolitan Police Service records for the dates 31/03/2022-31/03/2024
    B. Who will obtain them?
        a. The principal researcher will work with the 'Operating Agency' to obtain data from police systems.
    C. How will they be identified?
        a. Data on stalker offending in the Metropolitan Police Service force area will be obtained for 31/03/2022-31/03/2024. This will be cleaned to exclude juveniles, deceased suspects, those found subject to a false allegation, those entered into police systems in error, and those with no phone number or an invalid phone number.

D. How will each case be screened for eligibility?
   a. The data has been requested so that categories will be set for things like sex/gender, date of birth, date of offence, etc., which will be used to exclude non-eligible cases from the case sample as set out above.
   b. Having a known telephone number which is deemed to be valid based on the above criteria.
E. Who will register the case identifiers prior to random assignment?
   Principal investigator
F. What social relationships must be maintained to keep cases coming?
   N/A
G. Has a Phase I (*no*-control, "dry-run") test of the pipeline and treatment process been conducted? **No**

If so,

- how many cases were attempted to be treated
- how many treatments were successfully delivered
- how many cases were lost during treatment delivery

6. Timing: Cases come into the experiment in (check only one)

F. ~~A trickle-flow process, one case at a time~~ ___
G. A single batch assignment___
H. ~~Repeated batch assignments~~ ___
I. ~~Other (describe below)~~ ___

7. Random Assignment

A. How is random assignment sequence to be generated?

(coin-toss, every Nth case, and other non-random tools are banned from CCR-RCT).

Check one from 1, 2 or 3 below

~~1. Random numbers table → case number sequence → sealed envelopes with case numbers outside and treatment assignment inside, with 2-sheet paper surrounding treatment~~ ___

2. Random numbers case-treatment generator programme in secure computer <u>X</u>

3. ~~Other (please describe below)~~ ___

B. Who is entitled to issue random assignments of treatments?

Role: Research agency

Organisation: University of Cambridge, Institute of Criminology

C. How will random assignments be recorded about case registration?

Name of data base: Stalker RCT

Location of data entry: Cambridge University

Persons performing data entry: Barak Ariel

8. Treatment and Comparison Elements

A. Experimental or Primary Treatment
1. What elements must happen, with dosage level (if measured) indicated.

Element A: A single text message sent to all stalking offenders assigned to the experimental group that focuses on the impact of stalking on victims and associated third parties.

~~Other Elements:~~

2.   What elements must **not** happen, with dosage level (if measured) indicated?
Element A: The text message must not fail to be delivered due to errors in the recorded telephone numbers on police data. This would result in poor dosage levels. This will be analysed through a delivery report which will allow researchers to determine 'treatment as delivered'. Cases where treatment fails to deliver will be excluded from follow-up analysis.

~~Element B:~~

~~Element C:~~

~~Other Elements:~~

B.   Control or Secondary Comparison Treatment
      N/A

3.   What elements must happen, with dosage level (if measured) indicated.
Element A: The nudge intervention delivered to the experimental group must not be inadvertently sent to the control group.

~~Element B:~~

~~Element C:~~

~~Other Elements:~~

~~4.   What elements must not happen, with dosage level (if measured) indicated~~
~~Element A:~~

~~——~~

~~Element B:~~

~~Element C:~~

~~Other Elements:~~

9. Measuring and Managing Treatments

A.   Measuring
1.   How will treatments be measured?
      A.   The WireFast system used to send the text messages returns a delivery report that indicates if the destination mobile number was invalid or valid but failed to send. This will be used to measure treatment fidelity.
2.   Who will measure them?
      A.   Treatment measurement will be conducted by the principal researcher.
3.   How will data be collected?
      A.   The principal researcher will manually search Metropolitan Police Services crime indices and record with a binary measure (1/0) whether a further stalking offence was reported between 15/4-15/8/24 to generate re-offending data. In addition, the number of stalking offences at different levels (s2a, the two versions of s4a of the Protection from Harassment Act and s8 of the Stalking Protection Act will be

recorded and assigned their respective CCHI score to allow for harm level measurement).

4. How will data be stored? On MPS database
5. Will data be audited? Yes
6. If audited, who will do it? Principal Investigator
7. How will data collection reliability be estimated? Random sample of cases (k-selection)
8. Will data collection vary by treatment type? No. The RCT involves a single treatment type.

B. Managing

1. Who will see the treatment measurement data? Research team
2. How often will treatment measures be circulated to key leaders? Bimonthly
3. If treatment integrity is challenged, whose responsibility is correction? The Principal Investigator.

10. Measuring and Monitoring Outcomes

A. Primary Outcome – the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI.

B. Secondary Outcome 1 – the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI, comparing male and female suspects.

C. Secondary Outcome 2 - the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI, comparing suspects with female victims to those with male victims.

D. Secondary Outcome 3 - the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI, comparing suspects by age category.

E. Secondary Outcome 4 - the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI, comparing domestic abuse to non-domestic abuse suspects.

F. Secondary Outcome 5 - the effect of the nudge text on stalker re-offending and harm levels as defined by the CCHI, comparing male to female suspects and the time lapse between their first offence and the nudge text intervention delivery.

A. Measuring

1. How will outcomes be measured?
   i. For all primary and secondary outcomes listed above the treatment and control groups will be assessed to measure how many suspects did or did not re-offend in the follow-up period. This will be used to generate a total for re-offending prevalence, and translated into a percentage. Treatment and control groups will then be compared with one another. In addition, for each suspect, the follow-up analysis will evaluate the number and type of stalking offences; a CCHI score will be generated for each primary and secondary outcome in the treatment and control group, and this will be used to compare crime harm scores between the two groups. Finally, for both prevalence and crime harm, a t-test will be conducted, and p-values will be used to estimate whether the results observed are statistically significant.

2. Who will measure them?
   a. Research team
3. How will data be collected?
   a. After the six-month window has elapsed, a second data dump from MPS systems will be requested. This will be used to assess the above measurements.
4. How will data be stored? MPS dataset

5. Will data be audited? Yes

6. If audited, who will do it? Principal investigator

7. How will data collection reliability be estimated? Yes

8. Will data collection vary by treatment type? No

~~If so, how?~~

B. Monitoring

1. How often will outcome data be monitored? Outcome data will be monitored on a single occasion by manually searching crime data from MPS systems after the four-month window post-intervention has elapsed.

2. Who will see the outcome monitoring data? Research team

3. When will outcome measures be circulated to key leaders? Bimonthly

4. If the experiment finds early significant differences, what procedure is to be followed? N/A

11. Analysis Plan

   A. Which outcome measure is the primary indicator of a difference between experimental treatment and comparison group? The number of suspects re-offending and the CCHI score in the treatment and control groups.
   B. What is the minimum sample size to be used to analyse outcomes?
      i. 400
   C. Will all analyses employ an intention-to-treat framework?
      i. The study will only analyse results through a 'treatment as delivered' framework.
   D. What is the threshold below which the percent Treatment-as-Delivered would be so low as to bar any analysis of outcomes?
      1. Lower than 60%.
   E. Who will do the data analysis?
      i. Research team.
   F. What statistic will be used to estimate effect size?
      i. Student T-Test, unpaired sample, single-tailed.
   G. What statistic will be used to calculate p values? See F.
   H. What is the magnitude of effect needed for a p = .05 difference to have an 80% chance of detection with the projected sample size (optional but the recommended calculation of power curve) for the primary outcome measure? TBD

12. Dissemination Plan

A. What is the date by which the project agrees to file its first report on CCR-RCT? (report of delay, preliminary findings, or final result). With dissertation

B. Does the project agree to file an update every six months from the date of the first report until the final report? N/A

C. Will preliminary and final results be published, in a 250-word abstract, on CCR-RCT as soon as available? N/A

D. Will CONSORT requirements be met in the final report for the project? (See http://www.consort-statement.org/) Yes

E. What organizations will need to approve the final report? (include any funders or sponsors).
    a. Cambridge University Institute of Criminology, Metropolitan Police Service.

F. Do all organisations involved agree that a final report shall be published after a maximum review period of six months from the principal investigator's certification of the report as final? N/A

G. Does principal investigator agree to post any changes in agreements affecting items 12A to 12F above? Yes

H. Does the principal investigator agree to file a final report within two years of cessation of experimental operations, no matter what happened to the experiment? N/A