



College of  
Policing

[college.police.uk](http://college.police.uk)

# Piloting a new approach to domestic abuse frontline risk assessment

Evaluation report

Julia Wire

Andy Myhill

© College of Policing Limited (2018). This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit [nationalarchives.gov.uk/doc/open-government-licence/version/3](http://nationalarchives.gov.uk/doc/open-government-licence/version/3), or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: [psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk).

Where we have identified any third party copyright information, you will need to obtain permission from the copyright holders concerned.

This publication is available at: <http://whatworks.college.police.uk/Research/Pages/Published.aspx>.

The College of Policing will provide fair access to all readers and, to support this commitment, this document can be provided in alternative formats.

Any enquiries regarding this publication, including requests for an alternative format, should be sent to us at: [contactus@college.pnn.police.uk](mailto:contactus@college.pnn.police.uk).

# Executive summary

## Key findings

- Using the piloted risk assessment tool, first response officers' initial assessments of risk were less likely to be regraded, during a post-incident review, than assessments made using DASH.
- Victims in one force disclosed perpetrators' coercive and controlling and stalking and harassment behaviours at greater rates using the piloted risk assessment tool.
- First response officers recorded proportionately more crimes of coercive control during the pilot, though numbers overall were still small.
- Although the pilot showed an increase in officers documenting coercive and controlling behaviour, that outcome did not always translate into them assigning an appropriate risk grading. This finding was most pronounced for post-separation harassment and stalking.
- First response officers and their supervisors generally said they preferred the piloted risk assessment tool to the DASH. Some specialist support workers preferred DASH, but recognised its shortcomings in a policing context, while some preferred the piloted tool.

## Implications

- If adopted, the revised risk tool for frontline officers may improve both victims' disclosure and officers' primary risk assessments.
- If adopted, officers require an understanding of coercive control in order for the tool to be most effective.
- If adopted, the revised risk tool may result in a greater proportion of cases being graded as medium as opposed to standard risk, reflecting the ongoing nature of the abuse. Any impact on resources for secondary risk assessment would need to be balanced against potential benefits associated with earlier referral and intervention.
- If adopted, there should be national-level consideration of the potential impact of the revised risk tool on risk management processes, including referrals to partner agencies and Multi-Agency Risk Assessment Conference (MARAC).

## Introduction

The majority of police forces in England and Wales use the Domestic Abuse, Stalking and Harassment and Honour-based violence risk identification, assessment and management model (DASH), endorsed by National Policing leads in 2009, for the identification, assessment and management of risk in cases of domestic abuse. The DASH model was not subject to systematic empirical evaluation following implementation, and a review in three forces (Robinson et al., 2016) found inconsistencies in how it operated. The research also suggested officers struggled frequently to identify and document ongoing and dangerous patterns of controlling and coercive abuse.

Following publication of the DASH review research, the College, with input from survivors of domestic abuse, police practitioners, national support services, and leading academics,

designed a revised risk assessment tool for frontline officers to be piloted in three forces. The revised risk tool was intended to enable first response officers to identify ongoing abuse more effectively, and in particular coercive control, and make more consistently an appropriate initial assessment of risk of harm to the victim.

## **The pilot**

In contrast with the DASH, which employs yes/no questions with the expectation that officers will record contextual information relating to each, the revised tool asked victims to say how often specific behaviours occurred on a scale from 'never' to 'all the time' with the aim of creating a clearer picture of the nature and intensity of the abuse. Officers were expected to complete a single freetext section at the end of the process, giving their rationale for the level of risk and considering aggravating factors such as substance abuse and mental health, and other elements of vulnerability such as pregnancy.

The pilot ran for three months in divisions within three forces: Sussex, West Midlands and Humberside. First response officers and their supervisors received a training input prior to implementation that included interactive work based on scenarios using the revised risk tool. Officers in the pilot areas used the revised risk tool instead of the DASH for the duration of the pilot period.

## **Research methods**

The evaluation used multiple methods: case-file analysis, in-depth interviews, and direct observations of officers using the revised tool. The evaluation used a non-equivalent comparison group design. Comparison areas were suggested by the pilot forces to be as similar as possible to the pilot areas in terms of demographics and demand (volume of domestic abuse calls for service). Resulting samples were checked for similarity.

Data was collected for a time period both before and during implementation of the pilot, in both pilot and comparison areas. Due to errors in the data provided by Humberside, it was only possible to report findings from case-file analysis for Sussex and West Midlands. Data from interviews and observations undertaken in Humberside were reported.

The key outcome measure for the evaluation was researchers' assessments of whether the first responding officer had made an appropriate initial assessment of risk. Assessments were made by two researchers with specialist subject area knowledge, and based on a review of police case-files. To provide a degree of external quality assurance, a specialist domestic abuse support worker coded independently 25% of cases. There was a high degree of agreement between the risk grading decisions of the researchers and the support worker.

In-depth interviews were undertaken with first response officers, supervisors, police secondary risk assessors, and specialist support workers from partner agencies. In addition, researchers observed 17 shifts across the three forces in the areas where the revised risk tool was piloted.

## **Findings**

Results indicated higher rates of agreement between researchers and first responding officers on the level of risk to the victim when officers used the piloted risk tool.

- In Sussex, during the pre-pilot period, researchers made the same risk grading decision as officers in 57% of cases, rising to 71% during the pilot. In the comparison area, the level

of risk grading agreement remained consistent at around 56%.

- In West Midlands, researchers agreed with officers' risk grading in 53% of cases pre-pilot, rising to 73% during the pilot. There was no improvement in the comparison area.

## Changes to the distribution of risk gradings

There was an increase in the proportion of cases graded as medium risk in Sussex during the pilot, from 16% to 30% of cases. There was a corresponding decrease in cases graded as standard risk, from 82% to 66%. There was very little change in the distribution of risk gradings in the pilot area in West Midlands. The baseline proportion of cases graded medium risk by officers was higher in West Midlands than in Sussex so it is possible that in West Midlands there was movement in both directions.

It is likely that the increase in the proportion of medium risk cases in Sussex was a result of officers identifying patterns of abusive behaviour more frequently during the pilot (see below). It is not clear why a similar pattern was not found in West Midlands, but an existing policy of discretionary use of DASH in the pre-pilot period and in the comparison area may have meant the samples were less equivalent than in Sussex.

Some officers interviewed felt the piloted risk tool helped them better identify risk. A majority of supervisors and secondary risk assessors interviewed felt that officers graded risk more consistently with the piloted risk tool.

## Disclosure, identification and recording of coercive control

In Sussex, case-file analysis showed victims were more likely to disclose ongoing abuse when the piloted risk tool was used. In addition, more than twice as many victims in Sussex stated they felt the abuse they were suffering was getting worse when asked during the pilot period (45%) compared to the pre-pilot period (21%), suggesting that the piloted risk tool may be more effective in helping victims to identify and/or acknowledge the abuse they are suffering.

In West Midlands, there were no statistically significant increases in the disclosure of controlling behaviour between the pre-pilot and pilot periods. It is not clear why rates of disclosure did not increase as they did in Sussex, but less equivalent samples and smaller sample sizes may have had an impact. Rates of disclosure of coercive control in West Midlands were no worse with the piloted risk tool than with the DASH.

The majority of first response officers interviewed, as well as their supervisors and secondary assessors, felt the piloted risk tool made it easier to identify coercive and controlling behaviour. Keyword analysis of officers' incident reports and risk assessments indicated a greater tendency to identify and document coercive control – or the absence of coercive control – during the pilot. In both Sussex and West Midlands, there was an increase in the number of crimes of 'controlling and coercive behaviour' recorded during the pilot. These increases were from a low baseline, however, and the offence of coercive control was under-recorded in all pilot forces.

## Stakeholder perceptions of the piloted risk tool

First response officers, their supervisors and police officers and staff responsible for secondary risk assessments reflected positive views of the piloted risk tool, and generally preferred it to the DASH. No negative responses from victims to the piloted risk tool were reported by first response officers, or witnessed during the observations.

The views of specialist support workers were mixed in relation to the piloted risk tool. Those who said they preferred the DASH recognised its shortcomings in a police context. What they appeared to favour was a thoroughly completed DASH. In Sussex, a support worker who dealt only with high risk cases (where risk assessments tend to be completed most comprehensively) stated a preference for the DASH, whereas a support worker who dealt with only medium and standard risk cases much preferred the piloted risk tool. Specialist support workers who favoured the piloted risk assessment felt that it was more suited to first response officers, and that it was generally completed better than DASH.

## **Conclusions and implications**

The evaluation of a revised domestic abuse risk tool for first response officers showed positive results. Specifically, officers made more appropriate risk grading decisions when using the piloted risk tool than when using the current DASH tool. In addition, it appeared that the revised tool might mitigate some of the shortcomings of DASH when the DASH assessment is completed by frontline officers. Specifically, it appeared some victims were more prepared to disclose coercive and controlling behaviour to officers using the piloted risk tool.

### **What are the benefits of a revised risk tool for first response officers?**

The DASH review (Robinson et al., 2016) suggested DASH does not, in many cases, present a clear indication of the risk of harm to the victim when used by frontline officers. The piloted risk tool contained more focused and direct questions, based on the frequency with which specific abusive behaviours are experienced. This format appeared to encourage victims to disclose, and gave officers (and any subsequent supervisor or risk assessor) a clearer picture in some cases of the nature and intensity of the abuse. The single freetext section also appeared to be successful to some degree in getting officers to offer a clear rationale for their assessment of the level of risk to the victim.

### **How important is knowledge and understanding of domestic abuse?**

A revised risk tool alone will not result in officers making appropriate primary risk assessments in all cases. As well as better identifying and documenting ongoing abuse, officers must appreciate the risk this behaviour poses over and above discrete acts of physical violence. The pilot showed an increase in officers documenting coercive control, but that did not always translate into assigning an appropriate risk grading. This finding was most pronounced for post-separation harassment and stalking behaviours.

### **What are the possible implications for secondary risk assessment and risk management?**

If the revised risk tool were to be adopted for first response officers, there would need to be national-level consideration of the potential implications for police secondary risk assessment. The DASH review (Robinson et al., 2016) suggested, however, that the police secondary risk assessment process currently is not consistent or, in some cases, well-functioning. The revised risk tool may in fact have advantages for secondary risk assessors, and not just in that ongoing patterns of abusive behaviour may be better identified and recorded. Part of the difficulty for secondary assessors appears to be that relevant information is dispersed between different sources (the risk assessment and the incident report). Having both the risk assessment and the officer's rationale for their risk grading presented in the same place would be an advantage. This approach would also benefit specialist support services, who are most often sent the risk assessment for cases referred, but not the officer's incident report.

Perhaps the most significant implication for police secondary risk assessment would be any increase in the proportion of cases graded as medium rather than standard risk. Though it is not possible to know if the exact increase seen in Sussex would be replicated in other forces, any increase, while reflecting improved recognition of cases of ongoing abuse, may have implications for resources both for police and specialist support agencies. Any such impact would need to be balanced, however, against the potential benefit of reducing frontline demand if cases are referred more efficiently and escalation prevented through earlier intervention.

Another area in which the introduction of the revised risk tool may affect force practices is the process of referring cases to MARAC. The DASH review (Robinson et al., 2016) suggested some forces automatically trigger MARAC referrals based on 14 'yes' responses to the DASH. There is a lack of empirical evidence to support using a numerical 'cut-off' score. As the revised risk tool cannot, or at least should not, be scored in this way, MARAC referrals would need to be based solely on professional judgement. As the evaluation did not suggest the revised risk tool would result in a greater proportion of cases being graded as high risk, there would not likely be a significant increase in MARAC caseloads were it to be adopted.

## **Implications**

The key implications of the evaluation of the risk pilots are:

- If adopted, the revised risk tool for frontline officers may improve both victims' disclosure and officers' primary risk assessments.
- Officers require an understanding of coercive control in order to make effective structured professional judgements around threat of harm.
- If adopted, the revised risk tool may result in a greater proportion of cases being graded as medium as opposed to standard risk, reflecting the ongoing nature of the abuse.
- If adopted, there needs to be national-level consideration of the impact on police secondary risk assessment, and the potential impact of the revised risk tool on risk management, including referrals to partner agencies and MARAC.

# Contents

Executive summary .....	i
Contents .....	vi
Acknowledgements.....	vii
1. Introduction.....	1
2. The pilot.....	5
3. Research methods.....	9
4. Findings .....	16
5. Conclusions and implications.....	38
Appendix 1. Rationale for the development, structure and content of piloted risk tool .....	44
Appendix 2. Piloted risk tool .....	54
Appendix 3. Research methods .....	56
Appendix 4. Humberside data .....	58
Appendix 5. Logistic regression difference-in-differences model .....	60
Appendix 6. Suggested changes to the piloted risk tool.....	63
References .....	64



# Acknowledgements

Our thanks go to Alison Eaton, Simon Yates, Dan Dugan and Ben Newman of Sussex police; to Kat Sibley, Lydia Mason, Angie Whitaker and Louisa Rolfe of West Midlands Police; and to Phil Gadd and Anne Marie McLoughlin of Humberside Police. We also thank all of the officers and staff who gave up their time to be interviewed for the evaluation, and who assisted with this work in any other way.

We thank Professor Amanda Robinson (Cardiff University) and Dr Katrin Hohl (City, University of London) for their help and advice throughout the project.

Our thanks go to Hilary Fisher and colleagues at Women's Aid for help and advice and for facilitating valuable input from survivors of domestic abuse. Also Mary Mason (Solace Women's Aid) for facilitating specialist quality assurance of data coding, and to support worker Anat Toffell for agreeing to undertake this work.

We would also like to thank David Thomason (Cheshire Constabulary), Professor Marianne Hester (University of Bristol) and Associate Professor Lauren B. Cattaneo (George Mason University) for their extremely helpful peer reviews.

Finally, we would like to thank colleagues at the College of Policing for their valuable assistance with elements of the fieldwork, and for their reviews of final outputs, and members of the virtual project reference group for their interest in this work.

# 1. Introduction

Police forces in England and Wales began to introduce processes concerned with the identification and assessment of risk in cases of domestic abuse from the early 2000s. In 2009, in an attempt to make risk assessment more consistent among forces, National Policing leads endorsed a national model for risk identification, assessment and management: the Domestic Abuse, Stalking and Harassment and Honour-based violence risk identification, assessment and management model (DASH). The DASH model is centred on a twenty-seven question risk identification checklist, asked of victims in most instances by a first responding police officer. The information gathered during this risk identification interview informs both primary risk assessment and initial safeguarding, undertaken by the attending officer(s), and a secondary risk assessment undertaken by officers or staff in a central unit for all, or a proportion, of cases (see Robinson et al., 2016). Both primary and secondary assessors are required to use professional judgement to allocate a grade of 'standard', 'medium' or 'high' to indicate the risk of the victim being seriously harmed by the perpetrator.

Though it was adopted by the majority of police forces in England and Wales, the DASH model was not, following its roll out, subject to systematic empirical evaluation in relation either to its implementation or its ability to identify accurately and consistently the risk posed to victims by perpetrators. In 2014, Her Majesty's Inspectorate of Constabulary (HMIC; now Her Majesty's Inspectorate of Constabulary and Fire & Rescue Services, HMICFRS) published findings from a thematic inspection of police response to domestic abuse in England and Wales (HMIC, 2014a). In the context of a largely critical report, HMIC identified risk assessment as a specific area of concern. During fieldwork for the inspection, HMIC (2014a: 67) identified 'a range of different and inconsistent practices when assessing the risk to victims', and concluded that 'response officers often have a limited understanding of what it means to assess risk' (HMIC, 2014a: 69). Recommendation 6 of the report suggested the College of Policing 'consider the current approach to risk assessment with others, such as practitioners in forces, academic experts and organisations supporting practitioners and victims', and 'make an assessment of the sufficiency of the tools that frontline officers are given to assess risk' (HMIC, 2014a: 22).

The HMIC inspection also suggested that officers are 'often ill equipped to identify dangerous patterns of behaviour', especially where there is little physical violence but significant 'psychological intimidation and control' (HMIC, 2014a: 9). Recent published research has also shown officers often fail to identify coercive and controlling abuse, and tend to prioritise physical assault and injury during the current incident (see Robinson et al., 2017). Coercive control (see Stark, 2007) is a form of domestic abuse in which perpetrators use multiple abusive behaviours and tactics to entrap victims. In December 2015, a criminal offence of 'controlling or coercive behaviour in an intimate or family relationship' was enacted in England and Wales. Data published by the Office for National Statistics (ONS, 2017) shows that just 4,246 offences of coercive control were recorded in the year to March 2017 by the 38 forces who provide data to the Home Office data hub. To set this figure in context, the police service in England and Wales received more than 900,000 domestic abuse-related calls for service

in the year to March 2015 (HMIC, 2015: 28). If the service is to address the almost certain under-recording of coercive control, it is essential first that frontline officers are able to identify the behaviours that comprise this offence.

## A review of the DASH model

Partly in response to the HMIC recommendation, Cardiff University, in their role as part of the What Works Centre Consortium, undertook a review of how the DASH risk model was operating in forces. The resulting research (Robinson et al., 2016) was undertaken in partnership between Cardiff, the College and University College London, and involved extensive fieldwork in three police forces, comprising: a review of case-files, in-depth interviews with and a survey of officers and staff, and direct observations of officers attending domestic abuse incidents. Findings from this research supported many of the impressions from the HMIC inspection:

- Despite widespread support for the principle of risk assessment, the DASH model was implemented inconsistently both between and within forces.
- Officers sometimes used discretion not to submit a DASH risk assessment, specific questions were altered or omitted, and information was sometimes recorded in an inconsistent or incomplete way.
- Police practitioners frequently lacked an understanding of coercive control, focusing disproportionately on physical violence and injury at the current incident during both primary and secondary assessments of risk.

Indeed, numerous specific examples were identified during the fieldwork that reflected a lack of understanding of the dynamics of coercive control (see Robinson et al., 2016). Secondary risk assessment was found also to be variable, reflecting perhaps the nature of the original guidance to forces concerning implementation of the DASH risk model. While the original DASH checklist for police suggested first response officers undertake risk **identification** and specialist officers risk **assessment** for **all** cases, the 'practice guidance for police staff' suggested secondary risk assessment be undertaken only on cases graded by the first responding officer as medium or high (see ACPO, 2009). The guidance also suggested that secondary risk assessment should involve further contact with the victim to clarify or expand on information collected by frontline officers during the risk identification interview, which the DASH review indicated happened very rarely in practice.

A key issue with the implementation of DASH appears to have been that the checklist at the heart of the model attempts to perform multiple functions – the same twenty-seven yes/no questions serve as a frontline risk identification checklist/primary risk assessment, and also as a secondary risk-cum-needs assessment for police specialists and specialist support services. In practice, it appears that the DASH risk tool performs the latter function much better than the former (see Robinson, 2010). The initial yes/no nature of the DASH questions means that the success of the model rests with the ability of the interviewer to elicit detailed contextual information and record it on the risk assessment. Though there are exceptions, DASH risk assessments completed by

specialists – police or support workers – tend to be more comprehensive than those completed by frontline officers (see Robinson, 2010). The DASH review research (Robinson et al., 2016) confirmed anecdotal suggestions that some frontline officers lacked the knowledge, skill and time necessary to produce a comprehensive and accurate primary risk assessment. Yet it was also evident that some victims – for reasons that are both valid and understandable in the context of coercive control – are not prepared to engage in-depth with frontline officers in order to provide the contextual information to illuminate initial yes/no responses.

## A revised tool for frontline risk assessment

Following publication of the DASH review research, the College, with input from Cardiff University, designed a revised risk tool to be piloted in three forces and intended to enable first response officers to more effectively identify risk, and in particular coercive control, and make an initial assessment of the risk of serious harm posed to the victim by the perpetrator. A detailed description of the development of the tool, including rationales for the inclusion/exclusion of specific questions, can be found at Appendix 1. The underlying assumptions and key principles for the revision of the risk tool were:

- Coercive control, including post-separation stalking and harassment, is the most harmful form of domestic abuse and presents the greatest risk of serious injury and homicide (Campbell et al., 2003; Stark, 2007; Dobash and Dobash, 2015; Myhill, 2015; Sharp-Jeffs and Kelly, 2016; Monkton-Smith et al., 2017).
- Not all of the abuse that comes to the attention of the police under the wide-ranging national definition of domestic abuse is coercive control (Myhill, 2016; Butterworth and Westmarland, 2016).
- Frontline officers frequently take an incident-based approach to domestic abuse (Hoyle, 1998; Robinson et al., 2016); a revised risk tool for frontline officers should focus attention on dangerous patterns of behaviour.
- A revised tool should improve consistency and completeness of risk data by providing a straightforward and structured method of obtaining information from victims.
- A revised tool should improve the accuracy of subsequent risk assessment and referrals to specialist support services.
- A primary risk assessment should contain only questions that research evidence suggests are key to identifying risk of harm and making an initial assessment of safeguarding requirements; a more comprehensive risk/needs assessment is best undertaken by a specialist.

It was hoped, then, that the revised risk assessment would make it easier both for victims to disclose coercive and controlling abuse, and officers to identify and record it.

## **Structure of the report**

Chapter 2 describes the implementation of the pilot in three police forces. Chapter 3 details the research methods used to evaluate the pilots. Chapter 4 presents findings from the evaluation, and chapter 5 comprises a discussion of the findings and draws out implications for policy and practice.

## 2. The pilot

The revised risk tool was piloted in sub-divisions of three police forces: Sussex, West Midlands, and Humberside.

In contrast with the DASH, which employs yes/no questions with the expectation that officers will record contextual information relating to each, the revised tool was premised primarily on acknowledging that abuse is likely occurring and asking victims to say how often specific behaviours occur on a scale from 'never' to 'all the time'. Officers were expected to complete a single freetext section at the end of the process giving their rationale for the level of risk and to consider the victim's level of fear and entrapment, and any aggravating factors that might increase threat or vulnerability, such as pregnancy or alcohol, substance abuse or mental health issues. The key differences between the revised risk tool and the DASH are summarised in Box 1. The risk tool that was piloted can be found at Appendix 2.

### Box 1: Key differences between piloted risk tool and DASH

DASH	Piloted risk tool
Twenty-seven yes/no questions that ask whether specific abusive behaviours have ever happened	Sixteen questions most of which ask how often abusive behaviours happen, on a scale from 'never' to 'all the time'
Officers expected to provide freetext description following each question	Officers expected to complete single freetext rationale at the end of the risk assessment
Definitions of 'standard', 'medium' and 'high' risk taken directly from OASys risk assessment for general offending <sup>1</sup>	Definitions of 'standard', 'medium' and 'high' risk adapted to take into account the continuous nature of coercive control

The process of primary and secondary risk assessment remained unaltered during the pilot, save for the introduction of the revised risk tool. Though processes vary somewhat between forces (see Robinson et al., 2016), the basic process in all pilot forces was:

- First responding officer(s) attend a domestic abuse incident and complete a primary risk assessment, allocating a risk grade of standard, medium or high using professional judgement. They may or may not perform background intelligence checks on force systems.
- First line supervisor (Sergeant) reviews the circumstances of the incident and the primary risk assessment and revises the risk grade if they deem it necessary, using professional judgement. They may or may not perform background intelligence checks on force systems.

<sup>1</sup> OASys is the 'Offender Assessment System' used by the prison and probation services to assess the risk posed by offenders under supervision.

- Secondary risk assessment undertaken on all or a proportion of cases by a central unit, using professional judgement, and drawing on the primary risk assessment, information contained on force systems, and, in some cases, the Police National Computer and data from partner agencies.
- Referrals to specialist support services and/or MARAC made in accordance with force policies.

The only key change during the pilot was, then, the introduction of the revised primary risk assessment meaning supervisors and secondary risk assessors received a different information set, on a different template.

## Implementation

The pilot ran for approximately three months in each of the three pilot forces. Prior to officers using the revised risk tool, they received a three-hour training input. The training was delivered by force trainers and specialist public protection officers who prior to delivering the training attended a 'train the trainer' event hosted by the College. This event was to allow trainers to become familiar with and understand the training materials, and to try to ensure the training was delivered consistently between sites.

The training consisted of PowerPoint presentations describing the background to the pilot, and the rationale for and design of the risk tool. This input was followed by interactive group work during which officers worked through several scenarios using the risk tool, grading cases as standard, medium or high risk and producing written rationales to justify their decisions. During a final debrief session, officers' risk grading and rationales were compared with those of a group of specialist domestic abuse support workers who completed the same exercise prior to the training.

It should be noted that the training input, which involved some scenarios focused on coercive control, alongside the novelty/'newness' of the revised risk tool, is a potential alternative explanation for any positive results reflected in the pilot evaluation. This possibility applies especially to officers' identification and recording of coercive and controlling behaviours. This potential explanation is mitigated somewhat by the fact that all three pilot forces had implemented some form of training focused on coercive control in the year preceding the pilot. In addition, in Sussex there was no difference in scores for the key outcome measure over the first seven weeks of the pilot indicating there was not an immediate 'newness' effect that wore off as the pilot continued.

### Sussex

In Sussex, the pilot was implemented between December 2016 and April 2017 in the Eastbourne, Wealden and Lewes division of the force. The Hastings and Rother division acted as a comparison site.

### West Midlands

In West Midlands, the pilot was implemented from late January to May 2017 in the Birmingham East division. As the Birmingham East division is more densely

populated than the pilot division in Sussex, with a proportionately greater number of domestic abuse incidents, two teams (B and C) were chosen to implement the pilot; the remaining teams on the same division (A, D and E) provided the comparison condition.

A context specific to West Midlands was their policy (since discontinued) of discretionary use of the DASH risk assessment for some cases of domestic abuse. All officers prior to the pilot, and officers in the comparison teams during the pilot period, were only required to submit a DASH risk assessment where there had been a criminal offence recorded between intimate partners. For all other incidents – those involving intimate partners where no crime was recorded, and any incident involving family members as opposed to intimate partners – officers were permitted to use discretion as to whether a risk assessment was submitted. In practice, the researchers found this to mean that very few risk assessments were submitted for incidents that did not involve intimate partners, or for incidents that involved intimate partners but where a criminal offence was not recorded. In order to maximise the number of cases for analysis, officers using the revised risk tool during the pilot – but not those in the comparison area – were instructed to submit a risk assessment for all domestic abuse incidents. In practice, however, officers only submitted a risk assessment for approximately half of all incidents (suggesting the policy of officer discretion was difficult to ‘roll back’ for this specific period of time).

The context of discretionary use of DASH means that the sample of incidents in West Midlands was not as representative as that in Sussex. Specifically, in the comparison area, and before the introduction of the revised risk tool in the pilot area, it was more heavily slanted towards incidents where a criminal offence was recorded. The impact of this context on the evaluation findings is not clear, but it is possible that the sample in West Midlands was biased towards risk assessments for which officers might have been more likely to have recorded relevant information. Recent research has suggested many officers still equate identifiable criminal offences with greater threat of harm to the victim (see Robinson et al., 2016; Myhill, 2017).

A final issue with data collection in West Midlands was that first response officers were required to scan paper-based risk assessments on to an IT system. The day-to-day project lead for the force (an Inspector in the public protection unit) suggested that DASH forms may not always be scanned to the system due to operational or IT issues. Some officers recorded victims’ responses to the DASH questions on the force’s online crime portal as part of their incident and primary investigation reports, though the researchers found this practice to be variable. Though not likely to have affected a large proportion of cases, it is possible that for some incidents a DASH (or pilot) risk assessment may have been completed, but the information was not available to the researchers.

## Humberside

In Humberside, the revised risk assessment was piloted in the Bridlington division. Cleethorpes, also a seaside town, was chosen as a comparison site. The pilot ran in Bridlington from January to April 2017. There were, however, problems identified with data provided for the evaluation by Humberside Police. Risk assessment data that had been transcribed by an administration team was



found to contain significant errors. It was not possible to re-collect the data at source, as some paper-based risk assessments had been destroyed after transcription. The force has moved subsequently to a different system that eliminates the need for manual transcription of data. The errors in the data meant that the comparison condition in Humberside could not be used. See Appendix 4 for further detail, and for basic reporting of the case-file data from the pilot division. Interview data from Humberside, concerning for example officers' views on the piloted risk tool, has been included within the main body of this report.

## 3. Research methods

The evaluation of the risk pilot used multiple methods: case-file analysis, in-depth interviews, and direct observations of officers using the revised tool at domestic abuse incidents.

### Study design

Data was collected before and after the introduction of the revised risk tool both in a pilot area and in a comparison area in each force. This non-equivalent group design enabled comparisons between officers' risk grading before and after the introduction of the pilot risk assessment (referred to in this report as the 'pre-pilot' and 'post-pilot' periods respectively). The comparison condition permitted an assessment of whether or not officers who used the DASH risk tool at both points of measurement reflected the same differences as those using the piloted risk tool (see Appendix 3 for further discussion on study design).

Comparison areas were suggested by the forces to be as similar as possible to the pilot areas in terms of demographics and level of demand. All forces in the pilot had recently undergone some form of coercive control training. Forces were told not to change their practice in relation to domestic abuse in any other way for the duration of the pilot.

### Case-file analysis

Pilot forces granted the researchers full access to their record management systems for the duration of the pilot and evaluation.

The key outcome measure for the evaluation was researchers' assessments of whether the first responding officer had made an appropriate initial assessment of risk (standard, medium or high). Other potential outcome measures – such as revictimisation, or further police contact – were considered not to be viable. Revictimisation as an outcome is difficult to operationalise in relation to domestic abuse. Firstly, victims may or may not report further abuse, due to the behaviour of the perpetrator, the service they receive from the police, or other factors. Second, police data has been shown to be incomplete in relation to domestic abuse (see HMIC, 2014b; Myhill and Johnson, 2016), meaning outcome measures based solely on further arrests or convictions are unreliable. Thirdly, in-depth research has suggested positive outcomes for victims extend far beyond the discrete physical assaults that dominate police recorded crime data (Westmarland et al., 2010), and that desistance from abuse is a gradual process for the majority of perpetrators (Kelly and Westmarland, 2015). Finally, even if longitudinal follow-up interviews with victims were undertaken, there are too many factors beyond application of a primary risk assessment that may affect revictimisation and further contact with the police. Thus the most realistic and measurable outcome for this study was the quality of officers' primary assessments of risk.

Assessments of officers' risk decisions were made by two researchers with specialist subject area knowledge, and based on as far as possible the same information available to the first responding officer(s), principally:

- The circumstances of the incident (as described in the officer's incident report).
- The wider context of the relationship – as described in the incident report, and/or in documentation relating to any previous incidents.
- The information contained in the risk assessment.
- The offending history of the perpetrator.<sup>2</sup>

## Sample size

In Sussex, the aim was to achieve close to 100 cases in each group in the study – pre and post pilot, and pre and post comparison. One hundred cases was chosen as a benchmark as it was felt, based on a previous study (Myhill, 2016), that this sample size would yield meaningful and statistically significant findings.

As statistical significance was achieved even with smaller sample sizes, and due to the time intensive nature of the data collection, the sample size for the 'pre' groups was revised down to 75. Data was collected for all days in randomly selected weeks; natural variation in the number of incidents that occurred in each randomly chosen sampling week meant that groups contained slightly more or less cases.

**Table 1. Sample sizes: Sussex**

	Pilot area	Comparison area
Pre-pilot period	84	73
Post-pilot period	112	108

Data collection in West Midlands had additional challenges owing to data being held across multiple systems, and researchers having to manually identify whether incidents were attended by officers in the pilot or comparison groups and find cases where a risk assessment had been completed. The main outcome measure for the study achieved statistical significance in West Midlands with approximately 50 cases per group. This smaller sample size was, however, biased in the post-pilot area towards cases involving ex-intimate partners. It is possible this bias may have influenced the comparison of disclosure of controlling behaviours, as the revised risk tool asks primarily about current behaviours (thus a victim who may have disclosed both historical controlling behaviour and current stalking behaviour using DASH may only disclose current stalking behaviour using the revised tool). Further data could be collected in

<sup>2</sup> The offending history of the perpetrator, and the nature of any previous incidents recorded between victims and perpetrator, are available to officers by searching police systems. Observations of officers responding to domestic abuse incidents suggested that time spent researching the history of the case and the parties involved can vary.

West Midlands to see whether a larger and less biased sample produces findings for disclosure of controlling behaviour similar to Sussex.

**Table 2. Sample sizes: West Midlands**

	Pilot area	Comparison area
Pre-pilot period	51	53
Post-pilot period	51	53

### Sample characteristics

Comparison areas in the participating forces were chosen to be as close as possible to the pilot areas in terms of population demographics and the volume of domestic abuse incidents recorded by the police. As the sampling for the study was not random, data on basic socio-demographics was collected for the people involved in sampled incidents as a check on how well matched the pilot and comparison conditions were (see Table 3).

In Sussex, the key difference was a greater proportion of female victims in the comparison area. This difference is almost certainly explained by a higher proportion of incidents in the pilot area involving family members as opposed to intimate partners. Whereas incidents involving intimate partners tend to involve primarily female victims and male perpetrators, family incidents tend to be more mixed as they involve abuse between siblings and parents and their (aged 16 or over) children. Similarly, this higher proportion of family incidents in the pilot area likely explains the higher proportion of female suspects in the pilot area.

In West Midlands, the only key difference between the samples was a much higher proportion of incidents in the pilot area involving former as opposed to current intimate partners (see above). This difference (which may have evened out with a larger sample size) is not likely to be problematic for the main outcome measure as, with intimate partner relationships, post-separation harassment tends to represent an attempt at continuation of ongoing coercive and controlling patterns of behaviour (and presents a similar risk).

**Table 3. Baseline sample demographics for pilot and comparison areas**

	Sussex		West Midlands	
	Pilot area %	Comparison %	Pilot area %	Comparison %
<b>Victim sex</b>				
Female	73.8	87.7	92.2	86.8
Male	26.2	12.3	7.8	13.2
<b>Suspect sex</b>				
Female	25.0	16.4	7.8	13.2
Male	75.0	83.6	92.2	86.8
<b>Victim age</b>				
16-25	27.4	24.7	21.6	34.0
26-35	27.4	23.3	49.0	32.1
36-45	20.2	20.5	19.6	17.0
46-55	14.3	16.4	3.9	13.2
56+	10.7	13.7	3.9	3.8
Unknown	0.0	1.4	2.0	0.0
<b>Suspect age</b>				
16-25	33.3	19.2	21.6	22.6
26-35	27.4	30.1	39.2	37.7
36-45	13.1	16.4	19.6	26.4
46-55	15.5	15.1	11.8	7.5
56+	8.3	15.1	3.9	5.7
Unknown	2.4	4.1	3.9	0.0
<b>Relationship between victim and suspect</b>				
Intimate partners	35.7	50.7	27.5	50.9
Ex-partners	34.5	37.0	64.7	35.8
Family	29.8	12.3	7.8	13.2
<b>Total cases</b>	<b>84</b>	<b>73</b>	<b>51</b>	<b>53</b>

### Process for coding and tests for inter-rater agreement

The process for allocating a risk grade, using either the DASH or the piloted risk tool, required officers to use professional judgement. The information gathered through the risk identification questions asked of the victim, be it those on the DASH or the piloted risk tool, is just one of several considerations for officers when deciding whether to grade a case as standard, medium or high risk. The nature and circumstances of the specific incident, and intelligence from police systems on past incidents and behaviours, may also have a bearing, to varying degrees.

As the researchers were required also to use judgement when coding cases, and in order to streamline the coding process, a set of coding guidelines was developed by the researchers (see Box 2). These guidelines were designed to

reflect the prominence afforded to coercive control in both the original DASH guidance (ACPO, 2009), and the Home Office counting rules for recorded crime<sup>3</sup>, which classifies coercive control as the 'principal crime' over any physical assault up to Grievous Bodily Harm (GBH) with intent. The guidelines reflected, in other words, the fact that officers should prioritise coercive control in their assessments of risk to the victim. The guidelines also, however, accounted for physical violence that might be harmful outside a pattern of ongoing coercive behaviour.

## **Box 2: Risk grading guidelines**

### **Standard risk**

- No evidence of coercive/controlling behaviour, or harassment or stalking behaviours
- May be a 'pattern' of incidents, but they appear to be similar 'situational fighting' type incidents that are not escalating or especially harmful
- Can include 'situational' physical assaults, but not:
  - Assaults where the level of force or injury is severe
  - 'One-off' assaults where the violence used is inherently coercive, such as choking or pinning somebody down
- Does not include any kind of sexual victimisation

### **Medium risk**

- Evidence of coercive/controlling behaviour
- Evidence of post-separation harassment or stalking
- No clear evidence of control, but perpetrator is a serial abuser and victim is likely to be minimising (based on, for example, disparity between their initial call to police and what they report to first responding officers, or differences between current and previous risk assessments).
- Physical assault where the level of force or injury is severe, or where the violence used is inherently coercive, such as choking or pinning somebody down
- Sexual victimisation

### **High risk**

- As medium, but with aggravating factor(s), such as:
  - Credible threats to kill/cause serious harm
  - Planned or recent separation that has raised the threat level/put the victim in danger
  - Significant escalation in e.g. severity of physical violence
  - Significant mental health or alcohol/substance abuse issues that increase the risk of serious harm

<sup>3</sup> <https://www.gov.uk/government/publications/counting-rules-for-recorded-crime>

All cases were double coded by the two researchers who undertook the case-file analysis, and any disagreements were resolved by discussion. Fewer than 5% of cases were reclassified from the original researcher's risk grading.

In order to provide a level of external quality assurance of the coding, a specialist support worker from Solace Women's Aid (a voluntary/charity sector organisation that provides support to victim-survivors) was recruited to dip sample and code 25% of cases. These cases were drawn from across all the data collection sites (pre- and post-pilot; pilot areas and comparison areas). This individual was a specialist support worker for a number of years before progressing to training specialist victim advocates. She had considerable hands-on experience of risk assessing cases of domestic abuse and with police risk assessments.

The support worker based her risk grading decisions on the same information available to the researchers, and was provided with the coding guidelines used by the researchers. Cohen's kappa was calculated to indicate levels of agreement between risk gradings allocated by the researchers and the specialist support worker<sup>4</sup>. The kappa value was 0.61, indicating a 'substantial' level of agreement (Landis and Koch, 1977). See Appendix 3 for further discussion on tests for inter-rater agreement.

## In-depth interviews

In-depth interviews were conducted with key individuals and groups involved in the pilot (see Table 4).

**Table 4. Numbers of interviews conducted**

	n
Frontline response officers	17
First line supervisors (Sergeants)	7
Secondary risk assessors (police officers and staff in central units)	6
Partner agencies (specialist support services)	4

Interviews took place after the pilot had finished, and lasted 20-30 minutes on average. All interviews were digitally recorded and transcribed professionally. Thematic analysis of the data was undertaken in Excel following the 'framework' method (Ritchie and Lewis, 2012).

---

<sup>4</sup> A simple calculation of percentage agreement has been widely discredited as a meaningful measure of inter-rater agreement, in part because it does not take into account the likelihood of raters working with relatively small samples agreeing in a proportion of cases based solely on chance.

## Observations

Researchers observed 17 shifts across the three forces. Observations were only undertaken with officers using the piloted risk tool as, firstly, the principal aim was to see how the piloted risk tool worked in practice and, secondly, observations of officers completing the DASH were undertaken during a previous project (see Robinson et al., 2016). Researchers typically accompanied officers on late shifts (approximately 15:00 to 23:00), as late shifts are likely to see the highest volume of domestic abuse incidents.

Researchers accompanied officers for the entirety of their shift. As well as observing officers conducting risk identification interviews at the scene, researchers were sometimes able to observe officers completing their incident reports and other (electronic) paperwork, including the risk assessment, back at the station.

Observers were required as soon as possible after the shift to write up a narrative account of the shift they had observed, including brief descriptions of calls for service and any other activity. For domestic abuse incidents, observers were required to complete a structured data collection template to record information relevant to the evaluation (for example whether/how a risk assessment was completed; officers' rationales for choosing a specific risk grade).



## 4. Findings

- Using the piloted risk assessment tool, first response officers' initial assessments of risk were less likely to be regraded, during a post incident review, than assessments made using DASH.
- Victims in one force disclosed perpetrators' coercive and controlling and stalking and harassment behaviours at greater rates using the piloted risk assessment.
- First response officers recorded proportionately more crimes of coercive control during the pilot, though numbers overall were still small.
- Although the pilot showed an increase in officers documenting coercive and controlling behaviour, that outcome did not always translate into them assigning an appropriate risk grading. This finding was most pronounced for post-separation harassment and stalking.
- First response officers and their supervisors generally said they preferred the piloted risk assessment tool to the DASH. Some specialist support workers preferred DASH, although they recognised its shortcomings in a policing context, while some preferred the piloted tool.

The evaluation considered the impact of the piloted risk assessment on officers' risk grading decisions, as well as their identification and recording of coercive control. It also considered the rate at which victims disclosed behaviour associated with coercive control.

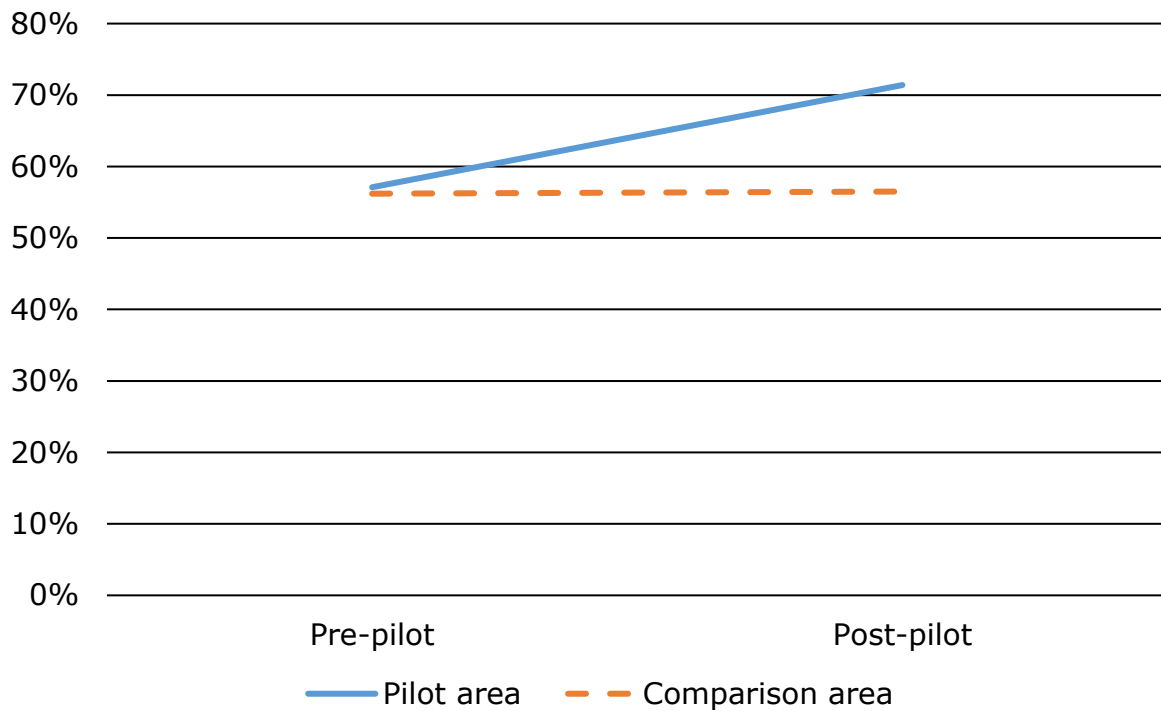
### **First response officers' primary assessments of risk**

The key outcome measure for the evaluation of the piloted risk tool was whether it resulted in officers reaching more appropriate assessments of the risk of harm to the victim. Specifically, whether the piloted tool allowed officers to recognise more consistently ongoing patterns of abusive behaviour. To determine this outcome, a sample of cases from randomly selected weeks during the pre-pilot and post-pilot periods were analysed, both for officers in the pilot area and for officers in comparison areas (see Tables 1 and 2 for sample sizes). Using the information collected with the respective risk tool (DASH or the piloted risk tool), information on the incident log concerning the original call to police, and the first responding officer's account and background intelligence on those involved, researchers assigned a risk grading (standard, medium or high) which was then compared to the grade assigned by the first responding officer.

Results indicated higher rates of agreement between researchers and first responding officers during the post-pilot period when officers used the piloted risk tool as compared with the pre-pilot period when DASH was used. In Sussex, during the pre-pilot period, researchers made the same risk grading decision as officers in 57% of cases. This level of agreement increased to 71% during the

pilot<sup>5</sup>. In the comparison area of Sussex, where the DASH was used throughout the pre-pilot and post-pilot period, the level of risk grading agreement remained consistent at around 56%.

**Figure 1. Percentage of cases where officers and researchers made the same risk grading decision, Sussex pilot and comparison area**

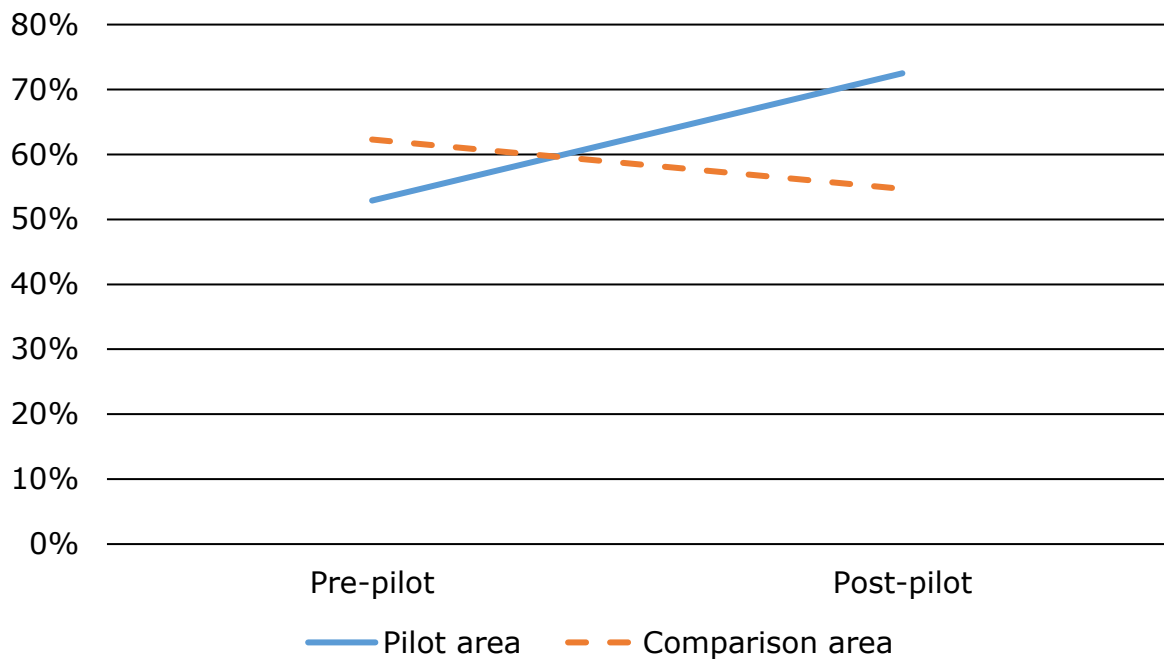


In West Midlands, a similar pattern was reflected. In the pilot area, researchers agreed with officers' risk grading decisions in a larger proportion of cases when the piloted risk tool was used (53% vs 73%)<sup>6</sup>.

<sup>5</sup> This increase was statistically significant;  $X^2(1, N=196) = 4.324, p = .038, \phi = .149$ .

<sup>6</sup> This increase was statistically significant;  $X^2(1, N=102) = 4.194, p = .041, \phi = .203$ .

**Figure 2. Percentage of cases where officers and researchers made the same risk grading decision, West Midlands pilot and comparison area**



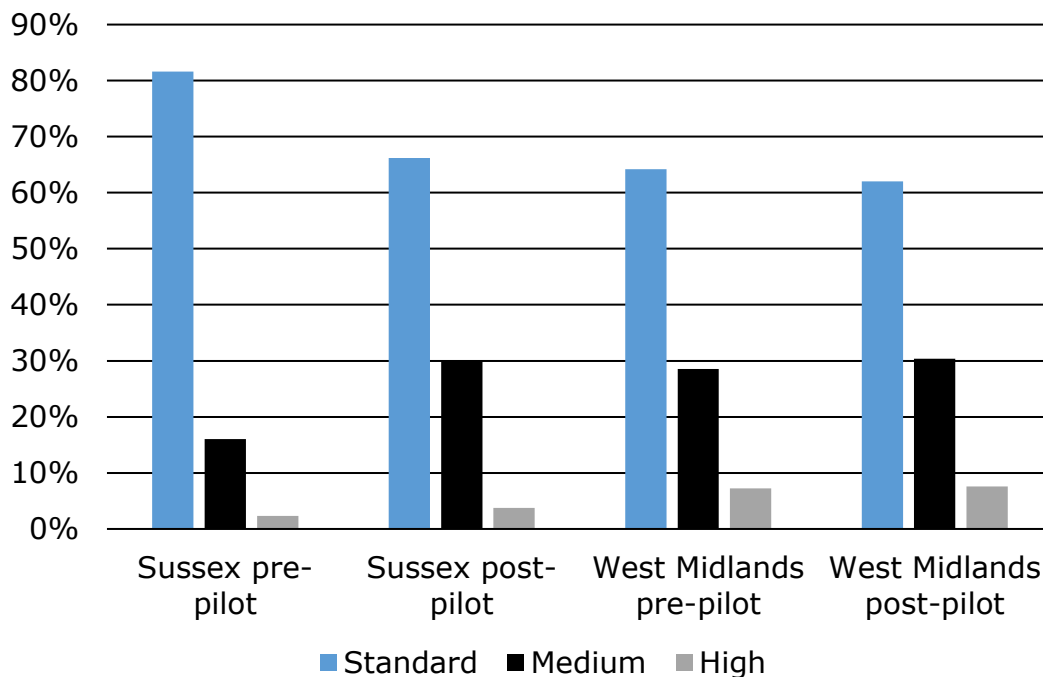
There was no improvement and no statistically significant difference between researcher and officer risk gradings pre and post pilot in the comparison area.

In order to confirm the results of the chi square analysis, a difference-in-difference analysis was conducted on pooled data from Sussex and West Midlands using logistic regression. This analysis suggested odds of agreement between officers' and researchers' risk gradings were 2.4 times greater in the pilot areas. The full model and findings are presented in Appendix 5.

## Changes to the distribution of risk gradings

Analysis of data provided by the pilot forces showed that, for the duration of the three month pilot period, there was an increase in the proportion of cases graded as medium risk in Sussex in the pilot area, from 16% of cases pre-pilot to 30% of cases post-pilot. There was a corresponding decrease in cases graded as standard risk, from 82% to 66%. It is likely this change was prompted by officers identifying patterns of abusive behaviour more frequently during the pilot (see below), and the piloted risk tool indicating that ongoing patterns of behaviour ought to be classified as medium risk (though it should be noted that officers did not, by any means, classify all cases where there was ongoing abusive behaviour as medium risk, indicating that they were still willing to apply their professional judgement).

**Figure 3. Risk grading in pilot areas of Sussex and West Midlands during the pre-pilot and post-pilot periods**



In West Midlands, there was a higher baseline proportion of cases graded as medium risk, and very little change in the distribution of risk gradings during the pilot period. There was, however, as described above, an increase in the level of agreement between officers' and researchers' assessments of the level of risk.

It is not possible to know for sure, but a potential explanation for the increase in medium risk cases observed in Sussex not being observed in West Midlands concerns the policy of discretionary risk assessment that existed prior to the pilot in the latter. In the pre-pilot period, officers in West Midlands were required to complete and submit a DASH primary risk assessment only in cases that involved criminal offences between intimate partners. In cases where they chose not to submit a full risk assessment, officers would allocate a risk grading of standard, medium or high based (we assume) on their assessment of the incident, the accounts of the people involved, and any intelligence from force systems available to them. Although the policy was one of discretionary use of DASH, in practice it appeared that officers very rarely submitted a DASH risk assessment in cases where a criminal offence was not present. For the duration of the pilot, then, officers were expected to conduct and submit a risk assessment also for cases recorded as non-crime incidents.

While in Sussex the movement of cases appeared to be primarily from standard to medium risk, it is possible that in West Midlands there was movement in both directions. Firstly, it is likely that, during the pilot, there were cases graded as medium that would previously have been graded as standard due to officers conducting a full risk assessment where previously they were not, and the pilot risk tool highlighting ongoing abuse in those cases. This view was in fact expressed by two supervisors in West Midlands. It is also possible, though, that there was a shift during the pilot that saw some cases previously graded as

medium being graded as standard; especially cases where there was a criminal offence, but no clear ongoing abuse. Previous research has shown that officers prioritise criminal offences, and particularly physical violence with injury, in their assessments of risk of harm (see Robinson et al., 2016; Myhill, 2017). Yet not all incidents involving discrete criminal offences may be representative of a wider pattern of controlling and coercive abuse (see Myhill and Hohl, 2016). This view was also supported by a supervisor in West Midlands, who suggested officers were more confident in grading cases as standard that they might previously have graded medium partly to “cover their own backsides”.

“I think as I said I think because the definitions are there they’re happier to call it standard whereas they weren’t before.” (Sgt)

## Officers’ perceptions of their risk decisions

While some of the first response officers interviewed did not think the piloted risk tool changed the way they graded risk, others felt it helped them to better identify risk and that the risk grading was more “clear-cut” and “spelt out” through the tool:

“I do not think the previous form really led you on to ... thinking about the risk as much ... like I say, it was more of a gut feeling that you ended up with.” (PC)

The question on the victim’s own perception of likelihood of threat of serious injury was identified by a couple of officers as resulting in more cases being graded as medium risk as opposed to standard:

“Sometimes although people aren’t particularly forthcoming and they might not be supportive of police involvement, but they’ll generally give you quite a genuine honest answer to that one. And if they’ve said that these things happen occasionally, and then they think that the risk of serious harm is an eight, obviously something’s going on, maybe they’re not telling you about.” (PC)

There was a sense among supervisors that first response officers graded risk more consistently with the piloted risk tool. Similarly, the majority of secondary assessors interviewed felt the piloted risk tool led to an improved initial grading of risk by officers:

“We tend to alter less of the new forms. They do tend to get the grading right on the new pilot.” (Police secondary assessor)

It cannot of course be ruled out that officers were making a greater effort in their assessments of risk due to their knowledge that a new tool was being tested and evaluated.

## Disclosure, identification and recording of coercive control

The piloted risk tool was designed to try to highlight the frequency and intensity of coercive and controlling behaviours and to move first responding officers away from an incident-focused mindset so as to be better able to recognise dangerous

patterns of behaviour. It was also hoped that by asking victims questions that acknowledged in the first instance the likelihood that abuse was occurring (as opposed to asking them whether it was occurring or not), and asking them to specify how frequently they experienced these behaviours, it would make it easier to disclose abusive behaviour to officers.

## Victims' disclosure of coercive control

Part of the reason that some officers felt that the piloted risk tool helped them to better identify coercive control may have been the result of victims disclosing more abusive behaviour when officers used that tool.

Case-file data was used to compare similar questions on the piloted risk tool and the DASH that focused on coercive and controlling behaviours to see if there was any difference in rates of disclosure of those behaviours<sup>7</sup>. In Sussex, this analysis suggested victims were more likely to disclose controlling behaviour, more likely to say they had been subject to stalking and harassment type behaviours and more likely to say they felt isolated when the piloted risk tool was used.

In the pilot area of Sussex, 42% of victims disclosed control during the pilot, compared to 25% of victims during the pre-pilot period when the DASH was used<sup>8,9</sup>. A similar increase was not evident in the comparison area.

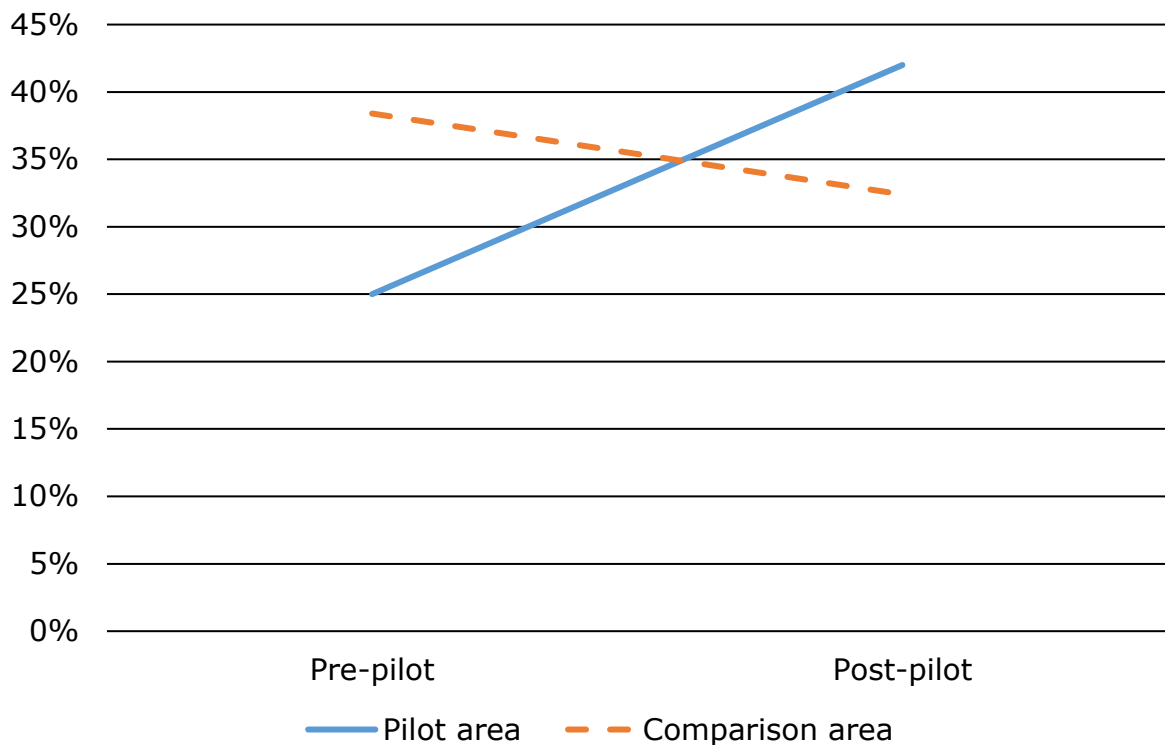
---

<sup>7</sup> Response options of 'occasionally', 'often' and 'all the time' on the piloted risk assessment were combined to indicate if the behaviour was disclosed. This was compared to the number of 'yes' responses to the equivalent DASH question.

<sup>8</sup> This increase was statistically significant;  $X^2(1, N=196) = 6.097, p = .014, \phi = .176$ .

<sup>9</sup> The question in the piloted risk assessment was 'How often does (...) control your daily activities, such as who you can see or how you perform household tasks?' The comparable question on DASH was 'Does (...) try to control everything you do and/or are they excessively jealous?'

**Figure 4. Percentage of cases where controlling behaviour was disclosed in Sussex Police pilot and comparison areas**

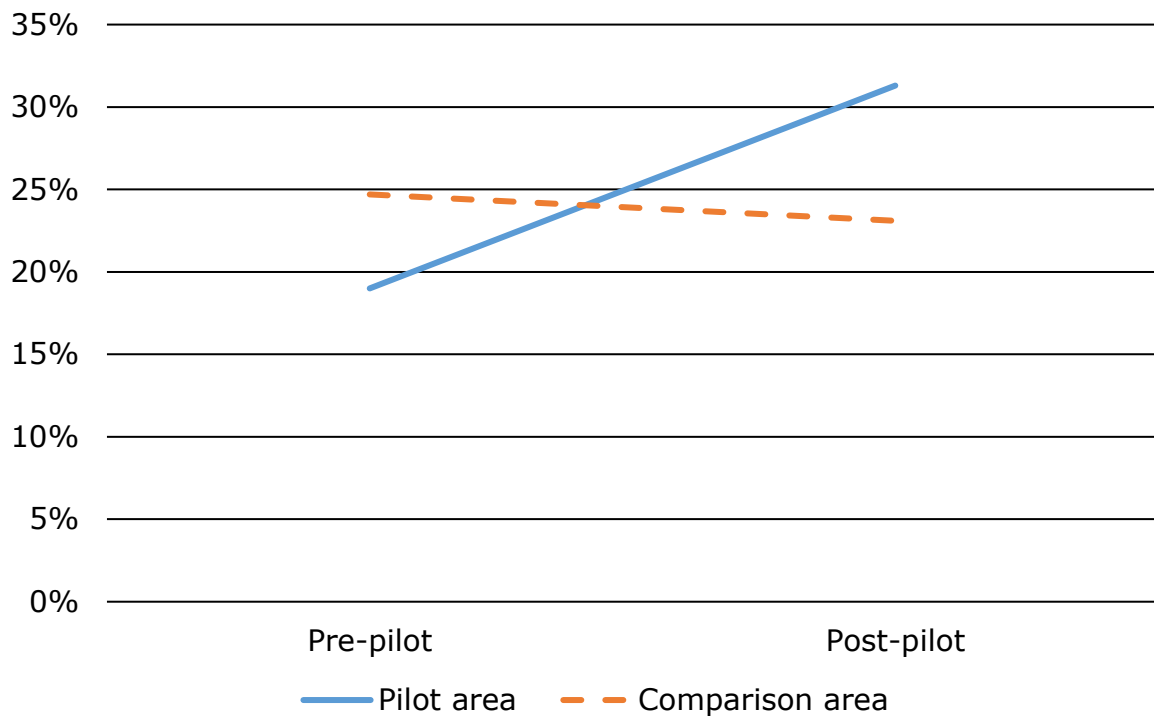


An increase, albeit one that just failed to reach statistical significance, was also seen in Sussex relating to the disclosure of stalking and harassment when the piloted risk assessment was used: 19% of respondents said they had been subject to stalking and harassment during the pre-pilot period, compared with 31% during the pilot<sup>10,11</sup>. Again, this increase was not replicated in the comparison area of the force.

<sup>10</sup>  $\chi^2(1, N=196) = 3.713, p = .054, \phi = .138$ .

<sup>11</sup> The question in the piloted risk assessment was 'How often does (...) follow or stalk you, or try to contact you when you do not want them to?' The comparable question on DASH was 'Does (...) constantly text, call, contact, follow, stalk or harass you?'

**Figure 5. Percentage of cases where stalking and/or harassment was disclosed in Sussex Police pilot and comparison areas**



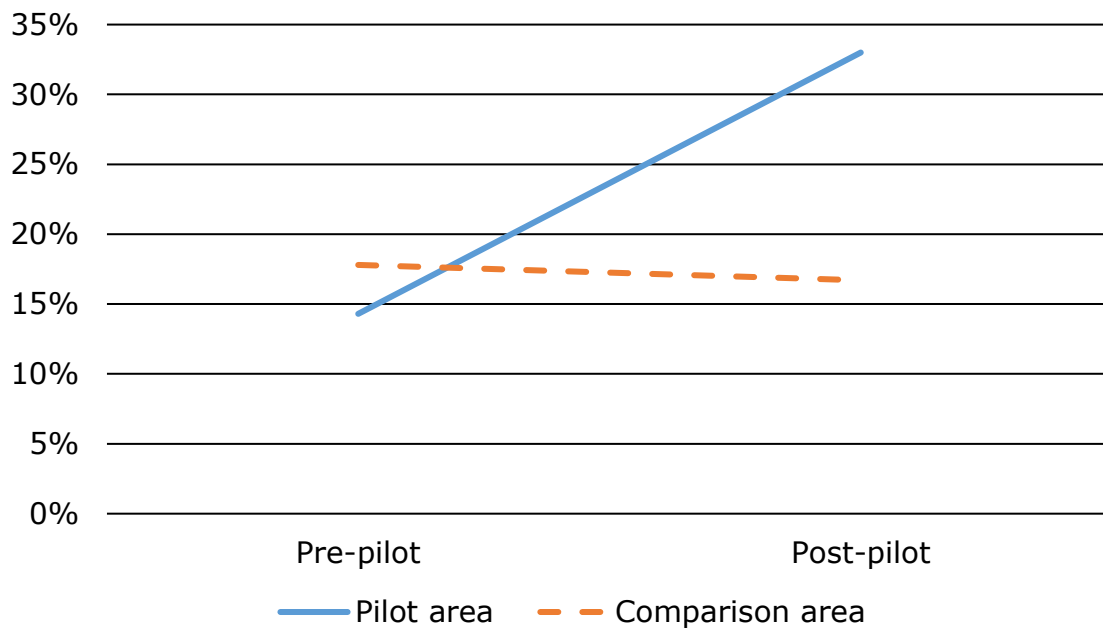
A relatively large and statistically significant increase was seen in the percentage of victims who said they felt isolated when asked using the piloted risk assessment, compared with the DASH. In the Sussex pilot area, 14% of victims said they felt isolated during the pre-pilot period compared with 33% during the pilot<sup>12,13</sup>. Again, there was no increase observed in the comparison area.

<sup>12</sup> This increase was statistically significant;  $X^2(1, N=196) = 9.000, p = .003, \phi = .214$ .

<sup>13</sup> The question in the piloted risk assessment was 'How often do you feel isolated or like you have no one to turn to for support?' The comparable question on DASH was 'Do you feel isolated from family/friends i.e. does (...) try to stop you from seeing friends/family/Dr/others?'



**Figure 6. Percentage of cases where victim isolation was disclosed in Sussex Police pilot and comparison areas**

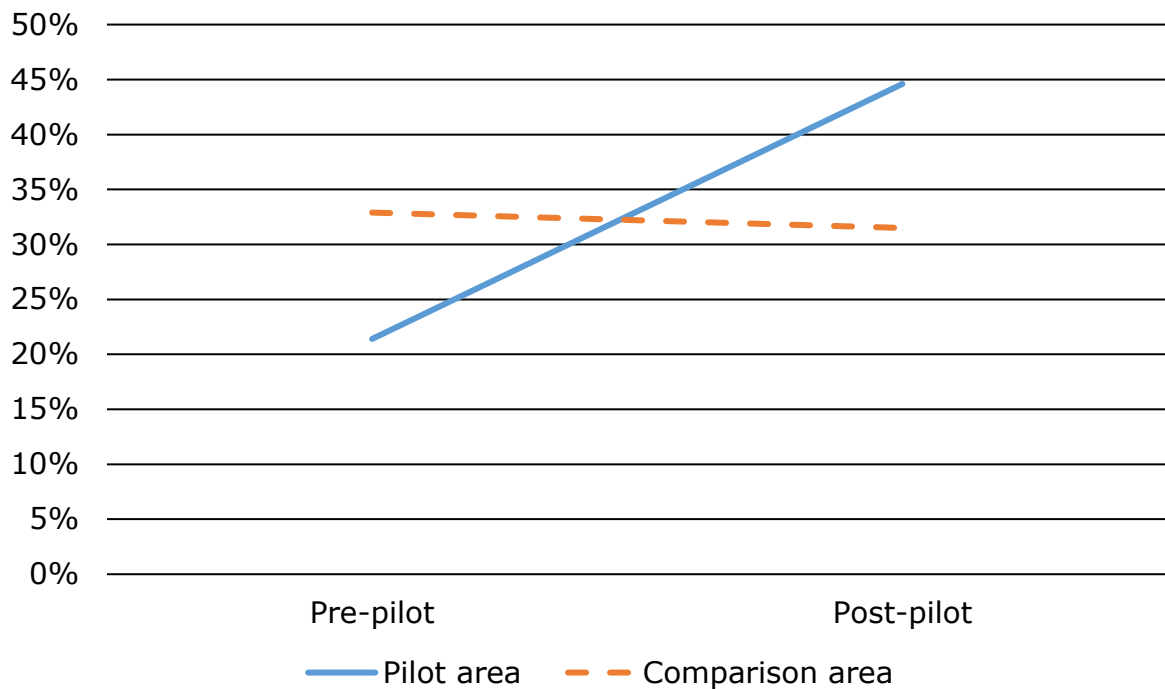


Interestingly, more than twice as many victims in Sussex stated they felt the abuse they were suffering was getting worse when asked during the pilot period (45%) compared to the pre-pilot period (21%)<sup>14,15</sup>. This finding was perhaps more surprising as the wording and response options for this question were the same on the piloted risk assessment as on the DASH. It might, therefore, be as first response officers described – the piloted risk assessment helped victims to identify more easily the abuse they were suffering.

<sup>14</sup> This increase was statistically significant;  $X^2(1, N=196) = 11.417, p = .001, \phi = .241$ .

<sup>15</sup> The question in the piloted risk assessment was 'Is the abuse you are experiencing from (...) getting worse?' The comparable question on DASH was 'Is the abuse getting worse?'

**Figure 7. Percentage of cases where victims said the abuse they were suffering was getting worse in Sussex Police pilot and comparison areas**



There was no difference in whether the victim stated the perpetrator threatened or attempted suicide or whether the perpetrator had made credible threats to kill, both of which have been identified as indicators of coercive control.

The increases in disclosure of abusive behaviours in Sussex were not observed in West Midlands. In fact, the proportion of respondents disclosing controlling behaviour specifically decreased from pre to post-pilot, though none of the comparisons were statistically significant. It is likely that these non-significant findings were the result of a smaller sample size, and that the apparent decrease in disclosure of controlling behaviour due to the bias in the pilot sample towards cases involving ex-partners (where controlling behaviour will have been superseded to a large degree by stalking and harassment).

### Victim’s own perception of risk

Research has shown that victims’ own perceptions of risk of further harm are as accurate as some structured risk assessment tools (see Wheller and Wire, 2016). Victims are well placed to know the imminent and future threat posed by the perpetrator. A question was therefore included on the piloted risk tool that asked victims on a scale of 0-10<sup>16</sup>, how likely they thought it was that their abuser would seriously injure them in future.

There were conflicting views on the utility of this question among first response officers. Some felt it was “the most useful thing” and “quite a good barometer”:

<sup>16</sup> Where ‘0’ is not at all likely and ‘10’ is very likely.

"I like the grading one to ten. It gives you a better understanding about how they're feeling ... and what effect the behaviour is having on their life." (PC)

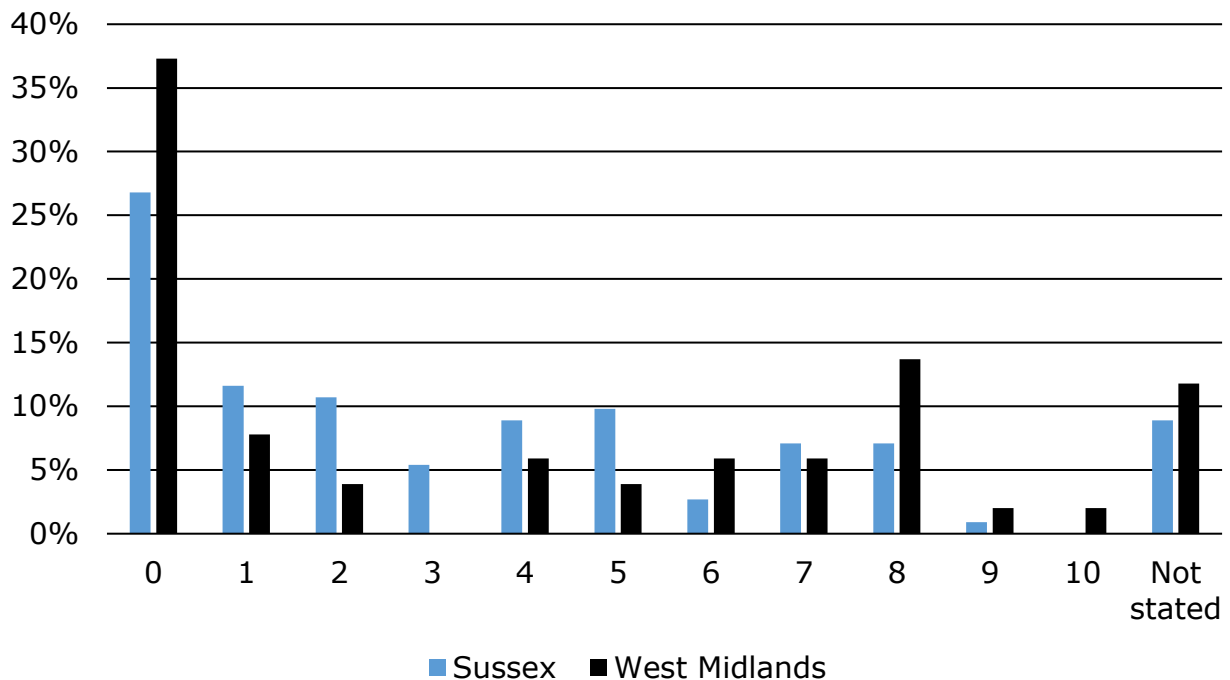
Others however, felt it was problematic:

"The one thing that I'd say I don't like about it ... is this 0 to 10 scale, because what I've found is people ... obviously you ask them it and very often people go ten, and you know that that isn't really what the situation is, they're not really terrified they will get a serious injury. But I don't know if people sometimes think that that means they're going to get, not more is going to happen to the offender, or that's what, I think they just put that to get that reaction a lot of the time." (PC)

"I've always found that it's been a difficult one for anyone to answer. It's sort of saying to someone, "Well, how long's a piece of string?" And they've always struggled and so most of the time they're either going with not stated or it will be really low or really high, depending on what the incident is. I don't think it gives a true reflection." (PC)

The notion that most victims give a high score at this question is not supported by the data (see Figure 8). Of the 112 pilot cases in Sussex, not a single victim said '10' to this question and only one did so out of 51 cases in West Midlands. While around a quarter of victims in Sussex (27%) and over a third in West Midlands (37%) said it was 'not at all' likely they would be seriously injured, officers were given guidance that a low score on this scale may not necessarily reflect the actual threat (due to victims potentially minimising the severity of the abuse, or it becoming 'normalised') and should be considered in combination with other information and responses.

**Figure 8. Responses to Q16: 'On a scale of 0 to 10, where '0' is not at all likely and '10' is very likely, how likely do you think it is that (...) will seriously injure you in the future?'**



Supervisors and secondary assessors were generally more consistent in their view of the utility of this question:

“The sliding scale with regards the one to 10 on the back, with regards how likely you think it is that they’ll severely injure you, that I find quite useful, because that pares them down ... They tend to give you an honest opinion on that one, I think ...” (Sgt)

“I liked the scale of 0-10 ... it gives us an indication because if they’re saying ‘all the time’, ‘all the time’, ‘all the time’ to all the questions but then saying no I don’t think he’ll seriously injure me it might say they’re so used to it, it’s such a norm in their life they don’t consider themselves at risk but when you read it clearly they are really at risk of serious injury.” (Police secondary assessor)

### Officers’ identification and recording of coercive control

It was suggested by some first response officers that the piloted risk tool enabled them to “get more of an understanding of the relationship as a whole” and look more widely at what had been happening:

“... what we’ve been guilty of in the past is assessing an incident on its current basis and looking at something as it presents there and then. What the new risk assessment forces you to do is look at the bigger picture.” (PC)

More specifically, the majority of first response officers interviewed, as well as their supervisors and secondary assessors, felt the piloted risk tool made it

easier to identify coercive and controlling behaviour as they felt it included specific questions that highlight it, such as, 'How often does (...) call you names, humiliate or degrade you?', and put it more at the "forefront". It was also felt that the piloted risk tool could help victims to identify to themselves when they are being controlled:

"... but I think it's made it a bit easier, as I said, in terms of someone who is at risk, perhaps, opening up a bit more about what someone in a relationship would probably see as more minor things ... but if someone sees it as more of a normal thing, I think ... it just helps someone speak up about it. Whereas, before with a more blunt question, they might have gone, "no, I don't think they're controlling me." They'll give details about things they wouldn't have done otherwise." (PC)

"You're identifying to yourself and to them that they're the victim of coercive control because when they're deep in it they're just ... to coin a phrase, they just can't see the wood for the trees, you know..." (PC)

Those who did not think it was easier to identify coercive control with the new risk tool did not feel it was any worse than the DASH. Instead, these officers all mentioned how it was not the risk tool in and of itself that helped to identify such behaviours; it was more to do with "speaking to the victim", what you "gauge when you have a conversation" and "being with the person".

Keyword analysis of officers' incident reports and risk assessments indicated a greater tendency to identify and document coercive control – or the absence of coercive control – during the pilot. In Sussex, officers in the pilot area mentioned 'control' in only four out of eighty-four (5%) incident reports in the pre-pilot sample; control was not mentioned at all in the freetext sections at the end of the DASH. In the first eighty-four cases in the pilot sample, where officers used the piloted risk tool, control was mentioned twelve times in incident reports and twenty-three times in the freetext section of the risk assessment. Taking into account cases where an officer mentioned control both in their incident report and on the risk assessment (n=7), control was mentioned in twenty-eight (33%) of cases.

In West Midlands, officers in the pilot teams mentioned control in only one out of fifty-one (2%) of incident reports in the pre-pilot sample; control was not mentioned at all in the freetext sections at the end of the DASH. In the fifty-one cases in the pilot sample, officers mentioned control in five incident reports and eleven risk assessments. Taking into account cases where the same officer mentioned control in both the incident report and the risk assessment (n=4), control was mentioned in twelve (23%) cases. Once again, the possibility that the novelty of a new tool and pilot project, in addition to a training input, contributed to these findings cannot be entirely discounted.

It is important to note, however, that officers identifying controlling behaviour did not necessarily result in them making an appropriate assessment of risk. Analysis of officers' incident reports and risk assessments revealed in some instances a lack of understanding of coercive control that was reflected also in the DASH review (Robinson et al., 2016). Some officers appeared not to appreciate the risk posed by ongoing abuse and highly controlling behaviour. Some also appeared to regard the risk to the victim as lower if there were no assaults or criminal offences reported at the current incident, or if the violence

used was not severe. The following is a summary of an officer's report of an incident they graded as 'standard' risk:

Male and female have known each other for eight years and have been together for one year. She says his behaviour is jealous and controlling. He checks up on her, checks her phone, and will ask for photos of what she is wearing before she goes out. He called her at the beach to ask how many men were there and whether she was wearing a bikini. He got angry in a bar when she said thank-you to the barman. He has grabbed her by the arms a number of times. His name calling is the most hurtful (slag, c\*\*\*, whore). This incident refers to him slapping her on the head when she refused his marriage proposal. Her main concern is that in ending the relationship he may continue to contact and harass her. She believes he has abused previous partners, but no trace [force systems] or PNC.

This lack of understanding was especially marked in relation to cases of post-separation stalking or harassment. In particular, there appeared to be, across the sites, a tendency for some officers to regard the risk of harm to the victim as intrinsically lower if they had separated from the perpetrator, or were no longer living at the same address.

"Due to the fact the relationship is over there are no control issues."

"The suspect has no control of the victim as they are no longer together or live together."

## Recording coercive control as a crime

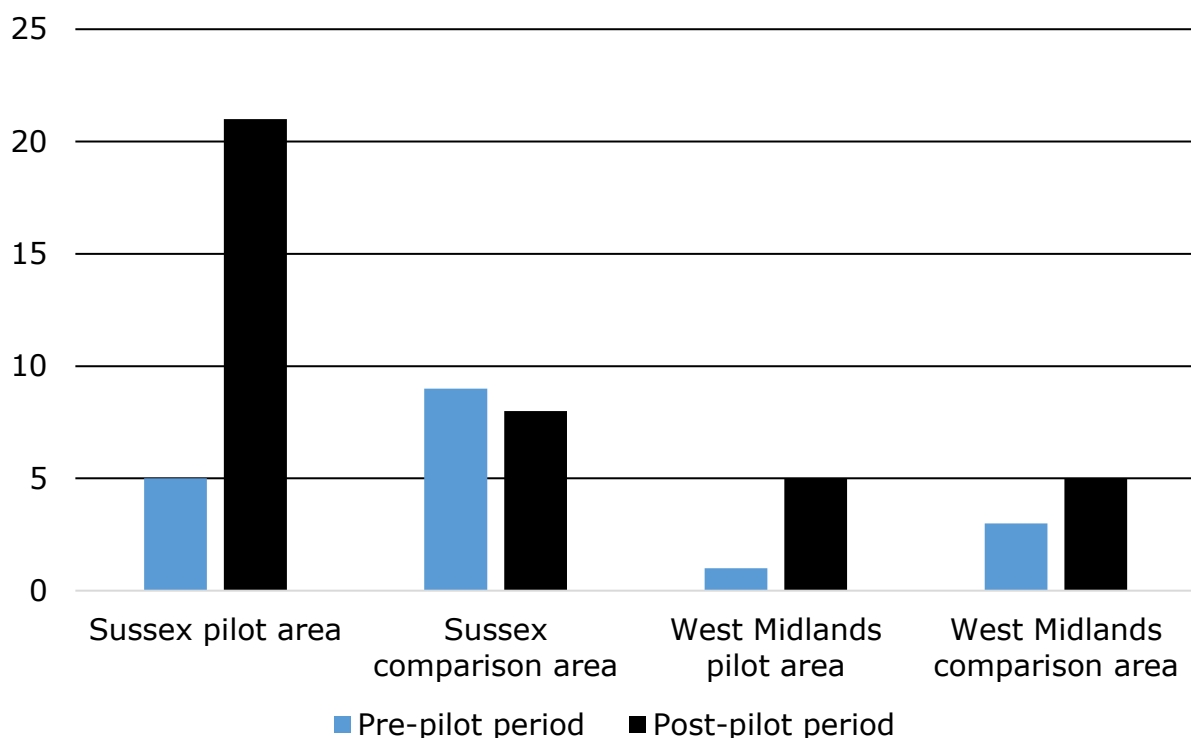
While officers felt the revised risk tool enabled better identification of coercive and controlling behaviours, and victims were disclosing such behaviours at a greater rate in one of the forces, data from across the pilot and comparison areas of both Sussex and West Midlands showed that despite an increase in the number of cases recorded as engaging in controlling/coercive behaviour in an intimate/family relationship the overall numbers were still very low. Only five offences of controlling/coercive behaviour were recorded in the pilot area of Sussex Police before the pilot; this number increased to 21 offences during the pilot when the revised risk tool was used. No corresponding increase was seen in the comparison area. In the pilot area of West Midlands only one offence was recorded during the pre-pilot period, rising to five cases during the pilot.

Analysis of officers' incident reports revealed numerous examples of cases where coercive control had been documented, and where controlling and coercive behaviour should have been recorded, but where either a non-crime incident was recorded or an offence over which coercive control should have taken precedence. The following example is a summary of key details from an officer's incident report for an incident recorded as 'threats to damage property':

Throughout their five year relationship female states male has been both physically and mentally abusive, with another police force being called to their address on several occasions. When the relationship first started female was living in hostel, but male took her to Kent, she believes to get her away from her family. States male is very controlling and has been throughout their relationship. He has constantly checked her phones and social media for messages and calls, doesn't like her to have male friends, and doesn't like her leaving home without him being there. Controlled the way she dressed and didn't like her wearing anything 'girly', and insisted

on her wearing high neck tops. Has her own bank account, but on several occasions she was made to give up some of her money. Throughout the relationship, male didn't want female to have contact with family and she was not allowed to give out her phone number. Would check her internet history every day to see if she had been talking to them. Regularly called her fat and ugly and said that she was no good. Due to this abuse she was prescribed anti-depressants, but when male found out he flushed them down the toilet.

**Figure 9 Number of offences of engaging in controlling/coercive behaviour in an intimate/family relationship in Sussex and West Midlands Police pilot and comparison areas before and during the pilot**



Further investigation would need to be carried out to determine the reason(s) behind the low-level of recording of coercive control, but a number of issues were discussed during the interviews. There was still a sense among a couple of first response officers that it was their role to deal specifically with the incident at hand and so while it might be possible to identify coercive and controlling behaviour through the tool, it would be more appropriate for specialist officers to "pick it up" and deal with it. Some supervisors felt there was a continued lack of understanding around coercive control among officers, and one mentioned "knock backs" from the central team responsible for applying a crime classification to an incident who they said would "want to crime an assault when the coercive behaviour is the more serious offence". On such occasions, it was thought control would still be recorded in the first response officer's crime report. Indeed, the researchers saw numerous examples across the forces of cases where coercive control was highlighted by either a first responding officer or by a specialist/investigator, but where the classification was a non-crime incident, and of cases classified as physical assaults or other offences where

coercive control could and should have been recorded as the 'principal' offence in line with the Home Office counting rules for recorded crime.<sup>17</sup>

## Stakeholder perceptions of the piloted risk tool

The review of DASH highlighted inconsistencies in the application and delivery of DASH, and found widespread support for a more focused tool for the frontline (Robinson et al., 2016). The piloted tool was designed with the intention of providing a straightforward and structured method of obtaining information from victims to inform officers' risk assessments. But although the piloted tool has demonstrated improved outcomes in relation to the identification of coercive control and the appropriateness of risk grading, was it viewed in a positive light by those using it?

### General views

First response officers, their supervisors and police officers and staff responsible for secondary risk assessments reflected positive views of the piloted risk tool, and generally preferred it to the DASH. While supervisors viewed the DASH as a 'necessary' process in responding to reports of domestic abuse, they felt it contained too many questions for an initial assessment of risk; words such as 'repetitive' and 'cumbersome' were used to describe the DASH. By contrast, supervisors, secondary assessors and first response officers felt the piloted risk tool was more specific and 'user-friendly'. The piloted tool was described as being easy to use, quicker to complete and with questions more relevant than some of those contained in the DASH:

"No I think it's alright ... You know, it's another form so you know ... and we're form-heavy so you know, but it is quicker and it is more relevant I think." (PC)

"The questions are more up to date as well, including social media and things like that into the questions. It has brought it more up to date." (PC)

"I think it's more user-friendly for the officers as well. I think if it's more user-friendly for the officers then we'll get a better completed form."  
(Police secondary assessor)

There were mixed views on whether the structure meant the piloted risk tool had to be asked verbatim, with one officer saying they felt as though they were "surveying" victims rather than speaking to them, whereas another believed the piloted tool "lends itself to being encompassed into a conversation as opposed to a form filling exercise".

### The question response options (frequency scale)

Similarly, the response options drew mixed feedback. Most first response officers spoken to said neither they nor the victims they dealt with had had any issues with the response options. A couple of officers did, however, feel that the options are subjective as "somebody's occasionally, could be somebody else's

---

<sup>17</sup> <https://www.gov.uk/government/publications/counting-rules-for-recorded-crime>



often" and that if something had happened once, it was not 'never' nor 'occasionally'. As secondary assessors will typically be reliant to some extent on what the officer has written on the risk assessment, some said they would have liked officers to provide more clarity around the responses (exactly when and how often something was happening, for example).

A specific issue with response options based on frequency is that of the 'one-off' physical assault. While in some cases (including some sampled for this pilot) victims may be minimising their abuse by saying they have never experienced previously physical violence, the wide ranging national definition of domestic violence and abuse means it is likely that at least some proportion of incidents that come to the attention of the police will have involved an isolated act of physical violence. In addition, previous research with survivors has indicated that for some it is the first physical assault that prompts them to report ongoing abuse to the police (Kelly, 1999).

## The single freetext section

Supervisors suggested that having a single 'general assessment of harm' freetext section at the end of the piloted risk tool was beneficial. Within this box, officers were asked to provide a summary of the risk posed to the victim by the perpetrator. Supervisors noted benefits of officers completing this both for the officer and for themselves in their role reviewing the completed assessments:

"A lot of the time they write a little bit of a background, then what's happened today and then their thoughts, which I think is really good, because it makes them think about it a bit more." (Sgt)

"... it gave the frontline officers the opportunity to rationalise it more which then made my job of reviewing it easier." (Sgt)

While not all of the first response officers expressed views as positive as the supervisors, only one said they did not like it (as they would have preferred a structured format to freetext). Others identified benefits of having to provide their risk grading rationale:

"I think it's quite good because it's meaning you've got to think about it. You're not just doing a tick box and saying "Yes, no, yes, no, yes no" and sometimes on the old form your rationale and things would be within the little comments filled at the end of those tick boxes. You're actually made to think about this and I don't ever leave it blank. I'll always put something in ... after everything else you've read, this is what I think, these are the reasons why, and that's why I've scored as I have." (PC)

"Well it just gives you this ideal because you just write down how you've got to where you are. How your rationale is at the job. It's now given you and you've got to fill in a bit of how you've got there which you didn't really have to do last time and it's just your thought process so if it does go wrong you can say well hang about, there's my thought process, that's how we've got to that on what was presented at the time because it's all through there and what was presented here ... so it's another safety tool

for them but it's a safety tool for us as well to say I got there because this and there you are, it's written down." (PC)

Examples of freetext rationales for different levels of risk, paraphrased from the original risk assessments, are presented in Box 3. It should be noted that detailed, coherent rationales were not provided in many cases, and that some 'rationales' resembled a 'cut and paste' of the circumstances of the incident from the officer's incident report. The quality of freetext rationales would likely be an issues for training and ongoing supervision, but the design of the piloted tool at least ensured the officer's rationale, where present, was attached to the risk assessment, as opposed to hidden somewhere in their incident report or subsequent update report(s).

### **Box 3: Example freetext rationales**

#### **Standard risk**

- Female was in a relationship with male for a year on and off ending some weeks ago. No previous [recorded incidents] between the two. Police were called by female's mother as she thought male was outside the house making noise although not committing offences and not being rude or abusive. Male did tell police he was outside. He said it was because he was drunk and wanted to get the female's attention as they had been texting since they broke up. Female has been advised that if she wants to stop contact from male then she needs to stop texting him. There are no threats and no harassing behaviour.
- This is the first time the victim has contacted police regarding the male. She has come home to find her husband with new girlfriend in the premises. Victim is aware divorce proceedings are in place. No physical violence used in this incident, just a verbal argument. Male has gathered belongings and been taken to a friend's address nine miles away. Victim has answered never to almost all questions, and perceived threat of severe injury is zero.

#### **Medium risk**

- At this stage the victim was asking for advice and reassurance regarding support should something happen in the future with her husband. There has been violence, however the last report was over 12 months ago. The concern is that the husband does seem to have a large amount of control over the victim. She does have her own money and can leave the house, but he will constantly text and phone her to see where she is. The victim is non-supporting of prosecution as she is fearful of repercussions. She has a good support network who will contact the police should the risk increase, or there be further assaults. The house is in both their names and she does plan to leave her husband. She is seeking legal advice next week regarding divorce. There is no immediate risk to victim, and medium risk would be appropriate.
- Offender used to push or slap her every few weeks when they were together. Used to check her phone often. There is conflict over child contact; victim wants a contact plan. One previous ABH, two previous non-crime incidents, and one previous child abuse non-crime. There is a pattern of abuse/violence, but this is only likely to result in serious harm if there is a change in circumstances.

### High risk

- I believe the level of risk posed to the family by the daughter is high based on the repeated comments that she wants them dead and that she is going to stab them; this appears to be aggravated by her concealing a weapon this evening and making the comments leading to her father having to arm himself in preparation to defend himself.
- The victim does appear to be scared, as she is unsure whether the threats that she received today will be carried out. The victim states that he is very controlling even though they have separated. They have never lived together and she does not allow him in the house when the children are there due to his unpredictable behaviour. He is described as being insane and is convinced that the victim is with other men. He calls and facetimes her, where he makes her prove she is in her house by getting her to walk around. He has turned up at her address early in the morning after she has been out with friends demanding to know where she has been and who she has been with. They were in a relationship on and off for about two years. He has previously pushed her and strangled her by pushing her up against a wall and putting his hands around her throat. She states that these incidents were unreported. The victim hasn't blocked his number due to the fact she would rather him ring her than turn up at her house. He has recently moved to within a ten minute walk from her house. He is addicted to cocaine and uses other recreational drugs. He is paranoid and convinced the victim is with other men.

Support workers interviewed also suggested that the freetext box on the piloted risk assessment tended on average to be completed more thoroughly and consistently than freetext sections of the DASH. It cannot be known for sure whether this finding was the result of officers being prepared to complete a single freetext section more thoroughly, and embracing as intended the direction to justify their risk decisions, or whether there was a 'Hawthorn effect' (officers making more effort because they knew they were involved in a pilot that was being evaluated).

### Incidents involving ex-partners

An issue that became clear when testing the revised risk assessment with survivors was that if they were no longer in a relationship with the perpetrator they might answer 'never' to some questions that at the time they were with the perpetrator they would have answered 'often' or 'all the time'. This finding is perhaps what prompted some of the specialist support workers interviewed for the evaluation to suggest the piloted risk tool was better at highlighting what is happening currently as opposed to the history of the abuse.

If the piloted risk tool were to be tested further or adopted by forces, any accompanying training and guidance would need to make clear that if a victim answers positively to the question relating to post-separation harassment and stalking (q6) but not many other questions they may still be at significant risk of harm. Officers should also be encouraged to record in the freetext section any

significant controlling behaviour or other abuse that occurred prior to separation from the abuser, especially if that abuse was never reported to the police.

Indeed, the police have been found to struggle more widely with responding to cases of post-separation harassment and stalking (see HMIC/HMCPSP, 2017). There was a strong theme in both pilot and comparison areas, and across the three forces involved in this pilot, for some officers (and indeed certain supervisors) to regard separation and/or not living at the same address as the perpetrator as reducing the risk to the victim. Clearly, there is a limit to what any structured risk assessment tool can achieve if officers lack understanding of the potential threat posed by (jealous and controlling) ex-partners.

## Incidents involving family members as opposed to intimate partners

The piloted risk assessment contained relatively few questions that might be considered not to be applicable to incidents involving family members as opposed to intimate partners (q7 - control through child contact, and q13 - separation). As with the DASH, officers could use their judgement to not ask these questions if they were not relevant to the situation at hand. Some first response officers still felt, however, that the piloted risk tool was less appropriate for incidents involving family members than for those involving intimate partners. While it is true that coercive control is more prevalent in intimate partner relationships, there are family relationships in which similar dynamics are evident, and family relationships are covered by the coercive control legislation in England and Wales.

## Sexual victimisation

Based on the balance of feedback received while developing the piloted risk tool, the decision was taken not to include a question relating to sexual victimisation (see Appendix 1). Very few police interviewees spontaneously mentioned the absence of this question, though one first response officer felt a question on sexual victimisation was needed to encourage disclosure, as did three of the six secondary assessors who were interviewed. Some of the specialist support workers interviewed also suggested they would have liked to have had a question on sexual victimisation.

The principal argument voiced in support of the inclusion of a question on sexual victimisation was that it might be the only time a victim is asked. The principal arguments against its inclusion were that comparatively few victims disclose sexual victimisation to first responding police officers (and indeed it is frequently the last thing to be disclosed even to specialist support workers<sup>18</sup>), and that sexual victimisation is unlikely to be occurring without other key indicators of coercive control that are asked in the revised risk tool (and identify in the first instance whether there is risk of harm). In addition, it would be very difficult if not impossible to design a question that captures adequately the complex and interrelated web of sexual abuse, coercion, assault and other forced sexual activity that typically comprises sexual victimisation in the context of coercive control (see Bagwell-Gray et al., 2015). And specialist support services and

---

<sup>18</sup> Based on conversations with both SafeLives and Professor Liz Kelly during the development of the revised tool.

academics consulted as part of the development of the tool felt a specialist would be best placed to engage a victim about these experiences.

Some victims may not be subject to a secondary risk assessment, however, either by the police or specialist support services. It is likely therefore that the inclusion or otherwise of a sexual victimisation question in a primary risk assessment will continue to prompt debate.

## Response from victims

No negative responses from victims to the piloted risk tool were reported by first response officers, or witnessed during the observations. Conversely, officers felt it to be “more victim friendly” as while it still asked personal questions, it did not go into “areas that could make them feel uncomfortable at the time” (a likely reference to the sexual victimisation question).

The piloted tool only included those questions that the research evidence suggested were key to identifying risk of harm and for the purposes of making an initial assessment. While first response officers noted this made it quicker to complete, due to the length and the way the questions were worded several also felt the piloted risk tool improved the level of engagement they had with the victim:

“...you can say to them from the outset “I’m going to ask you 12 (sic) questions” and it just makes them a little bit more inclined to participate, and be involved, and concentrate.” (PC)

“It was much easier I think to go through because it kept people’s interest for longer. I suppose it was over quicker which maintained people’s interest whereas the one before, by the time you got half way through the question set, people had lost interest or their mind had wandered ... Yeah their [the victim's] mind had wandered and you were just getting blank “Yes, no, yes, no, no, yes, yes” answers really, whereas this one now, you had to think about the answers because you had the options rather than just simple yes or no’s, and then to quantify what the answers were.” (PC)

The additional information given by the victim led a few officers who completed a paper-based form at the incident to suggest there was a need for freetext space beneath each question for officers to make notes at the time so specific details could be recalled more easily when writing it up electronically at a later point.

## The views of specialist support services

The views of specialist support workers interviewed in Sussex and Humberside<sup>19</sup> were mixed in relation to the piloted risk tool.

In Sussex, the ‘team leader’ and one of two support workers interviewed expressed a preference for the DASH. Interestingly, however, they were also aware of the limitations of the DASH, highlighting (as did all the partner agency interviewees) that DASH forms received from the police are of widely varying

---

<sup>19</sup> Interviews were not conducted with specialist support workers in West Midlands as West Midlands Police do not routinely send completed DASH risk assessments when making referrals to partner agencies.

quality. Support workers acknowledged that part of the reason many DASH forms lack sufficient freetext to place the yes/no responses in context is that some victims are not prepared to open up to first response officers in the way that they are to specialist support workers. Thus those who favoured the DASH had a slightly nuanced position in that they appeared to favour it when and if it was completed properly. Interviewees who favoured the piloted risk assessment felt that it was more suited to first response officers and that it was generally completed better than the DASH.

An interesting finding in relation to Sussex was that the support worker who dealt only with high risk police referrals favoured on balance the DASH, while the support worker who dealt with medium and standard risk referrals expressed a strong preference for the piloted risk assessment. It is possible that this finding reflects the impression of the researchers that first response officers tend to provide greater contextual detail on DASH risk assessments where they regard the risk of harm to be high than they do in relation to cases they regard as medium and particularly standard risk.

Interviewees in Sussex expressed concern that if the piloted risk assessment were to be adopted then the police and support services would be using sets of questions that were not consistent. They also suggested, however, that they routinely complete their own DASH with victims (a version developed specifically for non-police agencies that contains three fewer questions than the police DASH). The support service lead interviewed in Humberside did not view the police using a different question set as problematic – they much preferred the piloted risk assessment, and felt that it acted as a 'screener' that provided a good basis for undertaking a more comprehensive risk assessment using DASH.

Based on the findings of the pilot, and the views of those involved in its implementation, a number of revisions to the piloted tool should be considered. These are presented in Appendix 6.

## 5. Conclusions and implications

The evaluation of a revised domestic abuse risk tool for first response officers showed positive results. Specifically, officers made more appropriate risk grading decisions when using the piloted risk tool than when using the current DASH tool. In addition, it appeared that some victims were more prepared to disclose coercive and controlling behaviour to officers using the piloted risk tool.

### **What are the benefits of a revised risk tool for first response officers?**

The DASH risk model was endorsed by National Policing leads in 2009. That the model was adopted by the vast majority of forces has been regarded by many practitioners as an achievement in and of itself, as has the fact that the police and their partner organisations use a broadly identical risk model. Yet concerns around the practical functioning of the DASH model in policing have been evident anecdotally for some time, and these concerns have been borne out in the College's programme of research on risk assessment for domestic abuse (see Myhill, 2016; Robinson et al., 2016). In particular, the DASH, through a combination of victims choosing not to disclose and officers frequently not probing for or recording relevant information, does not in many cases present a clear indication of the risk of harm to the victim.

The piloted risk tool appeared to be helpful in this respect. More focused and direct questions, based on the frequency with which specific abusive behaviours are experienced, encouraged victims to disclose and gave officers (and any subsequent supervisor or risk assessor) a clearer picture in some cases of the nature and intensity of the abuse. The single freetext section also appeared to be successful to some degree in getting officers to offer a clear rationale for their assessment of the level of risk to the victim.

### **How important is knowledge and understanding of domestic abuse?**

It was clear from the evaluation, however, that a revised risk tool is not sufficient in and of itself to obviate the need for officers to have an understanding of coercive and controlling abuse and to appreciate the risk ongoing abusive behaviours pose over and above discrete acts of physical violence. The piloted risk tool appeared to be more successful than the DASH in highlighting ongoing patterns of abusive behaviour. For a risk model to be fully effective, however, officers must not only be able to identify the presence of coercive and controlling behaviour, but also recognise the risk it presents.

The revised risk tool requires, in common with the DASH, first response officers to use their structured professional judgement to make a primary assessment of the risk of serious harm posed to the victim by the perpetrator. Although the pilot showed an increase in officers documenting coercive control, that outcome did not always lead to them assigning an appropriate risk grading. This finding

was most pronounced for post-separation harassment and stalking behaviours; there was a clear tendency among some officers (and indeed supervisors) to associate separation/the parties no longer co-habiting with a reduction in risk.

## Alternative approaches to risk assessment

The College is aware of increasing interest, among both academics and police practitioners, in predictive approaches to risk assessment; approaches that could supplant the use of victim interviews and structured professional judgement. Such approaches are based on the identification of risk factors that are most predictive of future victimisation, through quantitative modelling of data from existing risk assessments and police case-files. Though the College is interested in the potential value of such approaches, there remain potential difficulties with predictive modelling of risk data, and the application of 'big data' approaches to risk assessment for domestic abuse may best be described as in its infancy.

In particular, the issue of outcome measures poses problems. Predicting future harm based on police data is complicated by issues of reporting and (under) recording. Due to the widespread under-recording of coercive control, outcome measures based on (re) arrests or convictions will currently bias data towards cases that involve physical violence and injury. Basing outcome measures on simple recontact is equally problematic, as recontact may be a positive outcome for some victims, suggesting they are more willing to engage with the police in order to reduce the risk of further harm. There are also ethical concerns that big data approaches may reproduce and amplify existing biases that may be introduced by officers' selective or incomplete recording of risk data<sup>20</sup>.

Such issues may be mediated somewhat by the addition of data from partner agencies, such as with an ongoing ESRC-funded project led by researchers at Manchester University<sup>21</sup>. Yet, at present, the College does not envisage predictive models replacing an initial victim interview and primary/secondary risk assessment based on structured professional judgement. Rather, the value of predictive models may lie in helping to inform structured professional judgements of risk at one or more stages of the process as predictive models can never say with complete certainty how an individual will act.

## What are the possible implications for secondary risk assessment and risk management?

If the revised risk tool were to be adopted for first response officers, there would need to be national-level consideration of the potential implications for police secondary risk assessment. It is not likely to be the case, however, that introducing a revised frontline risk tool would disrupt significantly a consistent and well-functioning police secondary risk assessment process in which the DASH risk assessment is the centrepiece. When observations were conducted for the DASH review (Robinson et al., 2016) in central units responsible for secondary risk assessment, the process fell some way short of the standard

---

<sup>20</sup> For a discussion of the general issues around 'fairness' in big data see <https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de>

<sup>21</sup> <https://gtr.ukri.org/projects?ref=ES/M01178X/1>



described in the academic literature on risk assessment (see for example Kropp, 2008). It was found, for example, that secondary assessors, some with little or no formal training on domestic abuse, spent a relatively short amount of time assessing each individual case. It was observed that the information contained on the DASH itself appeared to play a fairly limited role in some risk assessments, at least as compared to the officer's account of the circumstances of the incident, and other information held on police systems (see Robinson et al., 2016).

The revised risk tool may have advantages for secondary risk assessors, and not just in that ongoing patterns of abusive behaviour may be better identified and recorded. Part of the difficulty for secondary assessors appeared to be that relevant information was dispersed between different sources – primarily the DASH assessment and the officer's incident report. In particular, a secondary assessor may have to interrogate the incident report – which in some forces is spread over several different 'log' entries on electronic systems – in order to find the officer's rationale and other relevant context. Having both the risk assessment and the officer's rationale for their risk grading presented on the same document would be an advantage in this respect. This approach would also benefit specialist support services, who are most often sent the risk assessment for any case referred, but not the officer's incident report.

It is important to note that, in certain cases, it is possible that secondary assessors received a primary risk assessment based on the piloted risk tool that had less information and contextual detail than a well-completed DASH assessment. Yet the research evidence, including the review of DASH (Robinson et al., 2016) suggests such cases would be a clear minority compared to DASH forms that contain little more than yes/no responses.

## Increased demand for secondary risk assessment?

Perhaps the most significant implication for police secondary risk assessment would be if, as was evidenced most clearly in Sussex, the greater disclosure and recording of ongoing controlling and coercive abuse resulted in a greater proportion of cases graded as medium rather than standard risk. In forces where secondary risk assessment is only conducted for cases graded high or medium risk by the primary assessment, this outcome would have potential resource implications for central units conducting secondary risk assessments, and also potentially for partner agencies accepting referrals. This issue is, however, linked with the wider issue of the under-recording of coercive control as a crime, and must be balanced against the potential benefits of reducing frontline demand if cases are sent more efficiently down a referral path which may help to address ongoing abuse and prevent escalation through earlier intervention.

## Implications for referrals to MARAC

One area in which the introduction of the revised risk tool may affect force practices is the process of referring cases to MARAC. Currently, the SafeLives MARAC 'toolkit' for police, cited in the College's Authorised Professional Practice, suggests all cases assessed as high risk be referred to MARAC, but that referrals can also be made based on professional judgement (for cases at any risk level), escalation in the number of calls to the police, and where there is 'visible high

risk' based on 14 or more 'yes' responses to the DASH (SafeLives, 2015). The DASH review (Robinson et al., 2016) suggested some forces do automatically trigger MARAC referrals based on 14 'yes' responses to the DASH. There is a lack of empirical evidence for any numerical 'cut-off' score. Indeed, scoring the DASH is problematic because some questions – those relating to separation, pregnancy or recent birth, and abuse of children – do not apply to some victims, and the police DASH has three more questions than the SafeLives DASH used by partner agencies. As the revised risk tool cannot, or at least should not, be scored in this way, MARAC referrals would need to be based solely on professional judgement.

As the evaluation did not suggest the revised risk tool would result in a greater proportion of cases being graded as high risk, a significant increase in MARAC caseloads would not be likely were it to be adopted. Any increase in cases graded as medium risk, however, would increase to some extent the potential in some forces for more cases to be referred based on professional judgement, especially if these cases are referred to partner agencies and then to MARAC if assessed as high risk by a specialist support worker.

## **Future research**

In terms of what the police can achieve, improving the short, medium and long-term safety of victims depends on all aspects of police response; from the way the initial call is handled, to accurate recording and primary risk assessment, through secondary risk assessment and risk management in partnership with other agencies. The research conducted by the College suggests two key and significant gaps in the current evidence-base on the police response to domestic abuse.

### **The implementation of the coercive control law**

The pilot evaluation suggested the offence of controlling or coercive behaviour is being significantly and systematically under-recorded by forces. While there was an increase in cases recorded during the pilot, the numbers remained very low relative to the number of cases where either the risk assessment or incident report (or both) documented at least some degree of controlling behaviour. Numerous cases were identified during the case-file analysis where the principal offence appeared to be coercive control, but where the case had not been recorded as such. Most often, such cases were either recorded as non-crime incidents (even in some instances where coercive control was clearly described in the incident report), or were recorded as a lower offence (usually physical assault below the level of GBH with intent).

Further research would be required to understand the reasons for the under-recording of coercive control. The under-recording of coercive control has significant implications for the process of risk assessment and management. If, for example, a case is accurately recorded as coercive control, then it should make it much more difficult for officers dealing with subsequent calls to view it as a 'one-off' incident. The presence of coercive control should also help inform decisions around secondary risk assessment and referrals to partner agencies.

### **Risk management**

Though the research evidence-base on risk assessment has expanded significantly in recent years, there is still a dearth of evidence around risk management. The evidence-base for the effectiveness of MARAC is thin (Steel et al., 2011), and the College is not aware of published evaluations of other force risk management processes and models. In particular, there does not appear to be much longitudinal evaluation of outcomes for victims following their engagement with the police.

Force processes for secondary risk assessment and risk management likely vary significantly, and especially in relation to multi-agency and partnership working. Evaluation of different models of Multi-Agency Safeguarding Hubs, for example, would be welcome. In consideration of the findings from the DASH review (Robinson et al., 2016), a model of secondary risk assessment and risk management that may be worth testing and evaluating is one that minimises the police secondary risk assessment process in favour of referring cases to specialist support services in a timely way. Specialist support workers are best placed to undertake a thorough risk (and needs) assessment, both in terms of their level of training and the preparedness of victims to disclose. With this model, the police could concentrate their resources on criminal justice outcomes and preventative work targeted at perpetrators. This model would, however, work best in areas where all cases are referred automatically to specialist support services, regardless of both the level of risk allocated by the police and the wishes of the victim.<sup>22</sup> There may of course be resource considerations for partner agencies with this approach.

## Implications

The key implications of the evaluation of the risk pilots are:

- If adopted, the revised risk tool for frontline officers may improve both victims' disclosure and officers' primary risk assessments.
- If adopted, officers require an understanding of coercive control in order for the tool to be most effective.
- If adopted, the revised risk tool may result in a greater proportion of cases being graded as medium as opposed to standard risk, reflecting the ongoing nature of the abuse. Any impact on resources for secondary risk assessment would need to be balanced against potential benefits associated with earlier referral and intervention.
- If adopted, there should be national-level consideration of the potential impact of the revised risk tool on risk management processes, including referrals to partner agencies and MARAC.

---

<sup>22</sup> In some forces, referrals are made to partner agencies irrespective of the victim's wishes. In others, referrals are only made if the victim, when asked, consents to the referral. The latter practice is likely the result of concerns around data protection and use of personal data.

- There remain key gaps in the evidence-base around risk management and the implementation of the coercive control legislation.

# Appendix 1. Rationale for the development, structure and content of piloted risk tool

## Assumptions

The underlying assumptions that have informed the design of the piloted risk tool, and which reflect the research evidence-base, are:

- Coercive control, including post-separation stalking and harassment, is the most harmful form of domestic abuse and presents the greatest risk of serious injury and homicide (Campbell et al., 2003; Stark, 2007; Dobash and Dobash, 2015; Myhill, 2015; Sharp-Jeffs and Kelly, 2016; Monkton-Smith et al., 2017).
- Not all of the abuse that comes to the attention of the police under the wide-ranging national definition of domestic abuse is coercive control (Myhill, 2016; Butterworth and Westmarland, 2016).
- Frontline officers often take an incident-based approach to domestic abuse (Hoyle, 1998; Robinson et al., 2016).

## Key principles

The key principles for a revised risk tool were:

- A revised tool should improve consistency and completeness of risk data, by providing a straightforward and structured method of obtaining information from victims.
- A revised tool should improve the accuracy of subsequent risk assessment and referrals to specialist support services.
- A primary risk assessment should contain only questions that research evidence suggests are key to identifying risk of harm and making an initial assessment of safeguarding requirements; a more comprehensive risk/needs assessment is best undertaken by a specialist.

## Consultation

A programme of consultation was undertaken during development of the piloted risk tool. Initial meetings were held with SafeLives and Laura Richards, who were involved in the National Policing working group that developed DASH, and with Women's Aid, the national umbrella organisation for domestic violence voluntary/charity sector organisations. Through these channels, consultation was ongoing throughout the development period and as the risk assessment was refined.

Professor Amanda Robinson (Cardiff University), another member of the National Policing working group that developed DASH, was involved in the development of the piloted tool. Professors Liz Kelly (London Metropolitan University), Marianne Hester (Bristol University) and Betsy Stanko were also consulted on the content of the tool.

An early version of the revised risk assessment was presented and discussed at a workshop at the Women's Aid conference in July 2016. Also as part of the consultation and development period, College researchers conducted in-depth interviews with eight domestic abuse survivors (recruited through the Women's Aid panel), and fourteen frontline police officers.

## Key changes

A review of the DASH model in England and Wales (Robinson et al., 2016) suggested the current twenty-seven question tool does not produce consistent information on risk when used by frontline police officers. In part, this inconsistency is caused by officers either not completing the freetext sections of the form, or not providing adequate detail to set an incident in context.

The piloted risk tool tried to address the issue of inconsistent and incomplete data by:

- Reducing the number of questions such that only those key to **identifying** risk and making an initial assessment are asked at the frontline.
- Changing the response options primarily to scales in order to estimate the frequency of abusive behaviours and focus thinking on patterns of behaviour.
- Having a single freetext box where officers are asked to give their summary of risk (with prompts for key issues to consider).

## New question set

Table 1 shows where questions included in the piloted risk tool map across to those on the original DASH, and explains any changes in question wording.

**Table 1: Rationale for questions included in the piloted risk tool**

DASH question(s)	New question	Response options	Rationale
<p>Has (...) ever threatened to hurt or kill the children/dependents?</p> <p>Has (...) mistreated an animal or the family pet?</p>	<p>How often does (...) make threats to harm you or things you care about such as people, pets or property?</p>	<p>Never/occasionally/often/all the time</p>	<p>Frequent threats are generally employed in coercive control. Emphasises the personalised nature of the abuse. Draws in control through pets without having a separate question</p>
<p>N/A</p>	<p>How often does (...) call you names, humiliate or degrade you?</p>	<p>Never/occasionally/often/all the time</p>	<p>Captures the frequency of emotional abuse, which is a key tactic of coercive control not assessed directly by the DASH</p>
<p>Does (...) try to control everything you do and/or are they excessively jealous?</p>	<p>How often does (...) control your daily activities, such as who you can see or how you perform household tasks?</p>	<p>Never/occasionally/often/all the time</p>	<p>Original question conflates micro-regulation of everyday life (behaviour) and jealousy (possible cause of the behaviour). Most important in terms of identifying risk is to establish that the behaviour is present</p>
<p>Are there any financial issues? For example, are you dependent on (...) for money/have they recently lost their job/other financial issues?</p>	<p>How often does (...) deny you access to money, or control what you can spend it on?</p>	<p>Never/occasionally/often/all the time</p>	<p>Financial abuse has been identified as a key element of coercive control (Sharp-Jeffs, 2015). The current DASH question is often misinterpreted by officers and acts most often as a proxy measure for deprivation rather than establishing whether there is financial control within the relationship. The revised question addresses financial control more specifically</p>

Does (...) constantly text, call, contact, follow, stalk or harass you?	How often does (...) make you account for where you have been, or monitor your phone, email or social media to check up on you?	Never/occasionally/often/all the time	Captures the frequency of surveillance within the relationship. Updates it to include social media
Does (...) constantly text, call, contact, follow, stalk or harass you?	How often does (...) follow or stalk you, or try to contact you when you do not want them to?	Never/occasionally/often/all the time	Relates more specifically to post-separation stalking and harassment
Is there conflict over child contact? (please state what)	How often does (...) use child contact arrangements to control you or continue to abuse you?	Never/occasionally/often/all the time	There is potential for conflict over child contact between many ex-partners. Question makes specific the link between using child contact to continue the abuse and control
Do you feel isolated from family/friends i.e. does (name of abuser(s)) try to stop you from seeing friends/family/Dr/others?	How often do you feel isolated or like you have no one to turn to for support?	Never/occasionally/often/all the time	Original question conflates feeling isolated with controlling behaviour, which is covered by another question. Captures frequency of feeling isolated/not feeling able to access support
N/A	How often does (...) use physical violence towards you, such as pushing, slapping, punching or kicking?	Never/occasionally/often/all the time	Captures the frequency of physical violence, which is important for identifying coercive control and for reducing 'false positives' (e.g. 'one-off' or situational fighting between family members being graded medium risk just because there is an ABH level injury)
Has (...) ever attempted to strangle/choke/suffocate/drown you?	How often does (...) attempt to strangle, choke, suffocate or drown you?	Never/occasionally/often/all the time	Captures frequency of sub-lethal violence, which is important for immediate risk assessment. Even isolated/infrequent sub-lethal



			assaults are an indicator of heightened risk
Has (...) ever used weapons or objects to hurt you?	How often does (...) use or threaten to use weapons such as household items, knives or guns to hurt you?	Never/occasionally/often/all the time	Captures frequency of potentially sub-lethal/controlling violence, which is important for immediate risk assessment. Even isolated or infrequent use of weapons might be an indicator of heightened risk
Has (...) ever threatened or attempted suicide?	How often does (...) threaten or attempt suicide?	Never/occasionally/often/all the time	Captures frequency of controlling behaviour that has been linked to domestic homicide.
Have you separated or tried to separate from (name of abuser(s)) within the past year?	Are you planning to separate or have you tried to separate from (...) and has this put you in danger?	Yes/no	While separation in the context of coercive control is associated with domestic homicide, the question as worded in the DASH can result in inconsistent and possibly inaccurate data. For instance, the question is phrased in the past tense which can lead to confusion and inaccurate recording if a victim says that they intend to separate as a result of the current incident. Separation is also highly correlated with domestic abuse, but not all cases are in the context of coercive control – the revised question clarifies that the (planned) separation must have increased the threat
Is the abuse getting worse?	Is the abuse you are experiencing from (...) getting worse?	Yes/no	Escalation of abuse has been identified by research as associated with increased risk

<p>Has (...) ever threatened to kill you or someone else and you believed them?</p>	<p>Has (...) ever threatened to kill you and you believed they were capable of doing it?</p>	<p>Yes/no</p>	<p>Captures specific and credible threats to kill the victim. Threats to harm people the victim cares about is covered in an earlier question. The original DASH question potentially draws in perpetrators' threats to acquaintances and strangers</p>
<p>N/A</p>	<p>On a scale of 0 to 10, where '0' is not at all likely and '10' is very likely, how likely do you think it is that (...) will seriously injure you in the future?</p>	<p>0-10</p>	<p>Research has shown that victims' own perceptions of risk are as accurate as some actuarial risk assessment tools. Victims are well placed to know the imminent and future threat posed by the perpetrator. Officers should be given guidance that low scores on this scale may not necessarily reflect the actual threat and should be considered in combination with other information and responses</p>

## Rationale for excluding questions

Table 2 shows current DASH questions that were not included in the piloted risk tool and the reason(s) for their exclusion.

**Table 2: Original DASH questions that are not included in the piloted tool**

DASH question	Rationale for not including
<p>Has the current incident resulted in injury? (Please state what and whether this is the first injury)</p>	<p>Officers should ascertain this information anyway as part of their investigation into the incident. A survivor interviewed as part of the development of the revised tool suggested that as officers ask what has happened at the current incident prior to the risk assessment, being asked this question can make a victim feel as though they were not believed. The question also asks only about injurious violence, when research suggests much of the physical violence in coercive control is non-injurious but repeated (Stark, 2013). The more pertinent issue is to establish how often physical violence is used as a tactic of control.</p>
<p>Are you very frightened?</p> <p>What are you afraid of? Is it further injury or violence?</p>	<p>This question is very vague if there is no contextual information recorded. Survivors said there may be a difference between how somebody feels at the time of a specific incident, and how they feel more generally. Evidence from the fear of crime literature suggests it is more helpful to ask about likelihood of further victimisation than generic fear, as fear can be context specific. The follow-up question directs officers down the route of focusing specifically on physical violence.</p>
<p>Are you feeling depressed or having suicidal thoughts?</p>	<p>Depression is now extremely common, with an estimated 1 in 4 people in England experiencing depression in any given year<sup>23</sup>. The potential for this question to help discriminate between the possible levels of risk a victim faces is therefore not clear. A secondary assessment would be a better time to understand the victim's needs in relation to their mental health.</p>
<p>Are you currently pregnant or have you recently had a baby in the past 18 months?</p>	<p>Officers should, as part of their investigation of the incident, routinely collect details of any children present at the time of the incident or who normally reside at the address where the incident took place. Pregnancy/recent birth has been associated in some studies with the onset of abuse. The revised tool asks about abuse the victim has experienced/is experiencing currently. Pregnancy/recent birth may better inform a secondary risk/needs assessment.</p>

<sup>23</sup> <https://www.mentalhealth.org.uk/statistics/mental-health-statistics-uk-and-worldwide>

	Officers are encouraged to mention any relevant issues relating to pregnancy in the freetext summary at the end of the piloted assessment.
Are there any children, step-children that aren't (.....) in the household? Or are there other dependants in the household (i.e. older relative)?	As with pregnancy/recent birth, we do not feel it is necessary to ask this question at the initial identification/assessment stage, and feel that it would be better considered in a more thorough secondary risk/needs assessment. The underlying pattern of control identifies risk is present; a child that is not the abuser's has been correlated with increased risk of serious harm, but it is an aggravating factor.
Has (...) ever hurt the children/dependents?	Victims will often withhold information when asked this question by police officers, owing to fears of the potential consequences of disclosing. An analysis of DASH data from a random sample of incidents (Myhill and Hohl, 2016) showed a disclosure rate of 5% for this question. The actual rate would be expected to be higher considering the known overlap between abuse of women and children. More accurate responses are likely during secondary risk assessment. Officers are expected to record details of children, and may ask about their welfare outside of the risk assessment.
Has (...) ever threatened to hurt or kill the children/dependents?	As above, this question is also underreported. The Myhill and Hohl (2016) study found a disclosure rate of 2.5%. Victims have the opportunity to consider threats to children when asked about threats more generally in the first question of the piloted risk assessment.
Is the abuse happening more often?	Coercive control is by its nature 'continuous', and specific tactics of control will be frequent. A better indication of escalation is if the abuse is perceived to be getting worse (i.e. the perpetrator is having to increase the type or severity of abuse to try to maintain control).
Does (...) do or say things of a sexual nature that make you feel bad or that physically hurt you or someone else?	While research suggests sexual coercion is present frequently in coercive control, frontline police officers are not best placed to get full disclosure. This question is likely not to produce complete or consistent data. The Myhill and Hohl study found a disclosure rate of 10%. A higher level of disclosure is likely to be achieved by asking this question during a secondary risk assessment, resulting in more accurate information on which to base the full risk/needs assessment. Additionally, sexual coercion is unlikely to occur without the other controlling behaviour/abuse that is covered by the piloted risk assessment. Women's Aid, SafeLives and all prominent academics consulted agreed that this question should not be asked at the frontline.

<p>Is there any other person that has threatened you or that you are afraid of? (If yes, consider extended family if honour based violence. Please specify who)</p>	<p>This question was originally intended to act as a way to identify so-called Honour Based Abuse (HBA). We are not sure how effective it is, or how many forces then go on to ask additional DASH HBA questions. Academic experts have suggested 'multiple perpetrators' is not a useful gateway question, as many perpetrators recruit others, irrespective of HBA. A bespoke question for HBA would likely be difficult to draft, as (similar to 'rape') some victims may not identify with a potentially stigmatised term such as 'honour'. In addition, a bespoke question would not be applicable to the vast majority of victims, and when such questions are included in risk assessments officers can become cynical about it being a 'tick box' exercise. Probing for and identifying HBA may be more of a training and understanding issue.</p>
<p>Do you know if (...) has hurt anyone else?</p>	<p>A victim will not always be aware of a perpetrator's history and so might unintentionally provide inaccurate information as a result. If the perpetrator has come to the attention of the police for acts of violence against others (including previous partners) this information can be identified from police systems.</p>
<p>Has (...) mistreated an animal or the family pet?</p>	<p>Many officers believe this question to imply that men who are violent towards animals are more likely to be violent to people. The original DASH guidance in fact emphasises how (threats towards) pets are used to control victims and prevent them from leaving. Victims have the opportunity to consider pets when asked about threats more generally in the first question of the piloted risk assessment.</p>
<p>Has (...) had problems in the past year with drugs (prescription or other), alcohol or mental health leading to problems in leading a normal life?</p>	<p>Drugs, alcohol and mental health (the 'toxic trio') are highly correlated with most forms of criminality, so it is not clear that this question provides additional useful information in terms of <b>identifying</b> risk. Aggravating factors relating to substance use and mental health issues may be better considered at a second stage risk/needs assessment. Opinion was divided among stakeholders as to whether questions on the 'toxic trio' should be included. In practice, it was found to be very difficult to phrase questions to be specific enough to provide useful information in the absence of freetext clarification. Officers are encouraged to mention any relevant issues relating to substance abuse or mental health issues in the freetext summary at the end of the piloted risk assessment.</p>
<p>Has (...) ever breached bail/an injunction and/or any agreement for when they can see you and/or the children?</p>	<p>This information should be available through police records without officers having to ask it of the victim. If a victim has not reported a breach</p>

	<p>of bail to the police, we feel they are unlikely to disclose it as part of a risk assessment conducted by a frontline officer.</p>
<p>Do you know if (...) has ever been in trouble with the police or has a criminal history?</p>	<p>A victim may not be aware of a perpetrator's history and so might unintentionally provide inaccurate information. This information can be applied more accurately at a second stage risk assessment, using data held police systems. Frontline officers are expected to complete background checks during initial investigations.</p>

## Appendix 2. Piloted risk tool

	Never	Occasionally	Often	All the time	Not stated
1. How often does (...) make threats to harm you or things you care about such as people, pets or property?					
2. How often does (...) call you names, humiliate or degrade you?					
3. How often does (...) control your daily activities, such as who you can see or how you perform household tasks?					
4. How often does (...) deny you access to money, or control what you can spend it on?					
5. How often does (...) make you account for where you have been, or monitor your phone, email or social media to check up on you?					
6. How often does (...) follow or stalk you, or try to contact you when you do not want them to?					
7. How often does (...) use child contact arrangements to control you or continue to abuse you?					
8. How often do you feel isolated or like you have no one to turn to for support?					
9. How often does (...) use physical violence towards you, such as pushing, slapping, punching or kicking?					
10. How often does (...) attempt to strangle, choke, suffocate or drown you?					
11. How often does (...) use or threaten to use weapons such as household items, knives or guns to hurt you?					
12. How often does (...) threaten or attempt suicide?					

	Yes	No	Not stated
13. Are you planning to separate or have you tried to separate from (...) and has this put you in danger?			
14. Is the abuse you are experiencing from (...) getting worse?			
15. Has (...) ever threatened to kill you and you believed they were capable of doing it?			

	0	1	2	3	4	5	6	7	8	9	10	Not stated
16. On a scale of 0 to 10, where '0' is not at all likely and '10' is very likely, how likely do you think it is that (...) will seriously injure you in the future?												

## General assessment of the threat posed to the victim

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

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

*DRAFT*

## Initial assessment of threat

Standard  Medium  High

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

Risk of serious harm (Home Office 2002 and OASys 2006):

'A risk which is life threatening and/or traumatic, and from which recovery, whether physical or psychological, can be expected to be difficult or impossible.'



## Appendix 3. Research methods

This appendix provides further detail on the rationale for choosing a quasi-experimental design, the data collected for the case-file analysis, and the inter-rater agreement exercise for the case-file analysis.

### Why not a randomised controlled trial (RCT)?

An RCT design was considered, but ruled out for the following reasons:

- It would not have been possible in practical terms to randomise at the officer level, or at the point of dispatch. For domestic abuse incidents, many forces have a policy of either 'double crewing' (dispatching a car with two officers), or dispatching two single crewed units where possible so that the victim and perpetrator can be separated if necessary and spoken with individually. If two officers attended an incident, one assigned to the treatment group and the other the control group, there would be potential for contamination which could impact the fidelity of a trial. Similarly, randomisation of incidents at point of dispatch would be very difficult in operational terms and would have risked officers trained to use the pilot risk assessment tool approaching incidents assigned to the control group in a different way to how they may have done previously, or they may have used the wrong risk assessment tool in error.
- A cluster RCT would not have been feasible. While cluster sampling could perhaps have mitigated the issues outlined above, data from four areas is the absolute minimum that would be needed for this type of analysis, and the sample size would have needed to be very large. Due to the intensive nature of the data collection for the evaluation, this design was ruled out.
- At a higher level, exact matching of treatment and control sites would have been extremely difficult, if not impossible. The implementation issues outlined above dictate that the unit of randomization would likely be an entire division/command unit of a force, to minimise potential crossover and contamination. A non-equivalent group design mitigates the need for exact matching of areas.
- There were issues with the way force systems could be adapted to facilitate the piloted risk tool which necessitated taking a pragmatic approach to the research design.

### Case file data extraction

Data from randomly selected weeks was transcribed from police systems manually to an Excel spreadsheet. Data were collected for the following fields:

- Person reporting – victim, perpetrator or third party (family or other).
- Type of incident – intimate partners or family.
- Command and control log – summary of the original call to police.
- Officer's written account – summarised from incident reports.
- Whether an arrest was made/for what offence.

- Victim's responses to the risk assessment questions.
- Frontline officer's risk grading (standard, medium or high).
- Supervisor's risk grading (standard, medium or high).
- Previous incidents on the system involving the same people (single or multiple).
- Any other relevant information (drawn from all available case-file documents, including victim and witness statements (MG11s)).

## Inter-rater agreement

A kappa test indicated a 'substantial' level of agreement between the risk grading judgements of the researchers and a specialist support worker (see chapter 3). There were disagreements in a proportion of cases, however. That the level of agreement was not higher might be explained in part by the inherently subjective nature of risk assessment utilising structured professional judgement. The kappa test was run again in order to check that the specialist support worker's risk decisions were not aligned more closely with the frontline officers than with the researchers. This was not the case. The kappa value when comparing first response officers' risk judgements with those of the specialist support worker was 0.3, indicating only a 'fair' level of agreement. In other words, where there was disagreement between the researchers and the specialist support worker, the support worker's risk grading was not likely to have matched that of the frontline officer. In fact, closer inspection of the data revealed that a clear majority of mis-matches between researcher and specialist support worker concerned cases where the researchers (and in some cases the officers) had chosen 'medium' and the specialist 'high'. This tendency for specialists to grade risk higher certainly than police officers (see Robinson, 2010) can be witnessed in cases that are referred in to MARAC by a specialist support agency following risk assessment by the police that has not assessed the case as high risk.

A further interesting finding from the inter-rater agreement exercise was that the researcher and specialist assessments of risk were more closely aligned when officers had used the pilot risk assessment than when they had used the DASH. This finding was particularly marked in West Midlands, where officers' frequently recorded little contextual data when using the DASH. Though not possible to know for sure, it may be that the contextual information about frequency of behaviours provided by the piloted risk assessment permitted greater consistency of assessment.

## Appendix 4. Humberside data

In Humberside Police, the risk tool was piloted in Bridlington for a period of three months. Case-file data was manually transcribed by staff from the force for cases that were reported during the pilot period for both Bridlington and Cleethorpes (the comparison area) as well as data for cases reported during the three months immediately prior to the pilot period for both areas (the 'pre-pilot' data). Unfortunately, on review of the data, it became apparent that it contained numerous errors which rendered the comparison area data and pre-pilot data from Bridlington unusable. This section presents a high-level summary of the Bridlington pilot data, but no comparisons could be made between this data and any other time point or area.

### First response officers' primary assessments of risk

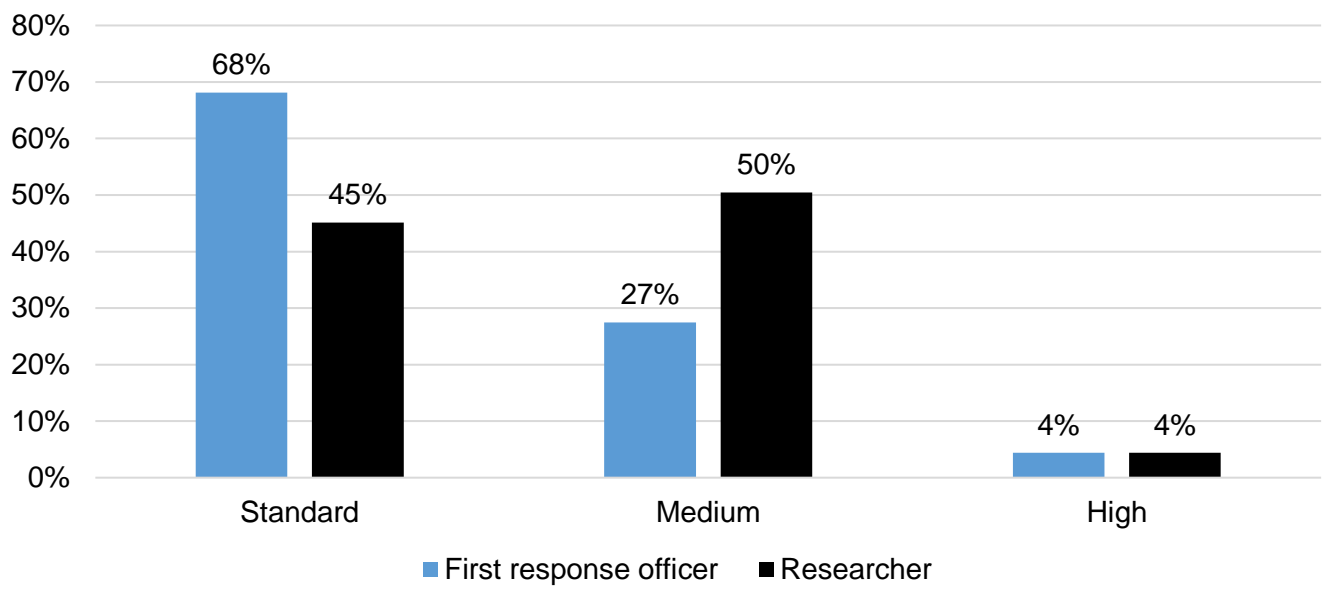
Cases were analysed from randomly selected weeks during the pilot period to determine whether the revised tool resulted in officers reaching more appropriate assessments of the level of threat posed by the perpetrator. As with the other two forces, all cases were double coded by two researchers using a pre-determined set of coding rules (see Chapter 3). As can be seen in Table 1 researchers agreed with the risk grading level selected by first response officers in 59% of cases.

**Table 3 Agreement in risk grading decisions between first response officers and researchers**

	Number of cases	Percentage of cases
Agree	67	59.3
Disagree	46	40.7
Total	113	100

There was a statistically significant difference in the proportion of cases graded as standard and medium risk, with first response officers grading a larger proportion of cases as standard risk than researchers and vice versa (see Figure 1). Of the 46 cases where there was disagreement over the risk grading, researchers selected a lower risk grading level than officers in 10 of those cases, upgraded the risk from medium to high in three cases and the rest were cases where the officer had graded it as standard risk and the researcher medium risk. A large part of this discrepancy in risk levels was due to the researchers identifying an element of post-separation stalking or harassment in several of these cases and so deeming the risk to be greater than standard. As mentioned, however, owing to the lack of data for the pilot area before the pilot period or comparable data for another part of the force, it is not possible to draw any conclusions as to whether there was any difference, positive or otherwise, in the risk grading decisions of first response officers using the revised risk tool.

**Figure 1. Risk grading levels for first response officer and researcher**



# Appendix 5. Logistic regression difference-in-differences model

A binary logistic model was used to estimate a difference-in-differences (DiD) effect.

The logistic model is slightly more complicated than the linear regression model. Its primary use is in predicting the probability of a dichotomous event (in our case the probability of two assessors agreeing), and is given by:

$$P(Y) = \frac{1}{1 + e^{-(Z)}}$$

$P(Y)$  = the probability of event  $Y$ .

$e$  is a mathematical constant.

$Z$  is described below.

It can be shown that the DiD can be estimated from the interaction term in the model below, as this specifically estimates the 'additional' probability of agreement when a case is in the pilot group post implementation.

$$Z_{it} = \beta_0 + (\beta_1 Time_t) + (\beta_2 Group_i) + (\beta_{DiD} * Time_t * Group_i) + \varepsilon_{it}$$

Where:

$Time_t$  is a dummy variable for the period pre- (coded 0), and post (1) the introduction of the piloted risk tool.

$Group_i$  is a dummy variable indicating pilot (coded 1) or comparison (coded 0) group membership.

$\beta_0, \beta_1, \beta_2,$  and  $\beta_{DiD}$  are coefficients for the constant, time, group and difference-in-differences terms respectively.

$\varepsilon_{it}$  is the 'error' term for group  $i$  at time  $t$ . The brackets have been added for clarity.

## Results

Data from the two sites was combined to produce a larger sample and to give a more general assessment of the DiD estimate. This was justified on the grounds that (a) the data from the two sites was very similar, (b) they were both from the same study and (3) a logistic model which included an 'area' effect showed that the area variable was not statistically significant. Results from the analysis are presented in Table 1.

**Table 1. Logistic regression analysis**

	B	SE	Exp(B)	95% CI for Exp(B)	
				Lower	Upper
Time	-.116	.241	.891	.556	1.428
Group	-.130	.250	.878	.538	1.435
DiD	.826	.344	2.284*	1.164	4.482
Constant	.353	.181	1.423		

\*p=<0.05

The value for the Constant gives a 'base' agreement rate of around 58%. The effect of moving from pre-pilot (time 1) to post-pilot (time 2) condition, holding constant any impact of the pilot itself, is small and not statistically significant. The variable representing membership of the pilot or comparison group ('Group') is also small and negative reflecting the observation that, all other things being equal, the level of agreement in the pilot group is slightly below that in the control (this can be seen in Figure 1). This variable is also not statistically significant, however, indicating that there is no evidence that the two groups can be regarded as different<sup>24</sup>.

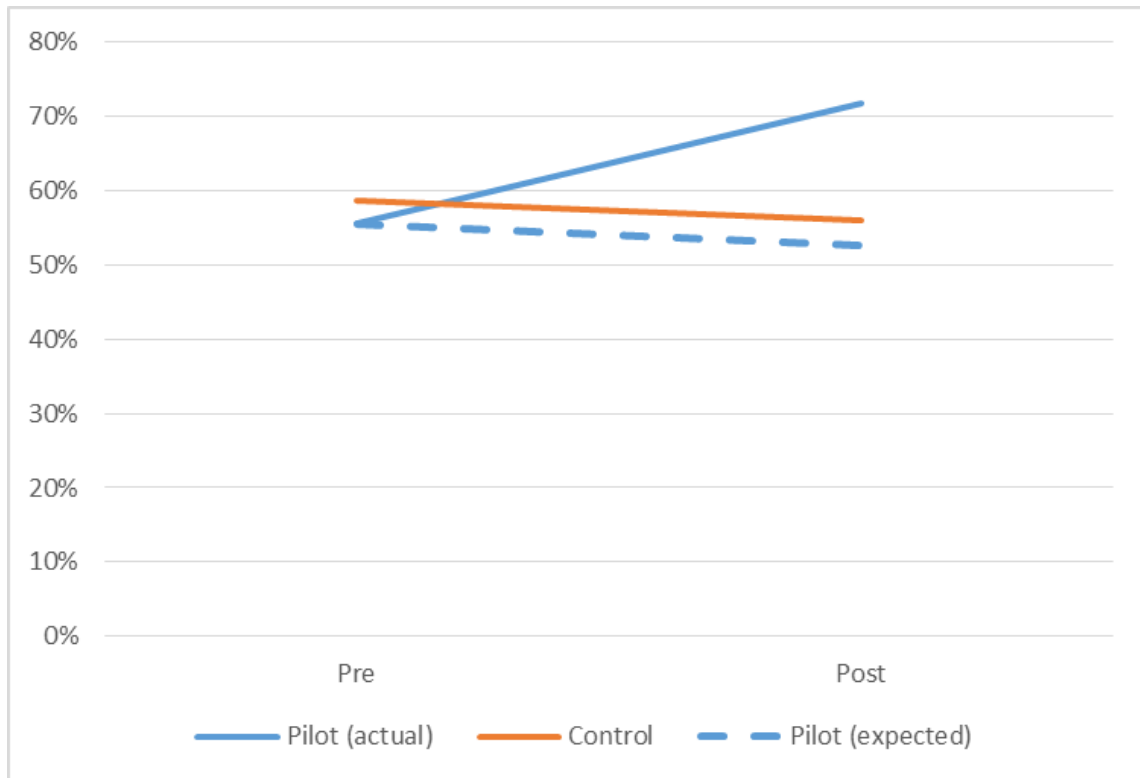
The largest and most influential variable in the model is the DiD which is statistically significant and shows that the odds of agreement were 2.3 times greater in the pilot group.

A visual representation of this model, showing the slight gap between the groups and including the observed and expected change from the pre-pilot to post-pilot periods, is shown in Figure 1.

---

<sup>24</sup> Technically, there is not enough evidence to reject the hypothesis that the two groups are different.

**Figure 1. Percentage of cases where officers and researchers made the same risk grading decision (combined data), including an estimate of the expected result in the pilot group**



The difference between the observed agreement rate in the pilot group, and the expected result from the model, which assumes the intervention has not been applied, is an increase of 19 percentage points (72% (observed) minus 53% (modelled)).

# Appendix 6. Suggested changes to the piloted risk tool

On the basis of the qualitative data collected for the evaluation of the pilots – primarily interviews with first response officers and direct observations of them using the piloted risk tool – there are revisions that should be considered if the revised risk tool were to be further tested or implemented more widely.

## 'Separation' question

Several officers commented that separating question 13 on recent separation of the victim and perpetrator into two parts would make it easier for victims to answer. A victim would then have say in the first instance only whether they have recently separated or plan to separate from the abuser; only if the answer was 'yes' would they then be asked if it has/will put them in danger.

## Abuse of children

A direct question relating to whether the perpetrator has hurt or threatened children should likely be included if the forthcoming domestic abuse bill contains (as proposed) provision for an 'aggravated' offence if children are involved.

## Control through child contact post-separation

Having frequency response options for this question did not always appear to make sense and so it is suggested that this becomes a yes/no question.

## Physical violence, strangulation, and use of weapons

A frequency scale appeared more appropriate for non-physical coercion than for physical assaults, as non-physical coercion is more likely to be perceived as continuous. While physical violence in the context of coercive control can be less severe but frequent (see Stark, 2007), it was also evident that discrete acts of physical violence, particularly severe acts of violence, are less likely to be perceived as continuous in nature, and one occurrence could have a very significant impact. In addition, there was evidence from the pilot (in line with Kelly, 1999) that some victims are prompted to report ongoing abuse the first time they are assaulted physically. It is suggested, therefore, that the response options of 'never', 'once only', 'occasionally' and 'often' are used for questions relating to physical assaults and use of weapons.

## Perpetrator's threats or attempts to commit suicide

Threats or attempts at suicide will not always be continuous in nature and so again it is suggested the response options are changed to 'never', 'once only', 'occasionally' and 'often'.

## So-called 'honour-based abuse' (HBA)

While there remain valid reasons for not asking a specific question about HBA, officers could be prompted to consider it prior to the freetext section.



# References

ACPO (2009) Domestic abuse, stalking and harassment and honour-based violence (DASH 2009) risk model: Practice guidance for specialist police staff. Report, Association of Chief Police Officers, UK.

Bagwell-Gray M, Messing J and Baldwin-White A (2015) Intimate partner sexual violence: A review of terms, definitions, and prevalence. *Trauma, Violence & Abuse*, 16(3): 316-335.

Butterworth K and Westmarland N (2016) Victims' views on policing partner violence. *European Police Science and Research Bulletin*, 13: 60-63.

Campbell J, Webster D, Koziol-McLain J, Block C, Campbell D, Curry M ... Laughon K (2003) Risk factors for femicide in abusive relationships: Results from a multisite case control study. *American Journal of Public Health*, 93(7): 1089-1097.

Dobash R and Dobash R (2015) *When men murder women*. Oxford: Oxford University Press.

HMIC (2014a) *Everyone's business: Improving the police response to domestic abuse*. Report, Her Majesty's Inspectorate of Constabulary, UK, March.

HMIC (2014b) *Crime-recording: Making the victim count*. Report, Her Majesty's Inspectorate of Constabulary, UK, November.

HMIC (2015) *Increasingly everyone's business: A progress report on the police response to domestic abuse*. Report, Her Majesty's Inspectorate of Constabulary, UK, December.

HMIC/HMCPSI (2017) *Living in fear – The police and CPS response to harassment and stalking*. Report, Her Majesty's Inspectorate of Constabulary and Her Majesty's Crown Prosecution Service Inspectorate, UK, July.

Hoyle C (1998) *Negotiating Domestic Violence: Police, Criminal Justice and Victims*. Oxford: Oxford University Press.

Kelly L (1999) *Domestic Violence Matters*. London: Home Office.

Kelly L and Westmarland N (2015) *Domestic Violence Perpetrator Programmes: Steps Towards Change*. Project Mirabal Final Report. London and Durham: London Metropolitan University and Durham University.

Kropp R (2008) Intimate-partner violence risk assessment and management. *Violence and Victims*, 23 (2), 202-220.

Landis J and Koch G (1977) The measurement of observer agreement for categorical data. *Biometrics*, 33: 159-174.

Monckton-Smith J, Szymanska K and Haile S (2017) Exploring the relationship between stalking and homicide. Report, Suzy Lamplugh Trust, UK. Available at: <https://www.suzylamplugh.org/Handlers/Download.ashx?IDMF=1a6cf4d9-0df5-42be-8b02-4bdb75fa264>.

Myhill A (2015) Measuring coercive control: What can we learn from national population surveys? *Violence Against Women*, 21(3): 355-375.

Myhill A (2016) Risk identification in cases of domestic abuse: Evaluation of a pilot project. London: College of Policing.

Myhill A (2017) Renegotiating domestic violence: Police attitudes and decisions concerning arrest. *Policing and Society*, doi: 10.1080/10439463.2017.1356299.

Myhill A and Hohl K (2016) The 'golden thread': Coercive control and risk assessment for domestic violence. *Journal of Interpersonal Violence*, doi: 10.1177/0886260516675464.

Myhill A and Johnson K (2016) Police use of discretion in response to domestic violence. *Criminology and Criminal Justice*, 16(1): 3-20.

ONS (2017) Crime in England and Wales: year ending March 2017. Report, Office for National Statistics, UK. Available at: <https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/crimeinenglandandwales/yearendingmar2017>.

Ritchie J and Lewis J (2012) *Qualitative research practice*. London: Sage.

Robinson A (2010) Risk and intimate partner violence. In Kemshall H and Wilkinson B (eds.) *Good practice in risk assessment and risk management* (3rd Edition). London: Jessica Kingsley Publishers, pp.119-138.

Robinson A, Myhill A, Wire J, Roberts J and Tilley N (2016) Risk-led policing of domestic abuse and the DASH risk model. What Works Centre: Crime Reduction Research.

Robinson A, Myhill A and Wire J (2017) Practitioner (mis)understandings of coercive control in England and Wales. *Criminology and Criminal Justice*, doi: 10.1177/1748895817728381.

SafeLives (2015) Toolkit for MARAC: Domestic abuse police officer. Report, SafeLives, UK. Available at: <http://www.safelives.org.uk/practice-support/resources-marac-meetings/resources-people-referring>

Sharp-Jeffs N (2015) A review of research and policy on financial abuse within intimate partner relationships. London: London Metropolitan University.

Sharp-Jeffs N and Kelly L (2016) Domestic Homicide Review (DHR): Case Analysis. London: Standing Together/London Metropolitan University.

Stark E (2007) *Coercive control: How men entrap women in personal life*. Oxford: Oxford University Press.

Stark E (2013) The dangers of dangerousness assessment. *Family and Intimate Partner Violence Quarterly*, 6: 13-22.

Steel N, Blakeborough L and Nicholas S (2011) Supporting high-risk victims of domestic violence: a review of Multi-Agency Risk Assessment Conferences (MARACs). London: Home Office.

Westmarland N, Kelly L and Chalder-Mills J (2010) Domestic violence perpetrator programmes: What counts as success? London: Respect.

Wheller L and Wire J (2016) Domestic abuse risk factors and risk assessment: Summary of findings from a Rapid Evidence Assessment. Ryton-on Dunsmore: College of Policing.