

**POL 4237**

**Sara Thornton**

**Wolfson College**

**Supervisor: Professor Sherman**

**Predicting Serious Domestic Assaults and Murder in the Thames Valley**

**Submitted in part fulfilment of the requirements for the Master's Degree  
in Applied Criminology and Police Management**

**January, 2011**

## **Abstract**

Thames Valley Police uses a risk assessment model to identify those cases of domestic violence where the risk of future harm is high. This study looked at all the cases on serious domestic assault and murder between 2007 and 2009 to establish how accurate the risk assessments had been in predicting the serious harm. In 55% of cases there was no prior recorded contact with the police. In only five out of 118 cases was the case assessed as high risk. Effectively there was an 80% false negative rate. In the same period 1740 other victims were assessed as high risk arguably resulting in a 99% false positive rate. A case control study was carried out to try to identify any risk factors that marked out those offenders who committed the most serious domestic assaults from other violent offenders. The case control study found that those who committed serious domestic assault and murder were less criminogenic than the risk pool of all violent offenders - contrary to the central hypothesis of escalating violence. The study also found that male offenders who committed serious domestic assaults were more than three times likely to be suicidal than other violent offenders.

## **Acknowledgements**

Several Thames Valley colleagues showed immense patience as I endlessly sought data from our computer systems. I am particularly indebted to Peter Milne and Penny Fraser for all those many spreadsheets and to Michelle Olufsen for completing many checks on the Police National Computer. Dr Katrin Muller-Johnson and Dr Barak Ariel gave sound advice on anything statistical and Professor Larry Sherman whose enthusiasm and wisdom provided huge encouragement.

Thank you all very much.

## Contents

<b>Introduction</b>	6
<b>Literature and theory review</b>	12
Prediction in criminology	14
Hindsight fallacy	17
Base rates	19
Overview of violence in domestic situations	20
The relationship between lethal and non-lethal violence	20
The need for specificity	21
The prevalence of prior record of domestic violence	24
The development and evaluation of risk assessment models	27
<b>The development of risk assessment in England and Wales</b>	33
<b>Method</b>	40
The research questions	42
The use of case control samples	45
Selection of the case control	47
<b>Results</b>	53
Descriptive analysis of the Thames Valley cases	54
Hampshire comparator	57
Exploration of the false positives	61
Case control samples	63
Male block	63
Female block	67
<b>Discussion</b>	72
<b>Conclusions</b>	87

## List of tables

Table 1: Description of risk levels	38
Table 2: Murders in the Metropolitan Police Area	39
Table 3: Offenders' criminal histories	48
Table 4: Case control variables	52
Table 5: Domestic violence offences in Thames Valley	54
Table 6: Gender and outcomes	55
Table 7: Most recent prior victim contact with police for domestic incidents	56
Table 8: Prior incident contacts between victim and police	56
Table 9: Prior risk assessment	57
Table 10; Domestic violence offences from 2007/9	58
Table 11: Recording of violent crime with injury offences in England and Wales 2009/10	58
Table 12: Prior contact and risk assessment in two forces	59
Table 13: Risk assessment in two forces	60
Table 14: Risk assessments by Local Police Area	62
Table 15: Male case control – significant variables	63
Table 16: Male case control – relative risk ratio	66
Table 17 Female case control – significant variables	68
Table 18: Female case control – relative risk ratio	70

## Introduction

Prevention has been at the heart of policing since Sir Robert Peel established the Metropolitan Police Service. Peel's first principle of policing was, "the basic mission for which the police exist is to prevent crime and disorder". A preventative approach resonates with our modern concept of a 'risk society' (Ericson and Heggarty 1997) trying to manage risks in order to prevent harm. A risk assessment approach to predicting and preventing violence is forward-looking, intelligence-led, and involves multi-agency problem solving. Domestic murder and other serious assaults can often look predictable with hindsight and therefore many claim they could have been prevented. But how does murder look with foresight? To what extent is it possible to identify those cases which will result in the most serious harm? Is it possible to "live forward but seek to understand backward"? (Gottfredson 1987).

In 2003 and 2004 there were three particularly tragic cases of domestic murder in the Thames Valley and in every case the victim's families have asked whether the death could have been predicted and, if so, whether it could have been prevented. In response to these and other cases the force has made a significant investment in training, specialist units, multi-agency arrangements and a new risk assessment model. This thesis will focus on the risk assessment model which has been implemented over the last five years.

Some analysts argue that, "there is a clear relationship between everyday and life-threatening interpersonal violence ... and ... domestic homicide" (Websdale 1994, p204) and that past behaviour is one of the best indicators of future behaviour (Monahan 1981). By this logic the police service has the

opportunity to intervene in escalating patterns of violence. But, in how many cases of life threatening violence had the victim called the police on previous occasions? And if they have done so are we able to distinguish the murderous offenders from the vast majority of offenders who do not go on to commit murder or other serious violence?

There is a considerable body of literature on using prediction in criminology. The extent to which criterion measures can be predicted by other measures (predictors) operating at an earlier time has a very obvious application to preventing re-offending in general and violence in particular. However it is clear from the research that understanding of the risk of domestic violence has lagged behind other areas (Kropp 2004, Campbell 2005b). The evaluation of risk assessment tools is also in its infancy. While risk factors have been identified, their accuracy in terms of predictions has been low with many false positives and false negatives (Heckert and Gondolf 2004).

Studies have analysed cases where murder and serious harm have occurred and have concluded that there are risk factors that made this outcome more likely. However in terms of predicting murder and serious harm most tend to commit the “hindsight fallacy”. (Sherman 1992b). While some factors may well apply retrospectively to cases of murder and serious assault they may well apply to many other cases where there is no escalation of harm relative to the absence of those factors. We can only say with confidence that smoking predicts lung cancer, for example, when we know that



proportionately more people who smoke contract lung cancer than those who don't smoke (Sherman 1992b).

For this reason Richard Doll and Bradford Hill's early report in the British Medical Journal (1950) demonstrated the link between smoking and lung cancer by comparing a group of patients who had cancer with a group of general medical and surgical patients. It is not the fact that only 0.3% men with lung cancer did not smoke which demonstrates the link between cancer and smoking. It is the fact that this compares with 4.2% of the general medical and surgical patients and that this difference is significant.

The current Association of Chief Police Officers' definition of domestic violence is, "any incident of threatening behaviour, violence, or abuse (psychological, sexual, financial or emotional) between adults, aged 18 or over who are or have been intimate partners or family members regardless of gender or sexuality". While this is a much broader definition than some that are used there is often concern that those under 18 are excluded. That said, it does include family relations other than spouse or partner and includes victims of either gender. This is the true nature of domestic violence in the home – it is not just about women being assaulted by their abusive husbands and care needs to be taken to avoid the use of powerful stereotypes.

A risk assessment approach known as SPECSS+ (Separation, pregnancy, escalation, community issues, stalking and sexual assault) was introduced in most forces including Thames Valley during 2005 and 2006. This was then

replaced by DASH (Domestic Abuse, Stalking and Harassment) in 2009. The assessment is completed by operational officers attending any domestic crime or incident, although in some forces this is restricted to domestic crimes only. The officer asks the victim a series of questions in order to complete a form and on the basis of those answers cases are then assessed as standard, medium or high. The tool is based upon analysis of previous domestic murders and seeks to assess the potential dangerousness of cases so that the high risk case can be identified and then managed within a multi-agency panel known as a Multi Agency Risk Assessment Conference (MARAC) (Richards et al 2008). Despite the literature on the reliability of risk assessment, DASH and its predecessor were introduced without any form of effective evaluation of its use in the pilot sites. This lack of evaluation raises basic questions about the accuracy and reliability of the forecasts produced by the risk assessment. The present study is a first effort to answer those questions, at least for one police force.

This thesis examines cases of domestic murder and serious assault in the Thames Valley to establish whether the victims were known to the police and if they were, how accurate effective the risk assessment tool was in predicting serious harm. The thesis also compares the offenders in these cases where murder or serious assault took place with a case control of other violent offenders in the Thames Valley. This comparison attempts to identify whether there are any risk indicators that might have helped to predict murder or serious assault.

This thesis is presented in six sections. It commences with a review of the literature and theory. The next section provides an analysis of the background to the development of the current approach in Thames Valley. The research questions are then presented and a section on research method and design describes the approach taken. A range of the considerations and issues are discussed. The research findings are presented in four sections, the analysis of the cases, a comparison of Thames Valley with Hampshire, an analysis of the false positives in Thames Valley and then the Thames Valley case control study. Then follows a discussion of the issues identified in the studies and some suggestions for future research. Lastly the thesis makes some overall conclusions and recommendations.

## Literature and theory review

Modern thinking is dominated by talk about the need to manage risk and in doing so control the future – it is the obsession of the late modern age. Yet while risk may be talked about, it is little understood and accurate measurement is elusive.

Malcolm Gladwell's book, *Blink*, was an unexpected best seller in 2005 and argued that "decisions made very quickly can be every bit as good as decisions made cautiously and deliberately" (p14). Such an approach has little time for evidence-based decision making and Gladwell argues that too much information causes people to make bad decisions. Yet many of his examples are not of the gut instinct of the lay person but are about the trained judgement of the professional who has been exposed to the data over the years. Moreover he misses the key question of how we can accurately know which assessment is correct when many assessments are made in a context of unknown variables and intrinsic uncertainty. The police officer will never know the truth about the level of violence in a relationship but is being asked to make a judgement about dangerousness despite that ignorance.

This review is organised in the following specific areas in order to provide the broad contextual background and situate the research in brief summaries of the current literature.

- Prediction in criminology
- Overview of violence in domestic situations
- The prevalence of prior record of domestic abuse

- The development and evaluation of risk assessment models

### **Prediction in criminology**

It is widely argued that “if one seeks to control crime behaviour one needs first to be able to predict it” (Gottfredson 1987) and therefore it is important to use the scientific knowledge of prediction to prevent crime. The extent to which criterion measures can be predicted by other measures (predictors) operating at an earlier time has a very obvious application to preventing re-offending.

Therefore there is a considerable body of literature on the use of prediction in criminology and related social sciences. Early research was in the field of mental health where predicting dangerousness or offending was highly desirable. It was argued that “if an actuarially valid array of risk factors for violence could be reliably identified, clinicians could be taught to incorporate these factors into their routine practice, and the accuracy of clinical predictions of violence among the mentally ordered would be commensurately increased” (Monahan and Steadman 1994, p9). However while many instruments have been developed over the years in this field there remain significant issues about the accuracy of these tools even in the hands of psychiatrists and psychologists (Scott et al 2008). In particular, it will always be difficult to predict infrequent events, such as murder, because they have a low base rate of occurrence in most populations.

Similarly, the use of risk assessment for sexual offenders has been more developed than in the field of domestic violence. In 2007 Hanson and Morton-Bourgon published a meta-analysis of 79 different studies where the ability of risk assessment to predict recidivism among sex offenders had been evaluated. They found that actuarial measures were the best predictors followed by structured professional judgement and then unstructured professional judgement which was consistently the most inaccurate. Interestingly there was no significant difference between results for empirically based actuarial judgements and conceptually-based actuarial judgements.

The use of risk assessment in respect of domestic violence is much less developed. In 2005 Jacqueline Campbell observed in her Vollmer Award address that there were only eight or nine investigations that could be subject to the type of analysis undertaken by Hanson and Morton-Bourgon (Campbell 2005b) – and not all of those were published.

Risk assessment models in use in domestic violence can be organised in the same three categories as sexual offending - unstructured clinical assessment, actuarial assessment and structured professional judgement. Unstructured clinical assessment is what professionals, including police officers, have used over the years to try to estimate risk of future harm (Kropp 2004). However there is a reasonable body of research which shows that people's judgements about uncertain events do not conform to probability theory or statistics (Kahneman and Tversky 1982).

“Actuarial risk assessment derives a quantitative estimate of the likelihood that an individual will act violently during a given period of time based on an evaluation of those characteristics that have been shown to statistically differentiate those who are violent from those who are not” (Logan 2005 p5 ). Paul Meehl’s work over fifty years ago found that statistical or actuarial evidence was more reliable than clinical evidence (Meehl 1954) in other words that “simple linear combinations of cues outdo the intuitive judgements of experts” (Kahneman and Tversky 1982). As criminologist Leslie Wilkins put it, “statistical evidence is, perhaps, the highest form of evidence in that, if an estimate of probability can be made, it can be stated simply and the probable error of the estimation can be known” (Wilkins 1985, p42).

However the methodology of early actuarial work has also been criticised by commentators (Farrington and Tarling 1985). In respect of domestic violence the actuarial approach has been criticised for its lack of practical value and the absence of an agreed risk assessment model (Kropp 2004).

In recent years risk assessment models have been developed which involve the use of a structured risk assessment tool by a professional. They do not provide actuarial assessments but provide a framework or checklist for professional judgement. Jacqueline Campbell argued that, “the judgement of an experienced and knowledgeable practitioner and a well-validated instrument or system along with the input of an abused woman herself will probably prove to be the best approach to lethality assessment” (Campbell 2005a, p1208). The structure is based upon theoretical and empirical



knowledge. While it therefore does have some empirical underpinning which makes it preferable to an unstructured clinical assessment it still allows for professional judgement. Nonetheless, there is still little evidence that such judgement is reliable. (Kropp 2004).

A few commentators have also questioned whether prediction is ethical. We all predict the behaviour of others to make sense of life but is a more systematic approach defensible? Wilkins (1985) argued that he drew his ethical boundary at recidivism and that predictive work should not be done around those who have not yet committed crime. Predicting domestic murder or re-assault is however fundamentally different from other forms of prediction because the identity of the victim is known. It is not the general population at large that is at risk but a known victim. This means that safety planning is absolutely key. Risk assessment enables victims to take steps to protect themselves which will paradoxically make it much more difficult to assess the accuracy of the risk assessments.

### ***Hindsight Fallacy***

“This fallacy lies in using hindsight, rather than foresight, to draw conclusions about causation” (Sherman 1992, p232). Hindsight fallacy is used to refer to the fact that not only does knowledge of the outcome affects judgements about probability but also that people are unaware of the extent to which this outcome knowledge has affected their perceptions (Fischhoff 2003). The tendency, and therefore weakness, for historians to view the past in a deterministic way is widely understood. Events are assumed to have

unfolded in a linear fashion and that the outcome could not have been otherwise. Hindsight rather than foresight is used to make judgements. Similarly accidents are scrutinised to “uncover or impose a pattern that will increase their perceived predictability and avoidability” (Fischhoff 2003, p304).

Any risk assessment for domestic murder needs to avoid the hindsight fallacy. Sherman identified the mischief that the hindsight fallacy had caused for the Kansas City Police Study on domestic violence (Breedlove et al, 1977 as cited in Sherman 1992b). The Kansas City Police study had shown that the police had been called to the address of the victim or offender in the two years preceding the domestic murder in 90% of cases (Breedlove et al, 1977 as cited in Sherman 1992). However this pattern would only be noteworthy if domestic murder occurs in a much higher proportion in addresses where the police have been called than addresses where the police have not been called. In this study there was no data on the overall number of calls to addresses. There was no case control sample. Moreover the fact that in Kansas City many people may live at the same address was overlooked. Therefore no conclusions can be drawn about the relationship between police being called to an address and domestic murder.

Fischhoff (2003) conducted four experiments which showed that knowledge of an outcome affects its perceived likelihood and that those making judgements believed that this inevitability was apparent with foresight. Not only did SPECCS+ and DASH not compare the cases with any control but the model is in danger of attributing inevitability where none exists. The brain wants to

make sense of the past, however such unconscious determinism undermines our ability to learn from the past. “Where information is limited and indeterminate, occasional surprises – and resulting failures – are inevitable” (Fischhof 2003, p311).

### ***Base Rates***

There is another potential difficulty with the assessments of risk when trying to predict very rare events which is covered in literature from Paul Meehl onwards. The research shows that base rates are either ignored or underweighted (Tversky and Kahneman 1982). SPECCS+ and DASH were developed in London where domestic murder is a very rare event with significantly less than 100 in a year among a population of over seven million. And the way in which cases are put into three categories of standard, medium and high is therefore potentially misleading because of the difference between proportion of deaths and number of deaths. Ken Pease (2010) illustrated the difficulty with the following analysis.

If the likelihood of death is 10% in high risk cases and 1% in standard cases it is still likely that there will be more deaths among standard cases. The *proportion* of deaths is highest in the high risk group, but the *number* of deaths is highest in the standard risk group. For example, if there are 1000 standard cases where the percentage of deaths is 1% there will be 10 deaths and if there are 10 high risk cases where the percentage of deaths is 10% there will be one death. Clearly there is huge temptation for officers to over

grade the risk in order to avoid the inevitable criticism when a death occurs among the standard risk group.

## **Overview of violence in domestic situations**

### ***The relationship between lethal and non lethal violence***

Theories about preventing domestic violence are predicated on a relationship between escalating domestic violence and lethal violence. However the research is less clear cut. Dobash et al (2007) compared a sample of British men convicted of lethal violence and a sample of British men convicted of non lethal violence in order to ask the question whether lethal violence is always associated with non lethal violence and whether the same factors were involved. They found that previous violence against the victim was less prevalent in lethal case than non lethal cases. In 41% of lethal cases there was no previous violence against the victim compared with 0% in non-lethal cases. While this research is based upon interviews with convicted murderers, serving life sentences at selected prisons, who may have under played their previous violence this still poses a challenge to the accepted wisdom that there is a progression from non lethal to lethal violence.

Those that killed had more conventional backgrounds than those who had not, with the killer's fathers more often in white collar jobs and mothers who were housewives. Those who used non lethal violence were more likely to have been brought up in a home where their father had alcohol problems and

physically abused them and their mothers. It is important not to restrict research to cases where there is a prior history but to look at those who kill but who have not previously used violence against the victim. The research found that “Some of the men who killed did not have problematic lives as children or adults, had no history of using violence to those victims or to others and were not drunk at the time. Men with these characteristics would be unlikely to be assessed as at risk of committing lethal violence and, as such, present a challenge to those who assess and manage risk” (Dobash et al 2007, p349).

This leads to a real concern that in risk assessing all the victims of non-lethal violence in such detail it is inevitable that many offenders who will commit really serious harm will not be identified. Moreover Jacqueline Campbell argues that the risk factors for murder are not exactly the same as risk factors for re-assault (Campbell 2005a).

### ***The need for specificity***

In the same way, Michael Johnson has argued that “we are trapped in overgeneralisations that assume intimate partner violence is a unitary phenomenon” (Johnson 2008, p3). He has developed a useful typology for domestic violence and has argued for differentiating between types of violence. He identifies four types of domestic violence:

- Intimate terrorism – the use by one partner of violence to gain control;

- Violent resistance – the response to the controlling behaviour;
- Situational couple violence – violence without the desire for control;
- Mutual violent control – both parties use of violence to gain control.

Johnson argues that much of the disagreement in the literature on domestic violence is caused by this lack of specificity about the type of domestic violence. This in turn is caused by the sampling that is done – if the research is based on surveys a very different picture will appear than if the research is based upon interviews in a shelter. Johnson illustrates this by analysing a study by Irene Frieze in Pittsburgh in the 1970s. Three groups had been selected – a survey of domestic violence victims who used a certain launderette, a group of women who lived in a shelter for battered women and a group of women who had filed for court orders. Johnson then used the questions from the research to assign cases to his typology.

In respect of men's violence he found that in the general survey sample 86% of cases were situational couple violence, 11% were intimate terrorism and 3% violent resistance. Whereas in the court sample 46% were intimate terrorism, 37% situational couple violence, 11% mutual violent control and 6% violent resistance. In the shelter sample 66% were intimate terrorism, 28% situational couple violence, 4% violent resistance and 2% mutual violent control.

It is clear that in the case of male offenders that intimate terrorism will dominate to a much greater extent among those surveyed in a shelter as

opposed to a general population of women who are suffering from domestic violence. Conversely situational couple violence will dominate those surveys which assess the general population of women suffering from domestic violence as opposed to a survey of those in a shelter.

In respect of women's violence he found that in the general survey sample 86% of cases were situational couple violence, 10% were violent resistance and 3% intimate terrorism. In the court sample 41% were violent resistance, 31% situational couple violence, 14% mutual violent control and 7% intimate terrorism. In the shelter sample 61% were violent resistance, 32% situational couple violence, 5% intimate terrorism and 3% mutual violent control.

Where the offenders are women there are many more cases of situational couple violence in the general survey than among those in the court or shelter sample. However in the shelter 61% of women were in the category of violent resistance with 41% of women in this category in the court sample. This suggests that women offenders who are in the shelter are dominated by those who have fought back and that many of them are also being dealt with in the court system – but that their presence in the general population is much smaller.

Any analysis that is undertaken as part of this study will be affected by the origin of the data used. Johnson argues that each type of domestic violence has different risk markers but that this is often overlooked. For example, many studies have identified risk markers after analysing surveys which

Johnson argues have a bias towards situational couple violence. Similarly if the sample is taken from criminal justice data Johnson's research makes it clear that there will be a bias towards violent resistance for women and intimate terrorism for men.

He also argues that as the risk markers vary so does the appropriateness of interventions. For example, restorative justice might be appropriate in situational couple violence but it would never be the right approach in cases of intimate terrorism.

Johnson's work provides a high bar for any risk assessment. Firstly, there is the issue of how the risk indicators been identified – survey, criminal justice system or shelter? Then there is the failure to identify different types of violence so that risk assessments may end up being a mix of risk factors for intimate violence and situational couple violence? Arguably such an approach would undermine the validity of any risk assessment and therefore will fail to identify those cases that are high risk.

### **The prevalence of prior record of domestic violence**

The accepted wisdom is that escalating domestic violence provides an opportunity for the police and other agencies to intervene to prevent further harm. However the prevalence of a prior history in cases of domestic murder varies widely across the many studies that have been undertaken. The early research in Kansas City in 1971-1972 showed that police had been called to



the **address** of the victim or suspect in domestic homicide cases over last two years in 90% of the cases and that police had been called to the address more than five times in 50% of the cases (Breedlove et al 1977). This research had a significant impact on the development of theory and practice.

Similarly, Neil Websdale's influential work on domestic murders in Florida in 1994 showed that 86.6% had a prior history of battering (Websdale 1999) and he argues that prior history was "widely recognised....although greater prevalence has not been clearly spelled out empirically" (p19). Websdale quotes the Detroit and Kansas City studies where 85-90% of cases had prior contact with police. However he does not look at other research that contradicts this nor does he examine why different groups might call the police more often or indeed how many individuals might live at one address in some neighbourhoods. His explanation for any low levels of recorded prior contact in any other study is poor record keeping by the police.

Since the early work in Kansas City there have been several other studies which found much less prior contact. A study based in Atlanta in 1984 showed that 30% of all domestic homicides had been preceded by an offence involving either the victim or the offender in the last four years (Saltzman et al 1992). Research in Milwaukee in 1987-1989 found that of 33 domestic homicides only one couple had a prior report of domestic violence, but that study did not examine other criminal history records as the Atlanta study did (Sherman et al 1991).

Research in Minneapolis in 1985-1989 showed that given the 52 homicides in the sample the rate of domestic homicide was 16.83 per 1000 for addresses with more than nine call outs (and this was considerably more than the .28 per 1000 where there had been no call outs). However while the risk is greater it remains the case that over 98% of the time any prediction arising out of prior reporting would be incorrect (Sherman 1992b, p7).

Research in Victoria, Australia, in 1988-1992 found that 90% had no prior contact in respect of the 82 cases of domestic homicide studied and that couples who report incidents almost never experience a homicide (Sherman and Strang 1992a). Analysis by the Canadian Bureau for Justice Statistics in 1991 found that 42% of cases had prior contact (Canadian Bureau for Justice Statistics 1992).

The span of results is quite staggering – from 3% of cases to 90% of cases having prior contact. There are clearly many variables which have not been controlled. For example, the units of analysis vary from addresses to couples to individual victims or offenders. The measures vary from incident dispatch systems to individual criminal histories. Also record keeping in various police departments will vary over time and location. The socio-economic make up of the population will vary between cities and countries and therefore the propensity of individuals to call the police. For example, those with few resources may have no one else to turn to. It cannot be assumed that there is a simple relationship between seriousness and calling the police which is true for all time and in all places – even with consistent measures and units of analysis.

## **The development and evaluation of risk assessment models**

Risk assessments for domestic violence were first developed in the United States about thirty years ago. Barbara Hart, who had a legal background, was the first to develop a lethality assessment based upon her work as a practitioner (Hart 1984). However Jacqueline Campbell's Danger Assessment, which was developed in the medical field, is probably the most widely used.

Jacqueline Campbell works in the medical field and her work sees domestic violence as a major public health problem to which a harm reduction model is an appropriate response. As Van Wormer and Roberts (2009) argued, "harm reduction is about preventing problems before they start and developing safety and treatment plans for persons who are at risk" (p18). Campbell's Danger Assessment was then developed into a "lethality screen for first responders" for law enforcement officers in Maryland. This assessment tool has 11 questions and is used to predict the danger and potential for lethality in situations. If a victim is assessed as high risk then they are referred to local programmes. Evaluation of this approach in 2006 and 2007 showed that 57% of victims screened were assessed as high risk and 54% of this group spoke to a programme counsellor on the phone. While one such victim was murdered the number of domestic homicides in Maryland in 2007 was the lowest since 1991 (Sargent and Campbell 2008). However the numbers are very small and this could reflect random variation.

While Campbell's original risk assessment was produced in the 1980s based on her work with practitioners, her National Institute of Justice research on risk assessment has an empirical base and used a multi-site case control study (Campbell et al 2003). This research was then used to revise the Danger Assessment. All the cases of femicide in 11 cities between 1994 and 2000 were examined to identify 545 closed cases where the perpetrator was a current or former intimate partner. In each of these cases a knowledgeable family friend or relative was identified from police records, approached and asked to participate. In 373 (68%) of cases a proxy was identified and in 307 cases they agreed to participate. Two criteria were used to exclude cases – age and no previous abuse – which left 220 cases in the study. A control group of 343 abused women was then identified and logistic regression used to estimate the association between identified risk factors and the risk of femicide.

The study includes both bivariate and multivariate analyses. In addition to the simple comparison between the cases and the control for all the variables the authors use logistic regression to identify the risk factors for intimate partner femicide. The paper argues that multivariate analysis which gradually added similar variables from a continuum ranging from the most distal, such as demographics, to the most proximal, such as the weapon used, is most insightful.

The strongest demographic risk factor for intimate partner femicide was the abuser's lack of employment similarly a university education was a protective factor. Ethnicity was not independently associated with risk of intimate partner femicide after controlling for other demographic factors. At an individual level the abuser's access to firearms and use of illicit drugs were strongly associated with intimate partner femicide but excessive abuse of alcohol was not. Having left an abusive partner after living together or having a child at home who was not the abuser's biological child also increased the risk of femicide. There was a nine fold increase in the risk when a combination of separation and a highly controlling abuser existed. Threats with a weapon and threats to kill were not surprisingly also associated with higher risks of femicide.

While stalking, strangulation, abuse during pregnancy, escalating violence, suicide, perceptions of danger and child abuse were more closely associated with femicide in the bivariate analysis, this was not the case in the multivariate analysis. Campbell explains this by arguing that "these characteristics of abuse are associated with previous threats with a weapon and previous threats to kill the victim, factors which more closely predict intimate partner femicide risk" (p1092). Importantly, previous arrest of the abuser for domestic violence was associated with a reduced risk of femicide.

Campbell's analysis would appear to be one of the few studies that used a control population to assist in the identification of risk and it has been evaluated prospectively. The report acknowledges the potential for the proxy

interviewed to be unaware of the complete history and it is also possible that knowledge of the outcome would both affect the judgement of those interviewed about the significance of pre-cursor events and their willingness to disclose information which might make them appear neglectful of the victim in the light of the outcome.

The Campbell Danger Assessment tool is not something that can be simply transplanted across the Atlantic for use in the Thames Valley. Apart from significant cultural, political and social differences between the two countries there are specific concerns about its suitability for the Thames Valley environment. For example, the study just looked at the murder of women by a current or former intimate partner whereas the identified domestic violence offences in Thames Valley include violence against men and among adult family members.

The Campbell study focused on urban populations which are demographically very different from the population of Thames Valley. Most important the Campbell study excluded all cases where there was no prior contact which in Thames Valley would mean excluding 55% of all cases. While it is understandable why this has been done because Campbell was looking to assist medical staff identify those patients that were most at risk it does lead to an incomplete view of the risk factors.

The Danger Assessment has been independently tested in two studies (Bennett et al, 2000 and Heckert and Gondolf 2004). There was some

predictive validity for re-assault, rather than lethality, but there was a high rate of false positives. As has been considered earlier, low base rates make lethality prediction very difficult and Campbell accepted, “it is always difficult to predict, with our current statistical models and limited resources for longitudinal research, a seldom occurring event” (Campbell, 2005a p1210).

The Danger Assessment has also been evaluated alongside three other risk assessments in use in the United States (Roehl et al 2005). A prospective study of the accuracy of four models – the Danger Assessment, DV-MOSAIC, the Domestic Violence Screening Instrument and the Kingston Screening Instrument for Domestic Violence was carried out to discover whether and how well they assess the likelihood of future violence. The study showed that while the assessments were better than chance there was a high level of false positives and that while the false negatives were not very high they are of great concern. The study did not compare the tools with expert practitioners’ unstructured assessments but does conclude that the ideal approach would be to have an experienced practitioner using a validated tool. Compared with other fields this level of prospective evaluation is extremely limited and very few have been evaluated by those other than their authors (Campbell, 2005a).

But what are the acceptable levels of sensitivity? Clearly as far as the victim is concerned a false negative is problematic, where the criminal justice system is concerned a false positive is of concern. The current models would appear to be work in progress and it is important that they are subject to

continuing rigorous evaluation - as Kropp argues, there is 'an unresolved schism between science and practice' (Kropp 2004, p682).



## **The development of risk assessment in England and Wales**

A succession of Labour Ministers had been determined to improve the response to domestic violence and this culminated in the Domestic Violence, Crime and Victims Act 2004. The developments of risk assessment within the Metropolitan Police and South Wales Police were against this general backdrop of political activity and focus.

In the Metropolitan Police the first predictors were identified from a study of thirty domestic violence murders between 2001 and 2002 (Richards 2006). The cases were analysed and characteristics identified with the intention of “identifying certain patterns and characteristics that could indicate potential lethality” (Richards 2004, p33).

Six high-risk identification markers were identified: separation, pregnancy, escalation, community issues, stalking, and sexual assault. It is not at all clear how the analysis of the 30 murders led to the identification of the risk factors. For example two cases out of 30 involved pregnancy or new birth and this has become a risk factor. Similarly “community issues and isolation” is a factor in 14 (47%) cases and becomes a high risk factor without any explanation of how such a judgement is being made and what is being counted in what is a very broadly drawn descriptor.

However there is no comparison of this group of cases with a control group of domestic violence cases that did not lead to murder. Without such a comparison it is not clear how the conclusions can be valid. It is also unclear whether the 30 cases selected were the whole sample in the period or, if they

weren't, whether there were any criteria for selection of those 30 cases. However this first study and the second study which will be described below does show the potential benefit of an analyst inside the organisation having access to the confidential data and therefore the opportunity to provide insights that would not be possible without access to the data.

The initial analysis of 30 murders was supported by a subsequent analysis of 400 other assaults termed "near miss" incidents and a review of international practice (Richards 2004). The analysis of "near miss" incidents does not make any comparison between the 30 cases of murder and this group of 400 as a control but it does look at 241 serious sexual assaults in some detail and then uses a group of other offences in a similar period in what is termed a control group. This control group included all cases of grievous bodily harm, actual bodily harm, kidnap, murder and attempted murder in January and February 2001 in the Metropolitan Police Area. From the published work it is not clear how this control group was used to draw conclusions.

In this analysis 241 sexual assaults with a domestic violence flag were analysed. The offences had all happened in the Metropolitan Police Area between January and April 2001. In 130 cases (54%) the victims had reported previous domestic violence to the police and in 49% of cases the offender had a previous criminal history (but as 44 offenders could not be identified that number is likely to be higher).

This analysis is used to corroborate the six high risk identifiers. For example, two observations are used to corroborate the claim that sexual assault is a good indicator of risk of lethality. The analysis showed that 16% of sexual assaults involved a weapon other than fists and the conclusion reached is that the offenders are using the weapons to injure the victim rather than for compliance. It is not at all clear how this conclusion can be drawn and even if it can, how does that result in sexual assault being an indicator of lethality? It is also asserted that sexual assaults tend to result in more serious injury than other types of allegations of domestic violence without showing a clear comparison between the sexual offences and the control group. Again even if a comparison had been completed between the sexual assaults and the other offences in the control sample it is not necessarily the case that sexual assault is an indicator of lethality. The section concludes “If victims are systematically raped and abused the risk of homicide increases” (Richard 2004, p17) – this may well be true but it is not a conclusion that can be reached from the analysis presented.

Separation is claimed to be another high risk identifier. The analysis shows that 116 (48%) of victims had separated or were separating at the time of the sexual assault. While it can probably be assumed that women who are separating were over represented in the group of women who were sexually assaulted in London in 2001, again it is not clear how this fact means that separation is therefore a high risk identifier for domestic murder. Lastly, 11 (5%) of victims were pregnant at the time of the assault and while that also means that pregnant women are over represented in the group of women who

were sexually assaulted in London in 2001 it is again not clear how being pregnant increases the risk of being murdered compared to those victims of assault where escalation never occurs. And in any event many victims of domestic violence are male and therefore pregnancy clearly irrelevant.

At later stage mental health, suicide-homicide, threats to kill, jealous and controlling behaviour and alcohol/drugs abuse were identified as high risk indicators. It is claimed that the model is similar to SARA, a model from the US, and is about prevention not prediction (Richards 2006). The original model was known as SPECSS+ and was piloted in the Metropolitan area, West Yorkshire and Thames Valley. While the literature suggests that the Police Standards Unit, which formed part of the Home Office was to evaluate the new approach it has not been possible to find any such evaluation.

At about the same time South Wales Police had been developing its own risk assessment. This was developed from a study of 47 homicides and relevant literature. A checklist of 15 yes/no questions was developed and any victim scoring over seven was considered to be high risk (Robinson 2006). This resulted in a similar but different risk assessment to that developed in the Metropolitan Police but again there was no comparator group.

In 2009 the SPECSS+ model was developed into the DASH risk assessment model which now consists of 27 questions covering 15 high risk factors. The complete questionnaire used in Thames Valley can be found at Appendix 1. When the risk assessment is complete the case is then categorized in the following way.

<b>Standard</b>	Current evidence does not indicate likelihood of causing serious harm.
<b>Medium</b>	There are identifiable indicators of risk of serious harm. The offender has the potential to cause serious harm but is unlikely to do so unless there is a change in circumstances, for example, failure to take medication, loss of accommodation, relationship breakdown, drug or alcohol misuse.
<b>High</b>	There are identifiable indicators of risk of serious harm. The potential event could happen at any time and the impact would be serious. "A risk which is life threatening and/or traumatic, and from which recovery, whether physical or psychological, can be expected to be difficult or impossible" Risk of serious harm (Home Office 2002 and OASyS 2006)

Table 1: Description of risk levels

## Evaluation

There was limited evaluation of the pilot sites and there has been very little evaluation of DASH or its predecessor SPECSS+. Sully and Greenaway (2004) looked at five cases of homicide in London and found that in five out of the six SPECSS+ indicators were present in some of the cases. This conclusion is itself not particularly insightful but the fact that they had selected five cases out of a possible 41 means that it is even less persuasive.

Humphreys et al (2005) completed a review of the implementation of SPECSS+ in the Metropolitan Police and West Yorkshire Police but there was little attempt to empirically test the model. Most of the data gathered was qualitative and consisted of interviews with officers and staff. That said they concluded that the operation of the risk assessment model was operating very differently in the four study sites.

Richards et al (2008) claim that there is evidence that the correct use of SPECSS+, in its original format, can reduce the incidence of domestic violence. Table 2 shows how rates of murder overall have been dropping over the last decade in London and while this does mean that the domestic murder rate has fallen the explanation could lie in factors other than the use of the risk assessment. The percentage of these murders which are flagged as domestic has remained at about 25% over the years.

2008/9	146
2007/8	156
2006/7	162
2005/6	168
2004/5	182
2003/4	204

Table 2: Murders in the Metropolitan Police Area

DASH does not claim to be able to predict violence but that it aims to prevent violence – a claim made by other approaches which can be described as structured professional judgement. This is particularly relevant in respect of false positives where although the risk may be assessed as high, successful intervention and management may have prevented escalation and further violence. This research will undertake a brief analysis of the false positives in Thames Valley.

## Method



The literature search found that while the broader research on prediction illustrated the fact that there is potentially great benefit in identifying effective risk assessment models there is very little evaluation of their use. Even in the United States there is only a modest amount of evaluation – in the United Kingdom there has been no effective evaluation of the use of DASH or its predecessor SPECSS+.

The method used for this thesis has two distinct parts. The first part was to select a number of Thames Valley cases where domestic murder or serious assault had taken place (the numerator). These cases were then analysed to answer the first two research questions. In effect this provided an evaluation of whether the SPECSS+ and DASH risk assessment process was predicting murder or serious assault in Thames Valley.

This analysis revealed a high level of false negatives and so while it was clear what was not working the key question was – what might work? While it might have been possible to simply transfer one of the more reliable models from the United States such as the Campbell Danger Assessment into the United Kingdom context that was not appropriate for a couple of reasons. Firstly, the profile of domestic murder is so very different in the United States. Most importantly the Campbell study had excluded all the cases where the victim had had no prior contact with the police – a factor which the numerator analysis had revealed to be significant.

As has been discussed in the previous chapter one of the most serious omissions in the development of SPECSS+ was the lack of comparison with any form of control group. It is easy to say that certain risk factors are present in a number of murder cases but how does the presence of those risk factors compare with a group where no murder took place? Therefore, the second part of this thesis presents a case control sample selected from an evidence based risk pool.

### **The research questions**

In how many cases did the victim have a prior history of reported domestic violence?

The use of SPECSS+ and DASH is predicated on the assumption that there are previous crimes or incidents to which the police are called which provide an opportunity to risk assess and intervene. The Metropolitan Police Study in cases of sexual assault showed that 54% of cases had previous victimisation reported to the police and so it was important to identify how many victims of serious domestic violence have had prior contact with the police in respect of domestic violence in Thames Valley.

Using the Cedar data base (the Thames Valley Police crime recording system) all the cases of domestic violence between 1 January 2007 and 31 December 2009 in the following categories were selected:

- murder,

- attempted murder,
- manslaughter,
- grievous bodily harm with intent

One hundred and eighteen cases were identified and all have been used in the analysis. While the research is particularly concerned about lethal violence the sample includes many near lethal offences of attempted murder and grievous bodily harm with intent. There were only thirteen cases of murder in the period and while a wider time span could have been used the study would have been undermined by the fact that the risk assessment has only been used force wide for five years. An alternative approach would have been to use data from more forces. But there are many small differences in the way that data is recorded in other forces and their data is less accessible. Lastly, the difference between murder and attempted murder may be as much about the speed and quality of medical care as the intent to harm. Thus a broader sample in Thames Valley was the best option for the present study.

The ACPO definition of domestic violence had been used to flag these offences so they were not limited to female victims or to those who have had an intimate sexual relationship. The cases included 51 (43%) male victims and 67 (57%) female victims. While it could have been possible to eliminate all the male victims this would have misrepresented the nature of domestic violence and would have nearly halved the chance of finding an appropriate indicator of harm.

Prior contact was then identified by searching the Cedar data base for all cases where the victim had contacted Thames Valley Police since 1 January 2000 until the date of the offence in the sample. While it could be argued that this excludes some prior victimisation this is a reasonable cut off given the fact that a minimum of seven years would have elapsed between contacts. It is always possible that victims may have made contact with the police but no record was made. The National Crime Recording Standard, introduced in 2003, and the robust auditing of its requirements make this less likely since 2003. The vast majority of victims for whom there was prior contact recorded had reported offences by the same offender (46 of the 53 cases or 87%). While there are cases of victims who have been abused by many offenders, the cases in the sample showed that most offending was done in the same relationship.

How accurate were the risk assessments based on prior history in terms of false positives and false negatives?

The 118 cases were then examined to see what the prior risk assessment was in order to assess the accuracy of such assessments using the DASH model or its predecessors. Given that the prior contact was assessed to the year 2000 which was before DASH or its predecessor SPECSS+ was introduced in six out of 53 cases there had been no risk assessment at the time of prior contact. SPECSS+ was introduced over a period in Thames Valley because it was first piloted in Oxfordshire in 2004 after a woman had been murdered in the police station car park as she came for help. The

approach was then introduced force wide at a later date. DASH was introduced in 2009.

A further search of the Cedar data base was then carried out to identify the number of high risk assessments made in all domestic violence cases between 1 January 2007 and 31 December 2009. Two thousand seven hundred and twenty one cases which had been assessed as high risk were identified and these were then analysed briefly to examine the issue of false positives.

Might the use of a case control sample produce more accurate risk indicators than analysis based on the numerator alone?

### **The use of case control samples**

The analysis of the Thames Valley cases identified significant shortfalls in the risk assessment currently in use. It was clear from the research undertaken that the risk indicators had been developed from the analysis of a group of murders without making any comparison with a larger risk pool. The work undertaken by Jacqueline Campbell and her team compared cases where the victim had been murdered with cases of domestic violence where there had not been lethal violence. In this respect she used a case control sample and then completed a bivariate and then a multivariate analysis. Campbell only looked at the victims and while a multi-method approach which looked at both

victim and offender might be the best approach given the constraints of time and resources for this thesis research the focus will be restricted to offenders.

Case control samples are observational rather than experimental. In a case control sample the researcher works from the outcome to the predictors in a retrospective study and assesses the **relative risk** between the cases and a control sample in respect of exposure to a particular predictive condition. They are used extensively in medical research where experimental approaches to risk factors (such as asking people to smoke) carry significant ethical issues. Similarly they are appropriate for studying the epidemiology of domestic violence.

Essentially, in a case control design the researcher works out how more likely the disease is in the group exposed to a risk factor as compared with the unexposed. The incidence rate among the exposed is calculated by dividing those who have been exposed to a risk factor and have the disease by the total number who have the disease. The incidence rate among the unexposed is calculated by dividing those who have not been exposed to the risk factor but have the disease by the total number that have not been exposed to the risk factor. The ratio between the incidence rates is then calculated by dividing the incidence rate in the exposed group by the incidence in the unexposed group. If  $R > 1$  then the association is positive if  $R < 1$  then the association is negative. (Schlesselman and Stolley 1982)

Case control designs do not provide incidence rates, they show the relative risk after exposure. This is because both the cases and the control have

been selected from the wider population. And as Monahan and Steadman (2003) argued, a true prevalence rate should be calculated by taking the known cases divided by the size of the population from which the cases are drawn. Thus the prevalence of lethal or serious domestic assault in Thames Valley over three years was 118 cases divided by three or an average of 39.3 per year, divided by the population of 2.2 million = 17.8 per million per year. Isolating risk factors from this outcome could identify sub-groups of the population for whom this rate is much higher but is beyond the scope of the current study.

### **Selection of the case control**

The early analysis of the numerator group completed above showed that only 45% of victims had a prior reported history of domestic violence. This suggested that prior cases were not the optimal denominator group from which to select a sampling framework from which appropriate case controls could be selected. Therefore a case control from a broader denominator group needed to be identified for the thesis research.

If looking for a risk pool among victims was not appropriate the obvious thing to do is to look at offenders. Several different approaches were tried before an appropriate group of offenders was selected. Analysis of the 118 cases showed that there were 123 offenders because four cases had more than one offender. However there were three offences with the same offender making the final total 120. There was an offence of grievous bodily harm, which appears to have been a brawl at a funeral, where there are five offenders.

For the purposes of this research those offenders who fall within the ACPO definition of domestic violence have been included.

<b>Previous criminal record</b>	<b>Number</b>	<b>% of cases</b>
Offender has prior criminal record	83	69.1
Offender has prior criminal record for violence	58	48.3
Offender has prior arrest	92	76.6
Offender has prior arrest for violence	71	59.1
Offender has no prior criminal record	37	30.8

Table 3: Offenders' criminal histories

Schlesselman and Stolley (1982) argued that controls must be similar to the cases in terms of potential for exposure. While having a prior arrest and having a prior criminal record has the highest prevalence in the group above and therefore it could be argued that these categories would provide the largest numerator group and therefore be the most appropriate. It is the very size of these groups which make them less attractive. As over 40% of all men have criminal records by the age of 50 (Farrington et al 2006) this would be a very large risk pool and could lead to many false positives. However, table 3 shows that those who have been arrested for violence are less likely than those arrested for other offences to produce false positives. They make up nearly 60% of the numerator even though arrests for violence are a minority of all arrests. This makes it a much more efficient group to examine and it was chosen as a basis for selecting the case control.



The category of prior arrest for violence included all offences within the Home Office category of violent crime which are listed at Appendix 2.

A sampling framework consisting of all those arrested for violence from 1 January 2007 to 31 December 2009 in Thames Valley was produced. This included offences of violence with or without a domestic violence flag and the Home Office category of violent crime listed above was used. This was the same definition that was used to identify the numerator group. A sample was randomly selected from this group of offenders using stratified sampling within age groups. Age was restricted to those cases where the age of the suspect fell within the range of the 120 domestic violence offenders. The age range is however very wide. The sampling framework found 49,000 cases with male offenders and over 9,000 cases with female offenders.

In the same way that an experimental and control group need to be similar in a randomised control trial so there should be similarity between the cases and the control in a case control sample. While a case control study cannot match on the unknown characteristics in the way that a randomised control trial, it can at least match on selected characteristics that are theoretically relevant. Ariel and Farrington (2010) argue that if it is known that participants may vary significantly then the control should not be simply randomly allocated but should be randomly allocated into blocks. This is known as randomized block design.

The sample was organised into two blocks, one for male offenders and one for female offenders. The reason for this was the very great difference in female and male offenders in terms of seriousness in the cases in the numerator – the risk of death per attack was four times higher where a male was the offender in the numerator. Jacqueline Campbell argued that the risk factors are different for men and women (Campbell 2005) and it was clear in Michael Johnson's work that the type of violence varied significantly between the two genders (Johnson, 2008). The female control sample was set at 100 cases and the male control sample was set at 150 cases. This was approximately double the number of numerator cases in each block and follows the advice of Schlesselman and Stolley (1982) who suggested that a case control should be two to three times the number of the cases.

These control samples were then tested for exposure in respect of the following criteria in comparison with exposure in the cases and a bivariate analysis completed. Case control studies can be analytic or exploratory where there are multiple hypotheses to test because there is no particular hypothesis. In this thesis the approach is exploratory as there is no particular hypothesis to test and the two groups will be compared in respect of exposure to a number of criteria.

While the selection of these criteria was limited by the availability of data there was a rationale for the selection of the criterion which are shown in table 4. Exposure to the criteria was assessed using the Police National Computer (PNC). The hypothesis for each area tested was: 'In comparison with

offenders who do not commit serious domestic harm those who do will be more likely or less likely to.....’

<b>Number of prior arrests</b>	There are some studies which suggest that those who commit serious domestic violence often have less substantial criminal histories than those who commit the less serious domestic violence (Dobash et al 2007)
<b>Number of prior arrests for violence</b>	As above
<b>Number of prior convictions and cautions</b>	Many who kill their partners are already in the criminal justice system (Stanko 2003).
<b>Number of prior convictions and cautions for violence</b>	As above
<b>Age at first arrest</b>	
<b>Age at first arrest for violence</b>	
<b>Age at first conviction</b>	
<b>Age at first conviction for violence</b>	
<b>Employment</b>	Campbell’s work in the US found that unemployment was one of the most significant risk factors for lethal attacks.
<b>Weapons</b>	The individual has used a weapon to commit an offence or intelligence suggests that he carry a weapon unlawfully. Again Campbell’s research found this to be a proximate risk indicator.
<b>Firearm</b>	Intelligence or conviction information exists to suggest that the individual has used, and may use or possess a firearm or imitation for the purpose of committing crime.
<b>Drugs</b>	This describes a very loose association with controlled drugs.
<b>Self harm – other than suicide</b>	The individual may cause harm to themselves.
<b>Suicidal</b>	Previous history or threats indicate that the individual may make a determined effort to commit suicide.

<b>Mental health</b>	The subject is known to suffer from a mental condition or order.
----------------------	--

Table 4: Case control variables

The results were then tested for significance using a series of T-tests for independent samples and Cohen's D (standardized mean difference) to measure effect size. The results obtained were largely not statistically significant and the effect sizes small. However there were some significant exceptions to this.

## Results

This chapter is divided into four sections. It commences with a descriptive analysis of the cases in Thames Valley which describe the effectiveness of the SPECSS+ and DASH process in predicting murder and serious assault. There is then a comparison between a similar sample of cases in Hampshire in order to answer the same question. It then looks specifically at the false positives in the period. Finally the results of the case control sample are presented.

### **Descriptive analysis of the Thames Valley cases**

One hundred and eighteen cases of murder, attempted murder, manslaughter and grievous bodily harm with intent (section 18 Offences against the Person Act 1861) have been identified between 1 January 2007 and 1 January 2010 in the Thames Valley Police area. This group includes both men and women and reflects the broad definition of domestic violence and therefore includes all family intimate relationships except for cases where the victim is under 18.

<b>Offence</b>	<b>Number from 2007/2009</b>	<b>% of sample</b>
Murder	13	11
Attempted murder	16	13.5
Manslaughter	1	0.8
GBH with intent	88	74.5
<b>Total</b>	<b>118</b>	<b>100</b>

Table 5: Domestic Violence offences in Thames Valley

Table 5 shows the cases which include 13 (11%) murders, 16 (13.5%) attempted murders, one (0.8%) manslaughter and 88 (74.5%) grievous bodily harm with intent. Fifty one 51 (43%) victims were male and 67 (57%) female

and the average age of the victim was 37 and the average age of the offender was 36. Table 6 below shows that the types of offences varied greatly between the male and female victims. As a consequence, for women the rate of death per attacks was 1 in 6 but for men it was 1 in 25. While some of this may be explained by the fact that men are on average stronger and larger than women there is also research which shows that men are more likely to use a gun or a knife (Browne 1987).

	<b>Murder</b>	<b>Attempted murder</b>	<b>Manslaughter</b>	<b>GBH with intent</b>	<b>All offences</b>
Female victims	11 (85%)	15 (94%)	1	40 (45%)	67 (57%)
Male victims	2 (15%)	1 (6%)	0	48 (55%)	51 (43%)
<b>Total</b>	<b>13 (100%)</b>	<b>16 (100%)</b>	<b>1</b>	<b>88 (100%)</b>	<b>118 (100%)</b>

Table 6: Gender and outcomes

In 51 (43%) cases the relationship was recorded as spouse or cohabitee (or ex), in 52 (44%) cases as lover, boyfriend or girlfriend (or ex), in 13 (11%) cases parent or family and in two (2%) cases civil partner (or ex).

There was prior recorded contact with police for 53 (45%) cases and there was little difference between male and female victims, 46% and 42% respectively. Table 7 shows that in 16 of the 53 cases (30%) the most recent previous contact was a non-crime domestic incident and of the remaining cases 16 were actual bodily harm (30%). In 87% of cases the previous offences were committed by the same offender.

<b>Victim's most recent prior domestic incident contact with the police</b>	<b>Number</b>	<b>% of sample</b>
Non-crime domestic incident	16	30.2
Actual Bodily Harm	16	30.2
Assault without injury	7	13.2
Criminal Damage	4	7.5
Threats to kill	4	7.5
Harassment	3	5.7
Racist incident non-recordable crime	1	1.9
GBH with intent	1	1.9
Arson with intent to endanger life	1	1.9
<b>Total</b>	<b>53</b>	<b>100</b>

Table 7: Most recent prior victim contact with the police for domestic incidents

Where there had been prior contact nearly half of the cases involved only one prior contact for a domestic incident with the police (table 8).

<b>Total number of contacts</b>	<b>Number of victims</b>	<b>Cumulative %</b>
20+	3	2
10-19	5	7
2-9	20	24
1	25	45
Sub total contacts	53	45
No prior domestic incident contact with police	65	55
<b>Total</b>	<b>118</b>	<b>100</b>

Table 8: Prior incident contacts between victim and police

In respect of the risk assessments 21 cases (40%) with prior police contact had been assessed as standard risk, 21 (40%) assessed as medium risk, five as high risk (9%) and there are six cases (11%) where the assessment is unknown (table 9). Only one of the thirteen murders had been assessed as medium risk with the remaining assessed as standard risk. Not one of the



murder cases with prior contact had been assessed as high risk. In respect of seven other murders there was no prior contact. This initial analysis suggests a high number of false negatives for the risk assessment tool as it was applied in Thames Valley. In respect of murder the false negative rate is 100% and in the case of non-lethal assault the false negative rate is 87%. The combined false negative rate is 90%.

<b>Risk assessment for victim on most recent prior domestic incident</b>	<b>Number</b>	<b>%</b>	<b>Valid %</b>
High	5	4.2	9.4
Medium	21	17.8	39.6
Standard	21	17.8	39.6
Prior contact but risk assessment not known or not recorded	6	5.1	11.3
N/A as no prior domestic incident contact with police	65	55.1	-
<b>Total</b>	<b>118</b>	<b>100</b>	<b>100</b>

Table 9: Prior risk assessment

### **Hampshire comparator**

The number of serious assaults and murders where there was no prior contact with the police was unexpectedly high. Similarly the level of false negatives was high and so there was always a chance that the risk assessment tools had been implemented in Thames Valley with some type of bias. Data was therefore obtained from Hampshire Constabulary in order to triangulate the results and to exclude this possibility. The force was selected primarily because of the easy availability of data but it is also a large non-Metropolitan force with a mix of urban and rural areas and in the “most similar

family” comparator which the Home Office use. Exactly the same crime categories were used and the same time period.

<b>Offence</b>	<b>Thames Valley</b>	<b>Hampshire</b>
Murder	13 (11%)	17 (9%)
Attempted murder	16 (14%)	9 (5%)
Manslaughter	1 (1%)	0 (0%)
GBH with intent	88 (74%)	154 (86%)
<b>Total</b>	<b>118 (100%)</b>	<b>180 (100%)</b>

Table 10: Domestic violence offences from 2007/09

The variation in the level of GBH with intent was noteworthy (table 10) and while it could represent different levels of offending may also represent different recording practices. Table 11 shows the comparative intensity of all types of most serious violence – those offences with a domestic violence flag and those without. It is clear that in Thames Valley there was either a lot less serious violence or a different approach was being taken to the recording of the serious offences – most probably a mix of the two. This however does not undermine the analysis on prior contact and risk assessment. It just means that the parameters for the numerator are probably slightly broader in Hampshire.

	<b>Most serious violence per 1000 population</b>	<b>Rank</b>
<b>Thames Valley</b>	0.374	5
<b>Hampshire</b>	0.488	15

Table 11: Recording of violent crime with injury offences in England and Wales 2009/10

The key area however was the high level of no prior contact as DASH and SPECSS+ are completely reliant on this as a risk pool. Table 12 below shows the comparison - 48% in Hampshire compared with 55% in Thames Valley. In both forces about half of the cases which result in serious harm the victims had no prior contact with the police. This clearly illustrates the weakness of any assessment tool which assumes prior contact.

The breakdown of risk assessments however reveals quite considerable differences in the two forces. Hampshire Constabulary has amended the DASH model to include an additional level of “very high”, “low” and “silver”. Overall table 12 shows that in Thames Valley there are fewer cases assessed as high risk.

	<b>Thames Valley</b>	<b>Hampshire</b>
Very high	0	21 (11.7%)
High	5 (4.2%)	10 (5.6%)
Medium	21 (17.8%)	11 (6.1%)
Standard	21 (17.8%)	14 (7.8%)
Low	0	1 (0.5%)
Silver	0	1 (0.5%)
Prior contact but risk assessment unknown	6 (5.1%)	35 (19.4%)
No prior contact	65 (55.1%)	87 (48.3%)
<b>Total</b>	<b>118 (100%)</b>	<b>180 (100%)</b>

Table 12: Prior contact and risk assessments in two forces

The percentage of false negatives also varies between Thames Valley and Hampshire (table 13). A false negative has been calculated by dividing the number of cases that were not assessed as high risk, and therefore not considered by the MARAC, by the total number of cases where there was both prior contact and a risk assessment completed.<sup>1</sup> This higher level of false negatives is linked to the comparatively lower level of cases assessed as high risk and suggests that while the model is clearly problematic in the accuracy of its prediction in both forces its use Thames Valley is potentially leading to even greater levels of inaccuracy.

	<b>Thames Valley</b>	<b>Hampshire</b>
<b>False negatives</b>	90%	63%

Table 13: Risk assessment in two forces

The comparison between murders in Thames Valley and Hampshire is interesting. While there were seventeen murders as opposed to thirteen in Thames Valley, three had been previously assessed as high risk in Hampshire whereas none had in Thames Valley. In seven murders in Thames Valley there was no prior contact and this was the case in eleven in Hampshire. Therefore if DASH is really focused on lethal attacks it is going to miss 54% of murders in Thames Valley and 65% of murders in Hampshire.

---

<sup>1</sup> Hampshire Constabulary use a category of “very high risk” for referral to MARAC and this is what has been used in the calculation.

## **Exploration of the false positives**

In order to look at the level of false positives a table was produced with all the cases that had been assessed as high risk between 1 January 2007 and 31 December 2009. In total there were 2721 cases where a high risk assessment was made in respect of 1745 different victims. In this same period five of these high risk cases resulted in domestic murder or serious assault which would mean a false positive rate of 99%. However such a conclusion is arguably very misleading. It is clear from the data that there are many re-assaults within this group and it could be argued that this undermines the conclusion of false positive. It is also the case that cases that were originally identified as high risk were not ultimately harmed because they were subject to MARAC safety planning and protection. Again, these would not be appropriately called false positives.

Table 14 shows significant differences across the Local Police Area in terms of incidence per 1000 population. The highest is in Reading at 3.2 per 1000 population to 0.28 per 1000 population in South Oxfordshire. There appears to be a greater concentration of high risk cases in the urban areas. However there are differences in the levels of risk assessment which suggest that are differences in implementation. For example Slough has 1.21 high risk assessments per 1000 population whereas Reading has 3.20. This variation across local units is similar to the early evaluation of SPECSS+ in the Metropolitan area and West Yorkshire Police (Humphreys et al, 2005).

<b>Local Police Area</b>	<b>High risk assessments</b>	<b>High risk assessments per 1000 population</b>
Aylesbury Vale	200	1.15
Bracknell Forest	63	0.54
Cherwell	129	0.92
Chiltern	99	1.09
Milton Keynes	487	2.06
Oxford	198	1.32
Reading	486	3.20
Slough	156	1.21
South Buckinghamshire	73	1.09
South Oxfordshire	36	0.28
Vale of the White Horse	38	0.32
West Berkshire	219	1.43
Wes Oxfordshire	54	0.53
Windsor and Maidenhead	52	0.37
Wokingham	148	0.91
Wycombe	283	1.72
<b>Total</b>	<b>2721</b>	

Table 14: Risk assessments by Local Police Area

This study has not looked at the effectiveness of the MARACs across the Thames Valley which will have considered all these cases over the three years. While the 2721 cases only refer to 1745 victims it is clear that there is a considerable workload for the partnerships. Moreover, given current concerns about police availability and bureaucracy it is important to note that such high numbers of cases require many hours to be spent in meetings sharing information and managing cases.

## Case Control Samples

### *Male block*

Table 15 shows comparisons between the male offenders in the serious domestic assaults studied and the case control sample drawn from the broader population of those arrested for violence. There are some striking similarities between the groups but also some differences some of which are significant. A series of T-tests for independent samples were conducted to test whether any of the differences in respect of prior criminal history were significant. In respect of five of the eight categories the difference was significant.

	<b>Offenders in the Thames Valley cases (<i>standard deviation</i>)</b>	<b>Offenders in the case control sample (<i>standard deviation</i>)</b>	<b>Effect size Cohen's D</b>	<b>P value of T test</b>
<b>Average age at first conviction for violence</b>	28.51 (12.86)	22.62 (8.20)	0.54	0.002
<b>Average age at first conviction</b>	24.52 (14.89)	20.29 (7.43)	0.48	0.03
<b>Average number of prior convictions and cautions for violence</b>	1.21(1.45)	2.77 (4.59)	-0.45	0.00
<b>Average age at first arrest for violence</b>	30.49 (14.32)	24.47 (9.99)	0.35	0.002
<b>Average age at first arrest</b>	25.69 (15.15)	22.09 (9.79)	0.28	0.07
<b>Average number of prior arrests for violence</b>	2.48 (3.32)	3.80 (6.53)	-0.25	0.04

<b>Average number of prior convictions and cautions</b>	6.78 (8.46)	6.69 (9.09)	0.01	0.94
<b>Average number of prior arrests</b>	9.32 (12.64)	9.29 (12.56)	0.002	0.048

Table 15: Male case control – significant variables

The average age at first conviction for violence in the serious domestic assault cases is 28.51 (SD=12.86) which is significantly higher than the average age of 22.62 (SD=8.20) in the control sample,  $t(78.90) = -3.18$ ,  $p=0.002$ ,  $d=0.54$ . The average age at first conviction for the serious domestic assault cases is 24.52 (SD=14.89) which is significantly higher than the average age of 20.29 (SD=7.43) in the control sample,  $t(83.47) = -2.19$ ,  $p=0.03$ ,  $d=0.48$ . Lastly, the average age at first arrest for violence in the serious domestic assault cases is 30.49 (SD=14.32) which is also significantly higher than the average age of 24.47 (SD=9.99) in the control sample,  $t(98.51) = -3.13$ ,  $p=0.002$ ,  $d=0.35$ . All three have a small to medium effect size.

The average age at first arrest in the serious domestic assault cases was 25.69 (SD = 15.15) but this average conceals the fact that of those with no convictions the average age is 44.7 and a few offenders with many convictions were arrested in their teens thus reducing the average. The average age of arrest in the case control sample is only 22.09 (SD = 9.79). This difference is not significant,  $t(101.12) = -1.84$ ,  $p=0.07$ ,  $d=0.28$ . There is a small effect size.



The first three findings show that the average age of onset of criminal career is later for those offenders who commit serious domestic assault than for those who commit all violence. Again, the evidence does not support the commonly held view that serious domestic assaults result from escalation over the years but suggest that these attacks are often much less predictable and perpetrated by those with less of a criminal history and a criminal history with a much later onset. In the cases with no prior violence, the serious assault came “out of the blue”.

In the serious domestic assault cases there was an average of 1.21 (SD = 1.45) convictions or cautions for violence compared with the case control where there was an average of 2.77 (SD = 4.59). This difference is significant,  $t(193.30) = 3.75$ ,  $p = 0.00$ ,  $d = -0.45$ . This constitutes a medium effect size. Similarly, in respect of prior arrests for violence in the serious domestic assault cases there was an average of 2.48 (SD = 3.32) compared to an average of 3.80 (SD = 6.53) in the case control sample. This difference is also significant,  $t(216.92) = 1.99$ ,  $p = 0.04$ ,  $d = -0.25$ . However the effect size is small.

These two significant differences are noteworthy because the commonly held view is that the most serious domestic assaults are a result of escalating harm and violence. In sharp contrast, the evidence shows that offenders in the case control have significantly more arrests and convictions for violence than the offenders who committed domestic murder and serious assault.

Therefore the evidence contradicts the commonly held view and the hypothesis upon which much of the risk assessment tools are based.

However in respect of total convictions and cautions there was an average of 6.78 (SD = 8.46) in the serious domestic assault cases compared with the control sample where the average number of convictions or cautions was 6.69 (SD = 9.09). This difference is not significant,  $t(153.57)=-0.07$ ,  $p=0.94$ ,  $d=0.01$ . The effect size is small.

In the serious domestic assault cases the average number of arrests was 9.32 (SD = 12.64) but this is very much skewed by two offenders who had 65 and 67 arrests and therefore pushed the average up considerably. In fact 74% of offenders had been arrested 10 times or less. Similarly in the case control sample the average number of arrests was 9.29 (SD = 12.56) and again the average is skewed by two offenders who had been arrested 57 and 60 times respectively. The difference is not significant,  $t(143.31)=1.96$ ,  $p=0.048$ ,  $d=0.002$  and the effect size is miniscule.

	<b>Offenders in the Thames Valley cases</b>	<b>Offenders in the case control sample</b>	<b>Relative risk ratio</b>
<b>Unemployed</b>	41.7%	38%	1.09
<b>PNC warning for weapons</b>	33.3%	23.3%	1.43
<b>PNC warning for firearms</b>	6.9%	4.6%	1.5
<b>Drugs, describes a loose association</b>	13.8%	16.6%	0.83

<b>with controlled drugs</b>			
<b>Self harm – other than suicide</b>	12.5%	7.3%	1.71
<b>Suicidal</b>	19.4%	6%	3.23
<b>Mental health</b>	16.6%	9.3%	1.78

Table 16: Male case control – relative risk ratios

Table 16 shows the differences in respect of the warning marks on PNC between the two groups. In respect of most categories the offenders in the serious domestic assault cases were more likely to have the warning marker than the case control. Thirty three per cent had a warning marker for weapons compared with 23% in the case control sample. Six per cent had a warning marker for firearm compare with four per cent in the case control sample. There were smaller differences in respect of markers for drugs and the levels of employment varied slightly.

However the relative risk ratio is significantly higher in respect of suicide, mental health and self harm. In the serious domestic assault cases the offenders are more than three times more likely to be marked as suicidal and nearly twice as likely to have mental health issues or a marker for self harm. The evidence suggests that these are predictors of serious domestic assault.

### ***Female block***

Tables 17 and 18 show the comparisons between the offenders in the serious domestic assaults and the case control sample drawn from the broader

population of those arrested for violence in respect of female offenders. A t-test for independent samples was conducted to test whether any of the differences in respect of prior criminal history were significant. In none of the cases was the difference significant and the effect sizes were small or less than small.

However in respect of the PNC warning for weapons there is a considerable difference between the female offenders in the serious domestic assault cases and the case control. In the latter group nearly one third of women (32.6%) had the marker compared with 7% in the case control. The relative risk ratio is 4.65 which means that those offenders who are known have used weapons are nearly five times more likely to commit domestic murder and serious assault than those who have not.

	<b>Offenders in the Thames Valley cases (standard deviation)</b>	<b>Offenders in the case control sample (standard deviation)</b>	<b>Effect size Cohen's D</b>	<b>P value of T test</b>
<b>Average number of prior arrests</b>	4.41 (8.98)	4.75 (12.16)	-0.03	0.85
<b>Average number of prior arrests for violence</b>	1.15 (1.81)	2.11 (5.88)	-0.22	0.15
<b>Average number of prior convictions and cautions</b>	3.17 (5.52)	3.36 (8.60)	-0.02	0.88
<b>Average number of prior convictions and cautions for violence</b>	0.80 (1.30)	1.64 (4.70)	-0.24	0.11
<b>Average age at first arrest</b>	29.24 (13.87)	28.37 (10.91)	0.06	0.71

<b>Average age at first arrest for violence</b>	33.21 (14.04)	30.70 (10.95)	-0.32	0.15
<b>Average age at first conviction</b>	24.03 (11.21)	27.64 (11.26)	0.19	0.31
<b>Average at first conviction for violence</b>	30.14 (14.29)	30.09 (11.57)	0.003	1.00

Table:17 Female case control – significant variables

In the Thames Valley cases the average number of arrests was 4.41 (SD = 8.98) but this is skewed by one offender who had 52 arrests. Similarly in the case control sample the average number of arrests was 4.75 (SD = 12.16) and again the average is skewed by two offenders who had been arrested 78 and 80 times respectively. The difference is not significant,  $t(117.96)=0.19$ ,  $p=0.85$ ,  $d=-0.03$ . The effect size is minimal.

In respect of prior arrests for violence in the serious domestic assault cases there was an average of 1.15 (SD = 1.81) but the range was not as great for all arrests with the highest number of arrests only 8. In the case control sample the average was 2.11 (SD = 5.88) with the highest number of arrests being 50. The difference is not significant,  $t(124.48)=1.44$ ,  $p=0.15$ ,  $d=-0.22$ . The effect size is small.

In the serious domestic assault cases there was an average of 3.17 (SD = 5.52) convictions and cautions compared with the control sample where the average number of convictions or cautions was 3.36 (SD = 8.60). This difference is not significant,  $t(128.17)=0.15$ ,  $p=0.88$ ,  $d=-0.02$ . The effect size is minimal. Similarly in respect of convictions or cautions for violence the cases had an average of 0.80 (SD = 1.30) compared with the control sample

where there was an average of 1.64 (SD = 4.70). Again, the difference is not significant,  $t(119.53)=1.61, p=0.11, d=-0.24$ . The effect size is small.

None of the differences in ages between the cases and the control group in Table 16 were significant for the female offenders. Only one area, average age at first arrest for violence, is there a small effect size,  $d=-0.32$ . While some of the averages are higher in the control sample rather than the cases the differences are not significant and, unlike the male cases, conclusions cannot be drawn.

	<b>Offenders in the Thames Valley cases</b>	<b>Offenders in the case control sample</b>	<b>Relative risk ratio</b>
<b>Unemployed</b>	54.3%	62%	0.87
<b>PNC warning for weapons</b>	32.6%	7%	4.65
<b>PNC warning for firearms</b>	2.1%	2%	1.05
<b>Drugs, describes a loose association with controlled drugs.</b>	13%	7%	1.85
<b>Self harm – other than suicide</b>	6.5%	13%	0.5
<b>Suicidal</b>	13%	10%	1.3

Table 18: Female case control – relative risk ratios

Table 18 shows that female offenders in the serious domestic assault cases are less likely to be unemployed than the control sample, the opposite of the male offenders. In respect of most categories the offenders in the Thames Valley cases were more likely to have the warning marker than the control sample and the relative risk ratio in respect of a warning marker for weapons is noteworthy as discussed earlier.

There were smaller differences in respect of markers for firearms, drugs, self harm, suicide and mental health. None of the significant differences in exposure that existed in the male offenders in respect of self-harm, suicide and mental health exist in respect of the women. In fact the percentage of those with the self-harm flag is half that in the cases than the control group.

## Discussion



The discussion of the results mirror the research questions and there is a concluding section on the need for national strategy to be evidence-based.

In how many cases did the victim have a prior history of reported domestic violence?

Both SPECSS+ and DASH are predicated on a view of domestic violence escalating over time from more minor offences to more serious offences. The assumption is that because many cases have prior contact with the police such contact can serve as an “early warning system” and an opportunity to intervene.

Studies have found a very wide range of levels of prior contact. One of the first studies in Kansas City (Breedlove et al, 1977) found a level of prior contact of 90% but many subsequent studies have found lower levels. Sherman (1992b) challenged the Kansas City results and the weight of evidence in all the studies reviewed is that there is much lower level of prior contact than has often been asserted.

This study has found that in Thames Valley only 45% of victims of the most serious domestic violence had prior contact with the police. Consequently in 55% of the cases there was no prior opportunity to intervene to prevent escalation. This evidence presents the first serious challenge in the UK to the hypothesis upon which current risk assessment models are based. The level of prior contact has been overestimated.

However it is not just that prior contact has been overestimated but that the assumption of escalation of violence over time is not borne out by the evidence. The case control study found that for male offenders the number of arrests, convictions and cautions for violence was significantly lower for those who committed domestic murder and serious assault than for the pool of violent offenders. While this analysis has not directly examined whether there is a pattern of increasing seriousness of violence prior to the lethal or near lethal assault, the total number of prior violence arrests is at least suggestive of a lack of escalation. For female offenders, there were also lower numbers of arrests and convictions and cautions for violence but the difference was not significant. This evidence presents a second serious challenge to the hypothesis on which current risk assessment models are based. Serious domestic violence offenders have been arrested, cautioned and convicted less often than the broader pool of violent offenders.

How accurate were the risk assessments based upon prior history in terms of false positives and false negatives?

In reviewing the South Wales Police approach to risk assessment and multi-agency work Amanda Robinson argues, “risk assessment in cases of domestic violence can be defined as trying to identify those victims who are most at risk of experiencing violence in the future” (Robinson 2006, p764). In other words, the risk assessment models seek to predict which cases will be subject re-assault including murder. The literature review quoted Paul

Meehl's research from over fifty years ago which argued that non-actuarial approaches will not result in accurate results (Meehl 1954). Neither SPECSS+ nor DASH are actuarial models and this study has shown that these risk assessments based upon prior history produce a significant number of false positives and false negatives. They do not produce accurate results.

The study found that within the serious cases reviewed that there had been a false negative rate of 90%<sup>2</sup>. The data from Hampshire showed that there was a 63% false negative rate. While there may be some inconsistent implementation in Thames Valley, the rates across all forces show a very high level of false negatives.

The study found that 2721 domestic violence cases had been initially assessed as high risk. There were many repeat cases in this group and the total number of victims was 1745. In terms of false positives, five victims out of 1745 were correctly assessed as high risk and were part of the initial numerator group of cases – a false positive rate of 99%. It will be argued by practitioners that false positives are indeed examples of where the MARAC process has succeeded. This study has not included an assessment of the MARAC process which is largely about information sharing and safety planning on high risk cases so it is not possible to comment on effectiveness of this process and the accuracy of the claims. While Robinson's study of MARACs in South Wales (2006) did find that those high risk cases which

---

<sup>2</sup> A false negative has been calculated by dividing the number of cases that were not assessed as high risk, and therefore not considered by the MARAC, by the total number of cases where there was both prior contact and a risk assessment completed.

were referred to MARAC did not have any contact in the subsequent six months in 30% of cases it is hard to make any conclusions as there was no comparison with high risk cases which were not referred to the MARAC.

Overall the results suggest that DASH and its predecessor SPECSS+ do not accurately assess risk. The literature review identified that there has been little evaluation of SPECSS+ and DASH but these results do corroborate the few studies that have taken place. Keri Nixon, in her unpublished doctoral thesis, commented on the use of SPECSS+ in Merseyside Police, “empirical tests of the SPECSS based risk assessment shows it to be unreliable” (Nixon 2009 p134). Robinson’s evaluation of the South Wales Police model looked at 146 high risk cases and checked for re-assault six months later. She found that only one risk factor, injury, significantly predicted repeat abuse (Robinson, 2006). The findings also corroborate the high rate of false positives found in the use of structured professional judgement models in use in the United States (Bennett et al 2000 and Heckert and Gondolf 2004).

There are many possible reasons for these inaccurate results which will be considered in this discussion. Some are in respect of the model itself – is it intrinsically flawed or just poorly implemented? More fundamentally – is the wrong research question being asked?

It was clear from the earlier analysis of SPECSS+ and DASH that the methodology was weak. The risk factors were identified from 30 cases of murder in London, there was no comparison with a broader risk pool and it is

not entirely clear why some factors were selected and others not. But Jacqueline Campbell's Danger Assessment model is based on research evidence and has been evaluated for reliability and validity. Yet it still had a high level of false positives and some false negatives (Bennett et al 2000).

Michael Johnson's work on a typology of domestic violence (2008) clearly sets out the different types of domestic violence but DASH is a blunt tool which assumes that there is one unitary phenomena. Surely any tool should take into account the fact that within the ACPO definition there are many different types of violence? There are offenders of both genders, there are same sex relationships and there are parent child and sibling relationships. Some may be about intimate terrorism but others will be about situational couple violence. Arguably the tool was developed with attacks on female sexual partners in mind and based upon an understanding of intimate terrorism rather than an appreciation of the breadth of domestic violence encompassed by the ACPO definition.

The application of a single risk assessment tool to such a wide range of violence does not appear to be justified in terms of the empirical knowledge available. There has been a tendency to put cases with any risk into the same ACPO definition of domestic violence – "just in case". That is in itself problematic but this is compounded by the use of a single risk assessment for all cases. Apart from anything else this is potentially a significant waste of resources!

However, even if the model is sound the inaccurate results may be caused by poor implementation in Thames Valley. Originally the risk assessment in SPECSS+ was completed by specialist staff in the Domestic Abuse Units. With the advent of DASH this is now completed by front line officers. This raises yet another question about the risk assessment process. How appropriate is it to rely upon the professional judgement of generalist operational officers rather than domestic violence professionals?

It is very clear when talking to operational officers that they do not feel comfortable completing the risk assessment. Questions of a very personal nature are asked and officers feel very embarrassed to ask them (Macvean and Ridley 2007). In Thames Valley the risk assessment needs to be completed for all cases of domestic incidents as well as crimes. This means that in 2009/10 18,386 risk assessments were completed for incidents and 12,490 risk assessments were completed for crimes. Some of these incidents may be a call from a third party such as a neighbour to a noisy argument. The completion of risk assessments in such circumstances is resource intensive and arguably an unjustified breach of privacy.

Kropp (2004) argues that in the United States despite widespread use of risk assessments there is an absence of minimum qualifications for those who complete assessments, standardised training, monitoring or best practice. The same would largely apply in Thames Valley – while there was some limited training this could not be described as adequately equipping staff. That said, even if the training had been a month long, which would have been

wholly inappropriate, it is questionable whether generalist police officers are then experts in the risks associated with domestic violence.

Trujillo and Ross's study (2008) of risk assessment in use in Australia shows that while officers may use a formal risk assessment that their views about a victim's fear and the frequency with which they have been called to an address impact on their assessment of risk. The study also showed that there was significant variation in the use of high risk assessments across the Local Police Areas which may be caused by different management cultures rather than any differences in respect of risk.

However there is a broader explanation which needs to be addressed – that this study has been asking the wrong question. The developers of SPECSS+ and DASH consistently argue that the tools are not about predicting murder but about deciding which cases are suitable for proactive intervention. The accompanying literature on DASH frequently repeats the assertion that it is not a predictive model its purpose is preventative (Richards 2008, p108). Does that interpretation undermine the logical premise that prevention depends on reliable prediction?

At a very simple level the police role is to prevent crime rather than quantify the likelihood of it occurring and so prevention may be a more appropriate aim than prediction. But surely any targeted prevention work must be based on prediction of harm which is sound? The risk assessment tool is being used to try to identify those cases where the risk is seen as the highest and therefore where preventative effort needs to be focused. In contrast, Jacqueline

Campbell was quite clear that her model was about predicting domestic violence – in fact she was even more specific, her model was about predicting domestic murder.

The rejection of the word “prediction” it is probably meant in a very narrow sense – that is to say the model does not aim to predict where harm will occur. But a more appropriate comparison is with a weather forecast for rain which is a prediction based on all that is known but does not mean does not mean that it will actually rain. And the parallel does not end there – as forecasts have improved as more information has been used so domestic violence prediction should improve as more information is available.

In their review of the approach to domestic violence in Thames Valley Police, Macvean and Ridley (2007) argued that “risk” should only be used in respect of mathematical actuarial models and therefore DASH does not measure risk but it is about a potential threat of harm. They suggested that because there is so much missing or incorrect information the assessments end up measuring uncertainty rather than risk. “It could be argued that what is referred to as the risk assessment in domestic violence is in fact a tool to assist in measuring the potential threat of harm to the victim and that the risk management plan is a plan that provides a set of actions to assist in reducing that potential threat of harm” (Macvean and Ridley 2007, p45). This argument is similar to Malcolm Gladwell’s argument in “Blink” that the information that is fed into risk assessments is too limited for effective risk management.



So if determining risk needs more information and the ability to calculate risk mathematically then what is DASH? Is it in fact misnamed as a risk assessment model? Macvean and Ridley are using threat to refer to some future harm. But are they saying anything more profound than exchanging the word 'risk' for the word 'threat'- that DASH should be described as a threat assessment model? And what would that mean if described in that way?

It is often the case that risk and threat are used interchangeably in the police service but they are not the same thing. Risk includes some element of likelihood – it is about severity of threat but also about the likelihood of that threat. In other words there always needs to be a denominator and a numerator when assessing risk. Inherent in the concept of a risk factor is the idea that its presence increases the likelihood of an outcome compared with a risk pool without that risk factor. To ignore this is to fall prey to the hindsight fallacy. This basic principle was somehow overlooked in the development of DASH and SPECSS+.

The brain wants to make sense of the past, but this unconscious determinism undermines our ability to learn from the past. As Fischhoff said in his paper on the influence of outcome knowledge on judgement, 'In hindsight we systematically underestimate the surprises which the past held and holds for us' (Fischhoff 2003, p311).

There have been many risk assessment processes introduced into policing in the United Kingdom in the last ten years and many will have the same

weakness. They will have identified so called risk factors with no regard for their presence in the wider risk pool and have fallen for the hindsight fallacy. Interestingly, there is evidence that the tide is beginning to turn against this approach, not necessarily because of the evidence that the risk assessments are not accurate but because of cost: they have become bureaucratic comfort blankets that are no longer seen as affordable. The recent 2010 ACPO submission to the Home Office consultation paper 'Policing in the 21<sup>st</sup> Century' stated,

“Good practice has been associated with long forms, the gathering of detailed information, pre-formatted questions and formulaic risk assessment processes. The impact is further exaggerated by the blanket application of these policies with very little room for officers to exercise judgement about the specific harm and threat they are encountering.” (p20)

The evidence of false negatives and false positives found in this study presents the third serious challenge to the current approach to risk assessment. While it may be that the implementation in Thames Valley has been problematic, it is more likely that the weak methodology of the development of these risk assessments lies at the root of this problem. Risk factors have been identified because of their presence in the numerator but they have never been compared with their presence in a wider risk pool or denominator. The concept of risk depends upon the presence of a denominator but these risk assessments have overlooked that fact and fallen into the hindsight fallacy.

Might the use of a case control sample produce more accurate risk indicators than analysis based on the numerator alone?

The study of the Thames Valley cases revealed that the use of SPECSS+ and DASH was not producing accurate risk assessments. The level of false negatives and false positives was high. The development of these risk assessments had been based upon the analysis of the numerator alone. Therefore the third research question was to discover if a case control sample would lead to the identification of more accurate risk assessments. Would it be possible to identify risk factors within the numerator group which were more prevalent than in the denominator group?

Given that only 45% of cases had prior police contact, a case control design using prior victims as a risk pool would exclude many cases. Most risk assessments in the criminal justice field have been about assessing the risk of re-offending. The approach to domestic violence has emerged from a medical harm reduction approach and this has meant that the focus has been on assessing the potential for harm to the victim. While there is evidence that this can empower the victim it may not be the best way to protect women. This study has therefore focused on the offender.

After careful consideration a risk pool of those who had been arrested for violence in Thames Valley in the same three year period was chosen. A random sample of the risk pool was identified and the numerator cases and the control sample were compared in two separate male and female blocks.

While there are some differences between the cases and the control sample most of the statistically significant differences are counter to the central hypothesis of escalating violence. For example, the male offenders in the case control were arrested more often for violence not less than those who went on to commit domestic murder and serious assault.

The results in respect of the presence of PNC markers for suicidal, mental health and self harm in respect of the male offenders are significant. Those who commit serious domestic assault are nearly three times more likely to be suicidal than other violent offenders. They are also nearly twice as likely to have mental health problems. This finding on the prevalence of mental health issues replicates the review in South Wales Police (Robinson 2006) and works in the US where those who murder their partners were more than four times likely to have mental health problems than the wider pool of murderers (Zawitz et al 1994).

This case control sample was compared with the 120 offenders in the Thames Valley cases which covered the breath of domestic violence offending types. It might be that if the group were divided into more specialised blocks, such as those male offenders who may be described as intimate terrorism cases there may be different, more significant, findings.

The literature review described the problems inherent in trying to predict those events for which there is a low base rate. Murder, and even serious assault,

has a low base rate but these risk models are based upon the approach to recidivism in general where the base rate is much higher.

Any risk assessment completed by police officers in respect of the risk faced by a third party is bound to have significant information gaps. As Macvean and Ridley argue, it is as assessment of uncertainty (2007). However, even if we knew all the information it would always be very hard to predict occurrences with such a low base rate. There is a clear relationship between base rate and predictability. The base rate of recidivism is high and therefore risk indicators are practical. But the base rates of domestic murder and serious assault are not high. Many cases appear to come out of nowhere and they are just not as obvious with foresight as the selected cases where there was prior contact suggest.

If a structured professional judgement model is not providing an accurate forecast might a sophisticated actuarial model involving non-linear data mining across huge data sets identify those elusive risk identifiers? If information wasn't limited would it be possible to predict accurately? Maybe, but given the diversity of situations and the dynamic nature of the domestic violence it is hard to see how a model will be found that applies to all domestic violence.

The study has raised several significant challenges for the current approach to domestic violence. Overall there is the need to ensure that policy is based upon evidence rather than theory alone. This is a significant challenge to

ACPO itself. How can a model be endorsed by ACPO without a sound evidence base and lacking any form of peer review and minimal evaluation? On the basis of this model the resources of domestic violence units have been rationed and reassurance has been given to victims. In this way there is significant potential to undermine the legitimacy of policing.

ACPO has made a leap of faith based on the work of a few analysts and police officers working in isolation from the wider research community. The National Policing Improvement Agency has published papers with no peer review. One of the most striking findings of this study has been the significant difference between the approach to evidence based practice in the United Kingdom compared with the United States. In the United States there has been a substantial piece of research with a large National Institute of Justice grant which has led to an evaluated tool.

These observations do not just apply to the approach to domestic violence but to many areas of strategy and policy development. In these days of austerity the need to focus precious resources on evidence based practice is more important than ever. The Coalition Government has a reforming agenda and yet there is no visible discussion of the relationship between policing and the universities and research community.

## Conclusions

The prevention of crime, and in particular the prevention of violence perpetrated on the vulnerable, is central to the mission of policing. The natural reaction to a horrific domestic murder is to reflect on what might have been done to prevent it. In 2003 and 2004 the tragic domestic murder of several victims in Thames Valley had focussed attention on police practice and the need to improve the response. The force was one of the first to adopt risk assessment in an attempt to identify those cases which were most likely to lead to tragedy.

This study was based around three research questions arising out of the use of that risk assessment in Thames Valley. The use of the risk assessment is based upon the premise that there are previous contacts with the police which provide an opportunity to assess the risk to the victim and intervene with other partners to prevent future harm. The literature review had shown that in international studies the levels of prior contact varied from 3% to 90%. The first research question therefore asked how often in case of domestic murder or other serious assault did the victim have prior contact with the police? One hundred and eighteen violent crimes which occurred between 2007 and 2009 were studied and in only 45% of cases was there any recorded prior contact. Therefore in more than half the cases studied there was no opportunity to risk assess and intervene.

Risk assessment models have been developed in many fields of criminology and have more recently been developed in respect of domestic violence. The literature review found that compared with other fields there have been few



evaluations of risk assessment models and “despite improvements in this field the prediction of violence remains an inexact science” (Scott et al 2008). There has been little evaluation of the accuracy of SPECSS+ and DASH and the study identified some concerns with the methodology adopted to develop them. Significantly the risk identifiers had been identified by initially reviewing 30 murders but there was no comparison with domestic violence cases which did not escalate to murder – there was no case control.

This fundamental flaw means that the developers of these risk assessment tools are not able to say that the risk is greater if the risk identifier is present because there is simply no basis for that claim. There has been no comparison with other cases. Jacqueline Campbell’s statistical approach does use a case control study but arguably does not establish accurate predictors as she excluded all the cases without prior contact.

The study found an 80% false negative rate with only five out of 118 cases being assessed as high risk – clearly in 65 cases there had been no prior contact at all. In terms of false positives there were 2721 high risk assessments made in the three year period concerning 1745 victims. While it might be tempting to say that there were therefore 1740 false positives (99%) such a conclusion would disregard the significant amount of multi-agency work that will have been undertaken to reduce the risk. However there will be significant levels of false positives where there has been no effective harm reduction activity carried out.

Given the lack of accuracy in the use of SPECSS+ and DASH the third research question then focused on the possibility that denominator based risk indicators might provide more accurate risk assessments. The numerator cases were analysed to identify an evidence based risk pool that captured more cases than prior victimisation (45%). A case control study was therefore completed which compared the 120 offenders in the original cases of domestic murder and serious assault with a case control sample of violent offenders (59%). This study showed that where there were differences between the 120 offenders and the case control the evidence contradicted the theory underpinning risk assessment because the case control offenders were more criminogenic.

The study has found that the risk assessment used in Thames Valley is not accurately predicting risk. However it has identified the possibility that a risk assessment that focussed on male offenders with a violent history who are either suicidal or self harm or have mental health problems might produce a more accurate alternative if the focus is to remain on predicting domestic murder and serious assault.

However domestic murder and serious assault have very low base rates in Thames Valley – 17.8 per million per year and this make prediction difficult. The study has discovered a significant level of repeat victimisation – 36% of all the high risk cases were repeat victims. The base rate is much higher for re-assault and this makes prediction more possible. This is where the focus of attention could be. The risk assessment models were developed based

upon studies of lethal attacks but may well be more accurate in predicting re-assault if they were to focus on risk identifiers for re-assault. While this may seem controversial it is better to promise more modest levels of prevention than promise the earth and fail to deliver.

The opportunity which technology affords to interrogate high volumes of data could be used to identify those cases where re-assault is much more likely than in other cases. While a small sample case control is useful to avoid the hindsight fallacy, a population-based forecasting tool would be more precise and comprehensive.

In its desire to do something to stop the tragic loss of life to domestic violence ACPO put its faith in the use of a risk assessment model based upon analysis of domestic murders in order to prevent domestic murders. This study has falsified the hypothesis upon which the approach is based. There is an opportunity to learn from this and to develop a new approach which is based upon evidence rather than theory.

## References

- ACPO (2010) 'ACPO response to Policing in the 21st Century' unpublished response to Home Office report
- Ariel, B and Farrington, D (2010) 'Randomized Block Designs' in the Handbook of Quantitative Criminology, New York: Springer Science and Business Media
- Bachman, R and Schutt, R (2003), *The Practice of Research in Criminology and Criminal Justice*, California: Sage Publications
- Bennett, L, Goodman, L, and Dutton, M (2000) 'Risk assessment among batterers arrested for domestic assault: the salience of psychological abuse', *Violence against Women*, vol 6, 1190-1203
- Breedlove, R, Kentish, J, Sandker, D and Sawtell, R (1977) 'Domestic Violence and the Police: Kansas City' in *Domestic Violence and the Police: Studies in Detroit and Kansas City*, Washington DC: Police Foundation
- Brown, A (1987), *When battered women kill*, New York: The Free Press
- Canadian Bureau for Justice Studies (1992) 'Homicide in Canada' in *Juristat Service Bulletin*, vol 14, no 4
- Campbell, J, Webster, D, Koziol-McLain, J, Block, C, Campbell, D, Curry, M, Gary, F, Glass, N, McFarlane, J, Sachs, C, Sharps, P, Ulrich, Y, Wilt, S, Manganello, J, Xu, X, Schollenberger, J, Frye, V and Laughan, K (2003) 'Risk Factors for femicide in Abusive Relationships: Results from a Multisite Case Control Study' *American Journal of Public Health*, vol 93, 1089-1097
- Campbell, J (2005a) 'Commentary on Websdale, Lethality Assessment Approaches: Reflections on their use and ways forward', *Violence against Women*, vol 11, 1206-1213
- Campbell, J (2005b) 'Assessing Dangerousness in domestic violence cases; history, challenges and opportunities', *Criminology and Public Policy*, 4, 653-672
- Dobash, R, Dobash, R, Cavanagh, K and Medina-Ariza, J, (2007) 'Lethal and Nonlethal Violence Against an Intimate Female Partner, Comparing Male Murderers to Nonlethal abusers', *Violence against Women*, vol 13, 329-353
- Doll, R and Hill, B (1950) 'Smoking and Carcinoma of the Lung', *British Medical Journal*, 739-748
- Ericson, R and Haggerty, K (1997) *Policing the Risk Society*, Buffalo: University of Toronto Press

- Farrington, D and Tarling, R (Ed) (1985) *Prediction in Criminology*, Albany: State University of New York Press
- Farrington, D, Coid, J, Harnett, L, Jolliffe, D, Soteriou, N, Turner, R and West, D (2006) *'Criminal careers up to age 50 and life success up to age 48: new findings from the Cambridge Study in Delinquent Development Home Office Research Study 299'* London; Home Office
- Fischhoff, B (2003) 'Hindsight: foresight; the effect of outcome knowledge on judgement under uncertainty', *Qual Saf Health Care* 2003, 12, 304-312
- Gladwell, M (2005) *Blink*, London: Penguin
- Gottfredson, D M in Gottfredson, D M and Tonry, M (1987) Ed *Prediction and Classification Criminal Justice Decision Making*, Chicago, University of Chicago Press
- Hart, B (1988) 'Beyond the duty to warn: a therapists duty to protect battered women and children' in Yllo, K and Bogard, M (Ed) *Feminist perspectives on wife abuse*, California, Sage
- Heckert, D and Gondolf, E (2004) 'Battered women's perceptions of risk versus risk factors and instruments in predicting repeated reassault', *Journal of Interpersonal Violence*, 19, 778-800
- Humphreys, C, Thiara, R, Regan, L, Lovett, J, Kennedy, L, and Gibson, A, (2005) 'Prevention not prediction? A preliminary evaluation of the Metropolitan Police Service Domestic Violence Risk Assessment model (SPECSS+)' unpublished report by the Centre for the Study of Safety and Wellbeing, University of Warwick and the Child and Woman Abuse Studies Unit, London Metropolitan University
- Johnson, M (2008) *Intimate terrorism, violent resistance, and situational couple violence*, New Hampshire: Northeastern University Press
- Kahneman, D and Tversky, (1982) 'Evidential Impact of base rates' in Kahneman, D, Slovic, P and Tversky, A (Ed) *Judgement under uncertainty: Heuristics and biases*, Cambridge: Cambridge University Press
- Kropp, P (2004) 'Some questions regarding spousal assault risk assessment', *Violence against women*, 4, 676-697
- Krug, E, Dahlberg, L, Mercy, J, Zwi, I and Lozano, R (2002) *World Report on Violence and Health*, Geneva: World Health Organisation
- Logan, C (2005) 'Risk Assessment and management in sexual and violent offending' unpublished report by the Liverpool University Senior Baxter Research Fellow

MacVean, A and Ridley, N (2007) 'An Evaluation of the Thames Valley Response to Domestic Violence' unpublished report to the Chief Constable of Thames Valley Police

Meehl, P (1954) *Clinical vs Statistical prediction; a theoretical analysis and a review of the evidence*, Minneapolis: University of Minnesota Press

Monahan, J (1981) *Predicting Violent Behaviour: An assessment of Clinical Techniques*, Beverley Hills: Sage

Monahan, J and Steadman, H (2003) 'Crime and Mental Disorder: an epidemiological approach' in Bean, P (Ed.) *Critical Concepts in Sociology*, London: Routledge

Nixon, K (2009) 'The Development and Validation of a Domestic Abuse Risk Identification and Management Tool' unpublished doctoral thesis

Pease, K (2010) 'Review of DASH' unpublished report to the National Policing Improvement Agency

Povey, D (ed) (2005) *Crime in England and Wales 2003/4: Supplementary Volume 1 Homicide and Gun Crime, Home Office Statistical Bulletin No 02/05*, London: Home Office

Richards, L (2004) 'Homicide prevention: Findings from the Multi-agency Domestic Violence Homicide Reviews', *The Journal of Homicide and Major Investigation*, vol 2, 53-72

Richards L, (2006) *'Getting away with it: a Strategic Overview of Domestic Violence, Sexual Assault and Serious Incident Analysis'* London: Metropolitan Police

Richards, L, Letchford, S and Stratton, S (2008) *Policing Domestic Violence*, Oxford: Oxford University Press

Robinson, A (2006) 'Reducing Repeat Victimization among high risk victims of Domestic Violence', *Violence against Women*, vol 12 No 8, 761-788

Roehl, J, O'Sullivan, C, Webster, D and Campbell, J (2005) 'Intimate Partner Violence Risk Assessment Validation Study: the RAVE Study Practitioner Summary and Recommendations; Validation Tools for Assessing Risk from Violent Intimate Partners' report submitted to the US Department of Justice

Sargent, D and Campbell, J (2008) 'Assessing lethality in domestic violence cases' *Gazette* vol 70 No 2

Saltzman, L, Mercy, J and Rhodes, P (1992) 'Identification of Nonfatal Family and Intimate Assault Incidents in Police Data', *American Journal of Public Health*, 82, 1018-1020

Scott, C, Quanbeck, C and Resnick P (2008) 'Assessment of Dangerousness' in Gabbard, G, Yudofsky, S and Hales, R (Ed) *The American Psychiatric Publishing Textbook of Psychiatry, fifth edition*, Washington DC: American Psychiatric Publishing Inc

Sherman, L, Schmidt, J Rogan, D and DeRiso, C (1991) 'Predicting Domestic Homicide: Prior Contact and Gun Threats', *Women battering: Policy Responses*, Cincinnati Ohio: Anderson Publishing

Sherman, L and Strang, H (1992a) 'Predicting Domestic Violence from Police Records' unpublished report to the Chief Commissioner Victoria Police

Sherman, L (1992b) *Policing Domestic Violence: Experiments and Dilemmas*, New York: Free Press

Stanko B, (2008) 'Managing Performance in the Policing of Domestic Violence', *Policing vol 2*, 294-302

Sully, P and Greenaway, K (2004) *An Investigation into the antecedents of domestic homicide with a view to its prevention*, London: City University

Trujillo, M and Ross, R (2008) 'Police Response to Domestic Violence, Making decisions about risk and risk management' *Journal of Interpersonal Violence*, published online before print

Van Wormer, K and Roberts, A (2009) *Death by Domestic Violence*, Westport: Praeger Publishers

Websdale, N (1999) *Understanding Domestic Homicide*, North Carolina: Northeastern University Press

Wilkins, L (1985) 'The politics of prediction' in Farrington, D and Tarling, R (Ed) *Prediction in Criminology*, Albany: State University of New York Press

Zawitz, M, Klaus, P, Bachman, R, Langa, P, Graziadei, H and Harlow, C (1994) *Violence between intimates*, Rockville: National Institute of Justice

## Appendix 1

1. Has the current incident resulted in injury?
2. Are you very frightened?
3. What are you afraid of? Is it further injury or violence (please give an indication of what you think (...) might do and to whom)
4. Do you feel isolated from family i.e. does (...) try to stop you from seeing friends/family/Dr or others?
5. Are you feeling depressed or having suicidal thoughts?
6. Have you separated or tried to separate from (...) within the past year?
7. Is there conflict over a child contact?
8. Does (...) constantly text, call, contact, follow, stalk or harass you?
9. Are you currently pregnant or have you recently had a baby (in the past 18 months)?
10. Are there any children, step-children that aren't (...) in the household? Or are there any other dependants in the household (i.e. older relative)?
11. Has (...) ever hurt the children/dependants?
12. Has (...) ever threatened to hurt or kill the children/dependants?
13. Is the abuse happening more often?
14. Is the abuse getting worse?
15. Does (...) try to control everything you do and/or are they excessively jealous?
16. Has (...) ever used weapons or objects to hurt you?
17. Has (...) ever threatened to kill you or someone else an you believed them?
18. Has (...) ever attempted to strangle/choke/suffocate/drown you?



19. Does (...) do or say things of a sexual nature that makes you feel bad or that physically hurt you or someone else?
20. Is there any other person that has threatened you or that you are afraid of?
21. Do you know if (...) has hurt anyone else?
22. Has (...) ever mistreated an animal or family pet?
23. Are there any financial issues? For example are you dependent on (...) for money/have they recently lost their job/other financial issues?
24. Has (...) had problems in the past year with drugs (prescription or otherwise), alcohol, or mental health leading to problems in leading a normal life?
25. Has (...) ever threatened or attempted suicide?
26. Has (...) ever breached bail/an injunction and/or any agreement for when they can see you or the children?
27. Do you know if (...) has ever been in trouble with the police or has a criminal history?

## Appendix 2

- ABH
- Assault on Police
- Attempted Murder
- Causing Death By Dangerous or Careless Driving under the influence
- Child Abduction
- Common Assault
- Conspiracy to Murder
- Cruelty/Neglect of Children
- Death by Careless or inconsiderate Driving
- Death by Driving - Unlicensed etc
- GBH with Intent
- GBH without Intent
- GBH/ABH
- Harassment
- Harassment/Public Order
- Manslaughter
- Murder
- Possession of article with Blade or Point
- Possession of Firearms with Intent
- Possession of Offensive Weapon
- Possession of Other Weapons
- Public Order
- Threats to Kill