Original Article

Managing Policing Demand for Digital Forensics through Risk Assessment and Prioritization in England and Wales

Dana Wilson-Kovacs and Jen Wilcox

Abstract Digital forensic helps to unlock the evidential opportunities presented by digital devices. Ninety per cent of all cases in England and Wales carry a digital element and identifying the devices with the most evidential value is central to dealing effectively and efficiently with the demand for digital forensic (DF) analysis. This article offers an ethnographically informed, empirical analysis of how police requests for in-house DF examinations are organized in 35 forces in England and Wales, and the mechanisms, contexts, and challenges of managing-related internal demand, which have been rarely examined systematically in both policing and DF practitioner literature. The analysis illustrates how an updated understanding of risk and prioritization is needed to address the growing demand for DF analysis. The findings call for updated, systematic approaches that encourage the forecasting of internal policing demand collaboratively and bridge between the operational and the technical sides of investigations.

Introduction

The fastest growing of all forensics disciplines, digital forensics (DF) helps to unlock the evidential opportunities presented by digital devices. For law enforcement agencies, the demand for DF analysis outstrips capabilities and brings new challenges (Tully, 2020). Backlogs are common (Casey *et al.*, 2009; Houck, 2020) and practitioners spend more time 'in the acquisition and preparation steps of an investigation as opposed to detailed evidence analysis and reporting' (de Braekt *et al.*, 2016, p. 68). Keeping up with the pace of technological change, including the growing data storage capacities and diversity of systems, requires costly infrastructural investments and ongoing workforce development that stretch available police resources.

With 90% of all cases in England and Wales carrying a digital element (National Police Chiefs Council [NPCC], 2020), dealing in a timely manner with the demand for DF analysis has become key to criminal justice outcomes. Identifying the devices with the most evidential value to establish the urgency and order in which cases should be examined is central to this process (Casey *et al.*, 2009; Horsman, 2017; Houck, 2020). To date, however, policing literature has rarely scrutinized demand management strategies in forensic support services (Rappert, Wheat and Wilson-Kovacs, 2021). This article offers the first empirical study of how police requests for in-house DF examinations are organized in 35 forces in England and Wales, and the mechanisms, contexts, and challenges of

Department of Social and Political Sciences, Philosophy and Anthropology, University of Exeter, Amory Building, Rennes Drive, Exeter, EX4 4RJ, Devon, UK. E-mail: M.d.wilson-kovacs@exeter.ac.uk

Policing, Volume XX, Number XX, pp. 1–12
doi: https://doi.org/10.1093/police/paac106

managing demand in dead-box[1] forensics. The analysis aims to answer two interrelated questions: first, how are submissions for DF analysis managed by forces? Second, what are the challenges and implications of current settings for the delivery of justice? Examining the under-researched, yet vital role of DF in criminal investigations, we argue that similar to mainstream policing (Laufs *et al.*, 2021; Walley and Adams, 2019), the complexity of managing demand in DF requires a flexible, dialogical, and collaborative approach between forensic services support and investigative teams. The article's main contribution to policing scholarship is to highlight the embeddedness of DF in investigations and the urgent need to analyse the changing nature of internal organizational demand for DF analysis more comprehensively.

The paper is organized as follows: Part 1 briefly introduces studies on managing demand in policing, before focusing on DF practitioners' approaches to risk assessment and prioritization. Part 2 provides the historical context of these two processes in DF in England and Wales, explaining the development and adoption of a tool by the National Policing Improvement Agency (NPIA) in 2012 to manage demand for DF in-house services. Part 3 sketches the background and context of the research on which the analysis is based. Part 4 examines how risk assessment and prioritization are undertaken by police forces and explores how tool limitations and operational oversight result in bottlenecks and tensions. Summarizing the findings, Part 5 reflects on their implications.

Demand, risk, and prioritization in policing and DF practitioner literature

Understanding demand in policing is much needed both practically and conceptually (Laufs *et al.*, 2021). Police forces usually associate demand with the volume of calls for assistance received from the public or the number of scene-of-crime investigations completed in a day (Walley and Adams, 2019), but the many ways in which demand is talked about highlight its complexity, diversity, and changing nature beyond these two measurements (Laufs *et al.*, 2021; Morris and Walley, 2022). Considering the increased complexity of investigative demand created by the use of mobile phones, Walley and Adams (2019) note how in contrast with the dynamic nature of demand, inflexible planning processes lead to errors in decision-making, queues, and backlogs. Consequently, a fragmented understanding of internal-facing organizational demand means that forces rely on coping strategies rather than proactive and systematic long-term approaches. This article focuses on how 35 forces in England and Wales manage policing demand for DF analysis, more specifically, on how in-house digital forensic units (DFUs) use risk assessments to prioritize cases.

Similar to policing, the difficulties of measuring risk, comparing assessment tools and evaluating their effectiveness (e.g. Gundhus, 2021), are reflected in the DFUs' risk assessment processes and impact on their demand management strategies. While forces use prioritization to distribute limited resources and rely on mathematical models and algorithms to respond to incidents (Brayne, 2020), few social science analyses have focused on these arrangements in forensic science support (Wilson-Kovacs, 2019).

DF presents several risks, including those related to the welfare of staff due to the violent and sexual nature of the material examined. When cases go to court with unprocessed evidential information or procedures not correctly adhered to, there is the additional risk of miscarriages of justice (Horsman, 2017; Houck, 2020; Tully, 2020). Insufficient training and a lack of appropriate investigative tools also pose a risk to victims and organizational reputation. Below, we focus on how perceptions of risk and understandings of prioritization serve as the platform to manage policing demand for DF analysis.

[1] Dead-box forensics refers to the forensic examination of digital devices that have been disconnected from their power source and physical location during an investigation. It differs from live forensics where examination relies on the target operating system running to perform the analysis.

These processes have been only partially covered in DF practitioner literature, which centres largely on their technical rather than organizational dimensions. For instance, prioritization is typically linked to triage, an administrative and technical process based on the identification and seizure of the digital devices with the most probative value to a case (Horsman, 2017). Triage assists with the selection of seized devices for in-depth examination or with their elimination from an investigation. Used at both crime scenes and in the DFUs, triage can increase efficiency and benefit large-scale, time-sensitive investigations with multiple devices (James, 2014). It can also mitigate failure demand, that is, the type of internal demand produced by the failure to deal with an inquiry appropriately in the first place (Morris and Walley, 2022).

Similarly, risk analyses in DF focus on securing the integrity of the information extracted, with various workflow models seeking to ensure the appropriate acquisition, examination, analysis, and interpretation of data (Casey *et al*., 2009). Typically discussed here are the benefits of accreditation, standardization, and automation (de Braekt *et al*., 2016; Tully, 2020), and the need to assess procedures, decisions, and actions through quality assurance processes that minimize bias and create consistency (Horsman, 2017; Houck, 2020).

Very few DF practitioner studies focus on how demand for DF analysis is internally managed. Their findings cover different national jurisdictions and reveal a lack of standardization, inconsistencies, tensions in case allocation and wasted resources (Alawadhi, 2019; James, 2014). Comparing arrangements in USA, Canada, Ireland, Australia, and South Korea, James (2014), for instance, notes how law enforcement agencies in the same country can rank cases for DF analysis differently. In South Korea, the absence of long-term planning and rapidly changing priorities led to ranking cases according to their public scrutiny and media coverage. In comparison, Canadian and Irish police forces used locally developed prioritization models based on variables such as offence types and number of devices submitted. Like the arrangements covered in this paper, these models rely on algorithms to produce overall scores that indicate the risk carried by a case. Although more systematic than the South Korean approach, the Canadian and Irish models led to tensions between higher ranking officers and practitioners, with officers said to ignore formal processes and request the prioritization of their cases above others. Even more comprehensive models, such as that used by Australian forces, which considered several factors, including variables like incident type, impact, response, resources, and budget, lacked detailed guidance for DF managers regarding the ranking of cases with similar high scores. While James' analysis is, at the time of publishing this article, 8 years old and arrangements may have changed, its findings highlight the compound need for proactive demand management tools and a collaborative approach between the different professionals involved in an investigation. As we discuss below, similar tensions pertain to England and Wales.

The matrix

Like policing itself, in-house forensic support in England and Wales has been subject to procedures designed to demonstrate value for money and increase accountability and transparency (Tully, 2020). Expenditure involving external providers (e.g. DNA laboratories) requires rationed and proportionate approaches, with crime scene samples being triaged before being sent for analysis. Analogous rationalities govern in-house support, such as DFUs, where demand typically outweighs resources (Rappert, Wheat and Wilson-Kovacs, 2021).

In 2011, the NPIA[2] issued the Guidance on Use of Hi-Tech Crime Unit (HTCU[3]) Case Prioritization Matrix (hereafter the Matrix), a national model of risk assessment and prioritization to manage the

[2] Established in 2006, the NPIA aimed to support police forces operationally by improving their practices. NPIA developed national services such as the Police National Computer, the National Fingerprint Database, and the Forensic21 programme.

[3] HTCUs have been renamed in recent years DFUs.

demand for DF analysis. The Matrix was part of the eForensics project, which belonged to the wider Forensics21 programme that sought to increase the effectiveness and efficiency of forensic support services. Together with the Association of Police Officers' Good Practice Guide issued in March 2012, the Matrix established the framework through which DFUs decided whether and how to accept cases and devices[4] for detailed analysis. Based on a model proposed by Parsonage and Parsonage in 2010, the Matrix was envisaged as a two-level tool to mitigate potential risks such as the loss of intelligence and evidence. It was to be applied when officers in charge of an investigation (OICs) submitted devices for examination to DFUs. Initially, a senior police officer (SIO) would review the information provided before formally authorizing the submission to ensure that the intelligence on the evidential opportunities of devices is strong and the DF analysis required by the officer is feasible. Then, a DF manager would double-check this information and accept, reject, or request further clarification from the OIC. This check would assess the value and importance of DF evidence in a case, with other factors, such as the seriousness of the offence, the risk of delaying or not examining a device, the status of investigation, and the ownership of the device, given weighting scores to reflect their importance. Based on collaborative decision-making between officers and DF practitioners, the model has given DFUs discretion over the prioritization of cases and managing demand internally.

While many police forces adopted the Matrix, several changes have occurred since. Following the disbanding of the Forensic Science Service and the phasing out of NPIA in 2012, there has been little centralized oversight of forensic support services. The establishment of the Forensic Capability Network in 2020 and the Transforming Forensics Programme aim to address this, but it is too early to assess their impact. Most recently, the National Police Chiefs' Council published its Digital Forensic Science Strategy, where a short-term goal (0–18 months) is to 'standardise priority methods and processes' (NPCC, 2020, p. 25). However, at the time of writing this, none of the forces we consulted were aware of these developments and there is no publicly available information to indicate this is taking place. Although the Matrix was to be reviewed regularly in line with force and area priorities, in practice, this has rarely happened. As we discuss below, given the unprecedented demand for DF analysis, the currency and applicability of the Matrix have been scrutinized by DFUs in relation to changes in sentencing—including new offence categories—and the introduction of accreditation and standardization of DF processes, among others.

Methodological background

Our analysis draws on data collected between 2018 and 2021 from 35 territorial police forces in the UK, as part of an Economic and Social Research Council (ESRC)-funded project on the development and application of DF in policing in England and Wales. It is based on approximately 270 hours of ethnographic observations at four DFUs affiliated to four police forces in England, 67 semi-structured interviews with stakeholders from these constabularies, including 12 DF and senior forensic managers, an additional 14 with DF managers from other forces, and numerous informal discussions with DF practitioners, gatekeepers, forensic and police managers, and various government and police stakeholders (see Table 1). The observations documented technical and administrative processes carried out by DF practitioners (including technicians, mobile device and computer examiners, gatekeepers, and managers) and officers. The interviews typically took 90 minutes and examined local practices and protocols, including how risk assessment and prioritization are undertaken and the challenges they raise. They were recorded, transcribed, and anonymized, with the extracts below referencing each interviewee's position in their organization and the interview number.[5] To supplement the qualitative data, relevant national policy and internal documents were

[4] At the time mainly computers, but the Guidance acknowledged that similar principles would also be applied to mobile phones.

[5] For instance, DF1/8 refers to a DF gatekeeper in Force 1, interview 8.

Table 1: Methodological overview

Methodology	Details and participants
Ethnographic observations	270 h observations at four forces with technicians, mobile device examiners and computer investigators, team, regional and senior managers, SIOs, OICs, detectives. and other police personnel.
81 Interviews	67 at the four forces observed (of which 12 with DF and senior forensic managers) 14 at other forces with DF managers
Survey	The four forces observed (including 11 responses from DF and senior managers)
Benchmarking	22 forces
Document analysis	Internal documents, Standard Operating Procedures (SOPs), and Service Level Agreements (SLAs) from nine forces National guidance for 43 forces

analysed. Additionally, a survey was conducted with 11 DF managers in the four police forces, and a national benchmarking request was made, to which 22 police forces responded. There are limitations to these responses, as they do not represent all constabularies and the amount of information supplied by each force varies in the level of detail provided.

Analytically, we used a thematic approach (Braun and Clarke, 2006) to examine how various stakeholders understand risk assessment and prioritization in relation to current and future demand. The initial 67 semi-structured interviews undertaken at the four DFUs examined the development and application of DF in policing in England and Wales. Demand management in the four DFUs studied emerged as a prominent concern in both the interview data and the ethnographic observations. An iterative analytical framework was adopted to hone in on this topic. To begin with, we ordered and thematized existing data and identified similarities and differences between the four forces studied. Once we established consensus over the emergent themes, these were used to produce specific questions about risk assessment for an internal survey of the four forces studied and the national benchmarking request. Subsequently, 14 additional interviews were undertaken with self-selected DF managers from other forces. The data collected through the survey, benchmarking, and second round of interviews were then thematically analysed, with new codes generated and combined to produce overarching themes. These findings were also used to strengthen the reliability of the field notes and initial interviews. To the best of our knowledge, the following analysis represents the first attempt to map current provision nationally and consider the nature, operation, and limitation of these processes in DF support services in England and Wales. The next section explains the application of the Matrix before examining key aspects of its use.

Table 2: DF risk assessment in 35 forces

Type of risk assessment	Number of forces
NPIA Matrix	25
Alternative risk assessment using THRIVE (2) and KIRAT (4)	6
No risk assessment but using case ranking process based on investigative frameworks and force priorities	4

Managing internal demand

Providing an overview of existing arrangements nationally, the benchmarking exercise revealed that 31 out of 35 participating forces used risk assessment tools (see Table 2). The four forces that did not employed a case ranking process based on investigative frameworks and force priorities to focus their resources more effectively on the most urgent cases. Twenty-five of the 31 forces that reported using risk assessment tools relied on the Matrix. Two of the remaining six preferred THRIVE—the Threat, Risk, and Harm instrument based on the National Decision Model and widely applied in operational policing (Laufs *et al*., 2021), and the other four forces used KIRAT, a tool designed to assess the risk of sexual re-offending. These alternatives were justified in terms of the currency of their variables, which were either obscured or missing in the Matrix, for instance,

victim/suspect vulnerability, court dates, the seriousness of the crime in relation to wider force priorities, and the impact on the community.

The 25 forces using the Matrix did so in either its original form or an altered version, some in combination with other risk assessment tools. Many of these forces were considering, developing, or implementing other instruments for assessing DF-related risk. This mixed picture suggests difficulties in updating the current model or finding new ones. Our interview findings support the benchmarking information: most participants in the forces that used the Matrix, criticized its effectiveness and lack of alignment with stakeholders' needs (e.g. officers, CPS, or Child Services). Participants argued that the Matrix's scoring mechanism and categories did not reflect the increase in the demand for DF examination related to other crime types. The Matrix made new types of offences difficult to risk assess because it left 'too much open to interpretation' (DF3/5). Its success was linked to available resources, the volume and types of submissions to DFUs, and the effective communication between operational and technical decision-makers (i.e. officers and DF practitioners). We examine these issues after briefly explaining the Matrix's submission, scoring, and oversight mechanisms.

The Matrix provides an evaluation of cases submitted to DFUs based on the nature of the offence, the value of other forensic evidence, the intelligence provided, the circumstances of the suspect, and the risk of harm. Its baseline builds on the most frequent type of offences DFUs encounter, that is, those related to the possession, production, and distribution of indecent images of children (IIOC), which constitute about 80% of submissions, a figure that remains constant across time and national jurisdictions (James, 2014; NPCC, 2020). However, even within this category, not all cases carry a similar risk, so to prevent bottlenecks, the Matrix helps practitioners establish the priority of a case, serving thus the dual purpose of risk assessment and prioritization.

In practice, the Matrix is embedded in the electronic systems used by OICs to submit devices to the DFUs for analysis. Procedurally, each OIC completes a formal risk assessment as part of their request for DF examination. This covers 38 yes/no questions that provide the assessor with information on various aspects of the case, such as the importance of DF evidence in relation to other information (e.g. suspect interviews, eyewitness accounts), the type of crime investigated, known risks, the identity and status of the suspect (e.g. in custody, on bail), and the likelihood that the devices submitted for analysis hold the evidence sought. OICs must choose the most appropriate options from drop-down lists of possible answers for every question. Each option has a built-in score that is invisible to OICs, to avoid the temptation to artificially enhance their scores. The sum of the values given by an OIC to each category represents the overall score, which indicates the risk posed by the case and suggests its priority. Once OICs submit their information, DF gatekeepers compare the overall score with the information provided and accept the submission or reject it, or ask the OIC for more details. If the latter, new information could result in a different score. Once revised, the case will be placed in the examination queue.

A score has two purposes. First, it will indicate nominally the time interval within which the analysis must commence. A score over 350 points, for instance, is high and indicates that work on the case should start within a month. A score between 250 and 349 is medium and specifies that analysis should begin in 2 months. A low score (150–249) suggests that the examination should commence within 3 months. Scores below 150 indicate that the case does not meet the DFU's threshold and should not be authorized for examination. Cases with low priority scores typically acquire one point for each day waiting in the queue, to ensure they advance priority wise. Second, the score also indicates the duration and extent of the analysis required. A high-score case requires more than 24 hours of in-depth specialist examination. Medium scores include 8–24 hours of analysis, and low-score cases less than 8 hours (this would typically cover triage, the extract of live data, and some limited examination). These allocations guide DF practitioners on the extent of their input and time spent on each case.

Rather than a one-size-fits-all algorithm, the Matrix requires human oversight. The guidelines of the original Matrix pilot[6] illustrate this plasticity

[6] Thanks to Leslie Charleston for making these documents available.

by acknowledging that the timescales given are 'to manage the expectations of investigating officers' and provide 'a rough indicator' for practitioners. As few DFUs could accomplish the timescales proposed, guidelines suggested that they should be treated as 'aspirational' rather than 'cast in stone'. It is also noted that 'different types of investigation require alternative strategies, some taking longer than others'. These guidelines capture the uncertainty inherent to each investigation and recognize the difficulty in providing accurate estimations regarding the length of examinations. Yet, their flexibility has been obscured in the subsequently published national guidance. Here, aspirational completion times have become target timeframes for commencing work. This shift in emphasis has given the Matrix an additional dimension as a performance measurement mechanism. Below, we illustrate how increasing tool limitations pertaining to the Matrix's variables, and operational considerations related to the ways in which the Matrix has been used to compensate for escalating demand, changes in legislation, and organizational uncertainty, have impacted its effectiveness.

Tool limitations—scenario and risk definition

DF gatekeepers often described the scenarios covered by the Matrix as 'too narrow' (DF2/6), explaining that not all types of offences are covered, and some of the offences considered are 'not afforded enough score' (DF4/3). For instance, the Matrix does not include offences such as harassment, stalking and suicide, or accounts for new categories, such as up-skirting. These omissions make it difficult to prioritize related cases. Other developments and legislative changes since the Matrix were introduced, such as bail conditions, hearing dates, the impact on victims, and CPS pressures to undertake rapid DF analysis needed for remand, were also reportedly difficult to incorporate into its use, as one DF gatekeeper explains:

> There is no way of prioritising a case if a court case is imminent, or no way of prioritising if a suspect has been arrested and the data on their phone can secure a remand. This is a situation we have had several times relating to people arrested for live abuse and the data on their phone has secured a remand which safeguards their victims, but this is not covered appropriately on the risk assessment, and I have had to write up decisions to allow me to accept these phones as a priority. (DF4/8)

DF practitioners described the Matrix as 'too IIOC heavy' (DF2/2), limited in its range and not allowing 'OICs to identify and select specific risks' (DF1/5), as another gatekeeper explains:

> There is a vast amount that can be included in 'ongoing potential risk of physical/sexual harm', for example if there is a suspicion that a suspect has been contacting an underage victim it can go into this category, it can also go into this category if the suspect is contacting a named victim and has previously assaulted them but there is no suggestion of an immediate risk, it can also go into this category if the suspect is at risk of retribution from an OCG[7] member if their drugs are stolen. It seems that a large part of the cases end up in this category even though their risks are vastly different. I have also had OICs try to use this category for a possession of IIOC case where there is no intel of any contact offending because there is a chance there might be contact offending if they are viewing IIOC, this is an example of ones where I have had to change the risk. There is also no guidance in this section regarding what is a serious risk, or other less serious risk, and the matrix scores are only 1 point apart. (DF1/3)

This extract captures some of the current classifying challenges and illustrates how assessments based on the present risk factors lead to a clustering of different scenarios, offence types, and scores, making it difficult to prioritize cases. Consequently, gatekeepers called for widening the understandings of risk that align with those used by command units,

[7] Organized Crime Group.

so OICs can recognize risks easily and provide more tailored case information. In addition, the Matrix's lack of guidance on how to deal with different types of risk leads to situations that do not 'allow OICs to clearly express the force priorities their case may be supporting' (DF4/7). These insights show how the Matrix needs to be more responsive to force-wide understandings of risk.

Risk adjustments

None of the 26 DF managers interviewed understood how the Matrix's algorithm worked, and only 2 had been involved in the original NPIA pilot. Organizational memory of how and why the Matrix was developed nationally was fragmented. Consequently, some forces adapted the Matrix to address stakeholders' concerns about its relevance, re-prioritize the identification of risk and use resources more effectively. Less radical approaches consisted of forces retaining all the Matrix's variables and measures of risk but altering the timescales for the commencement of analysis. For instance, whereas in the original guidelines, different risk levels meant that high-risk cases would start within a month, medium cases within 2 months, and low risk within 3 months, updated versions of the Matrix effectively doubled the times to commence analysis for medium and low priority cases to 3 and 6 months, respectively.

More drastic approaches involved amendments to the values carried by each of the 38 yes/no original questions, so that DF gatekeepers could understand better the algorithm behind the assessment. Other extensive changes included analysis turnaround times: whereas the Matrix indicated the time taken before analysis is undertaken, some forces modified this to indicate the time within which DFUs should complete analysis, thus increasing waiting times. Designed in collaboration with policing stakeholders from various units of command, this change tried to accommodate different types of risk and had various turnaround times, ranging from 7 days to 9 months according to the urgency of the case. Adaptations also incorporated factors such as bail and court times and sometimes combined variables pertaining to threat and risk to reduce subjective interpretations of what they may entail. At the time of writing our article, the forces that implemented such extensive changes had little information to share about the effectiveness of their new arrangements.

Operational considerations—gatekeeping

Several tensions were reported around gatekeeping, which is key to the admission of cases to DFUs. The NPIA guidelines stipulate gatekeeping should occur twice: once at the police level, when requests for examinations are submitted by OICs and approved by SIOs, and then at the DFU level, where submissions are verified by DF gatekeepers. Our observations suggest that this arrangement was not consistent: SIOs did not always review OICs' submissions to DFUs, a situation reported in other forces as well. Sometimes, SIOs had little to no time allocated for such an evaluation and performed this task as an additional duty, alongside a demanding workload. Many SIOs also had a limited understanding of what DF analysis can achieve, resulting in disproportionate requests for DF analysis (Wilson-Kovacs 2022). This led to DF gatekeepers having to follow up with OICs for more clarification and information, which was described as 'an onerous task' (DF2/1) and a 'very time-consuming exercise' (DF5). In the forces studied, DF gatekeepers verified the examination target details (i.e. what do officers need from an exhibit), established that the correct powers of seizure were used, and checked that an SIO approved the submission. Nationally, practices varied widely between forces, and the ownership of these checks could be hotly disputed. An examination of performance data in the four DFUs observed showed that almost a third of cases were placed on hold, with the required more case information before being allocated to the queue. Other forces reported similar issues: in some, 75% of all cases submitted needed re-assessment, re-scoring, and/or additional information.

Revising the automatic scores produced by the Matrix, especially when they did not match the information given by OICs, produced further tensions. Checking scores involved repeated attempts to establish rapport with the OIC and clarify the details. As previously noted, the Matrix guidance

advises that scores should not be revealed to officers, to prevent OICs from augmenting their case. Because officers cannot see the scores of their submissions, DF gatekeepers widely reported that some chose answers that increased the likelihood of their submission being accepted. For instance, when asked whether digital evidence is key, secondary, or peripheral, some OICs would choose the first option even when evidence other than digital was available and relevant. Some DF gatekeepers explained this practice with reference to the opacity of the Matrix and OICs' fragmented understanding of the submission process. Enhancing scores to increase the chances of having submissions accepted for analysis and dealt with quickly, was also acknowledged by some of the officers interviewed, who attributed it to being pressured by their supervisors to submit devices for DF examination even when investigative strategies were still under construction.

Operational considerations—discretion

Discretion is an important aspect of how the Matrix is used, as the timeliness with which a case is considered for analysis and the number of items examined. The submission system cannot be openly circumvented: as described above, examination requests must be submitted electronically, and submissions must score at least 150 points to be analysed by DFUs. Given the amount and quality of information OICs submit, DF gatekeepers must keep an open mind in their interpretation of case details and adjustment of scores. Those we interviewed agreed that OICs faced difficulties in submitting accurate and comprehensive case details and noted that despite guidelines explaining the process, officers do not understand what information is required. This can be because OICs lack the time to enter full case details into the system, and/or are not familiar with the submission procedure. Occasionally, DF gatekeepers remarked that OICs and SIOs expect to have their cases examined as soon as they are submitted because of their experiences with more established in-house forensic support services, such as CSI and fingerprinting. Often, these expectations are held by unrealistic Service Level Agreements (SLAs) between the DFUs and the forces they serve and dulled by the demand faced by DFUs.

Informal interactions between members of staff can impact on the number of items accepted for analysis and the speed with which cases are analysed. For instance, the units dealing directly with child sexual abuse cases are usually co-located with DFUs, because of the frequency with which their investigations need DF assistance. Physical proximity affords interaction, fosters familiarity, and presents the opportunity for face-to-face negotiation, especially when more work than that stipulated in the SLA is required (as some SLAs limit the number of items DFUs accepted for analysis). When neither intelligence nor triage can reduce the number of exhibits holding potential evidential information, some officers would use persuasive tactics to have all their items accepted, as one officer explained:

> When I go in there… I'll buy them cakes… I do whatever to say 'look I know it's only three, but these (extra ones) are only little… can you just have a look… and (Team Leader) can't say no to you. (OIC3)

Here, cake and civility act to soften rules and allow more items to be accepted for examination. Similarly, anecdotes of cases being moved up the queue following interventions from SIOs to senior forensic managers were recalled.

Case interpretation and capture

Although by design the Matrix requires human input and interpretation, little guidance has been provided on their nature and extent. For this reason, some of the DF gatekeepers interviewed describe assessing risk as 'a very subjective area' (DF2/2). We explained above how instead of ensuring a more accurate reporting of case details, the invisibility of the Matrix's scores for OICs has the opposite effect and leads to extensive checks. Scoring a submission can also differ between DF gatekeepers. The same case information can be evaluated differently by DF gatekeepers from the same DFU, which can impact on whether a case is deemed high, medium, or low priority, or indeed, whether it is accepted in the queue. DF gatekeepers noted that when such discrepancies occur, in-team discussions help clarify the priority of a case.

Interpreting risk using the Matrix relies on submitting correct and comprehensive case information. Filling in the form used to capture the details of a case '(m)akes (officers) think about their case, (and) helps open discussions about specific areas of case they may not otherwise have thought to tell you' (DF2/6). However, the 'form needs to direct them' (DF1/4), which appears not to be the case, both in the forces studied and others whose managers we interviewed. DF gatekeepers asked for better contextual assistance for OICs to complete their submission, more training with submission requirements, and clear step-by-step guidance and comprehensive explanations of the process for OICs. However, opinions on the usefulness of the information required from OICs varied. For some DF gatekeepers, the OIC was more aware of their own cases and the risk they carried. For this reason, unlike DF practitioners who did not 'know every detail' (DF3/1), OICs 'hold overall responsibility for the entire case' (DF1/1):

> We can only assess the risk from the information provided by the OIC and may not be aware of situations outside of the main circumstances of the case that may have an effect on the level of risk of the case. (DF4/8)

Alongside responsibility, these DF gatekeepers argued that OICs have a greater overall understanding of risks outside the crime, for instance:

> A father not being able to see children is a risk not immediately apparent for an IIOC possession case, or if the suspect or victim is at risk of physical harm due to the offence, e.g. drugs and cash being seized form dealers lower in the chain could be at risk, and this would not be apparent just reviewing the circumstances of the crime. (DF4/4)

While some DF gatekeepers mentioned that the move to the electronic submission system may have inadvertently relaxed oversight procedures for OICs, they all agreed that officers needed to understand the proportionality of their own requests for examination and to demonstrate a reasonable line of inquiry in their submissions. As risk assessing outside one's jurisdiction, across command units, and using different tools can lead to tensions and discrepancies, it was argued that gatekeepers should not be responsible for decisions about the proportionality of cases and other legal and/or policing aspects.

In contrast, a smaller number of DF gatekeepers believed that delegating the risk assessment process to the DFU would guarantee more consistency than relying on OICs for interpretation:

> We need to ensure what the OIC puts down in the exam target is clear and comprehensive so that we can assess the risk better. Doing the risk assessment ourselves ensures consistency rather than relying on 100s of officers to interpret the guidance. We will always need to discuss with officer when things are unclear but feel this would be a more efficient process. (DF2/3)

Notwithstanding, all DF practitioners called for dialogue 'when things are unclear' (DF1/3): a process seen as more than the simple exchange of information and described as 'discussion' (DF3/6) and a 'collaborative effort' (DF3/4) that simply 'just needs to be better' (DF1/7).

It was widely agreed that the scoring system and the assessment variables of the Matrix must be reviewed to include other offence types and clarifications on how such offences, for instance, harassment, should fit into the current assessment. More guidance was also required to assist DF gatekeepers with the prioritization of cases with similar high priorities and scores, and assessing risk, prioritization, and the interpretation of variables such as 'position of trust' and 'threat of future harm'. Finally, many forensic managers requested more information on how the Matrix is used by other forces to identify and adopt evidence-based practices: '(The Matrix) should be reviewed with current practices from a select few forces nationally to see if they employ a better method or best elements of a method that we can employ' (DF2/6).

Concluding remarks

While the constant flux of patterns in demand for investigations impacts on the internal demand for specialist services and processes provided by forensic science support to police forces (Walley and Adams, 2019), analyses of such demand are scarce. This article offered an exploratory and empirical analysis of the challenges raised by current DF risk assessment and prioritization procedures in 35 police forces in England and Wales. Since 2012, many of these forces have used a national risk assessment tool—the Matrix—designed to evaluate cases submitted to DFUs for analysis. However, a rise in requests for specialist DF examinations across different types of offences, and several socio-legal changes in the last decade led to stakeholders questioning the usefulness of the Matrix and seeking various alternatives to cope with investigative requests.

In its outer-facing capacity, the Matrix is an algorithmic tool that helps to evaluate the risk carried by each submission to DFUs and serves as an essential entry point for submissions. In its inner-facing capacity, it orders cases using a specific understanding of risk based on the most frequent type of offence DFUs encounter (i.e. those involving IIOC and their possession, distribution, and making). This approach results in offences whose risk is either not recognized or diminished by the algorithm used, being given less importance. To adjust for this and accommodate requests for the DF examination of non-IIOC-related cases, lengthy discussions with OICs and revision of the information submitted are typically undertaken. The Matrix scores are also adjusted to accommodate legislative changes and other policing priorities.

The Matrix's application opens to scrutiny different and sometimes conflicting, understandings of internal demand and risk. In line with critiques on the use of predictive policing and scholarship on algorithmic management (e.g. Brayne, 2020; Gundhus, 2021), we showed how rather than being a dispassionate and objective decision-making tool, the Matrix requires human input, interpretation, and discretion. Blurred lines of ownership and accountability, a fragmented operational understanding of what DF can achieve, what information DFUs need to process submissions, and the difficulty of justifying Matrix-driven decisions to policing stakeholders, lead to tensions between DFUs and operational units and demonstrate the need for more systematic approaches to how this type of internal demand is managed.

Like other studies on managing internal demand in policing (e.g. Laufs *et al.*, 2021; Morris and Walley, 2022), our analysis highlights how demand through risk assessment and prioritization processes can result in additional, unplanned activities that slow down investigations. Moreover, a lack of centralized guidance and monitoring nationally has led to considerable uncertainty about how forces deal with multiple, changing priorities and the surge in requests for DF examination at a time when investigation has come to rely increasingly on them (Walley and Adams, 2019). Lessons are rarely shared between forces, and even within the same constabulary, evaluation of these processes and foresight are often missing. This uncertainty limits the identification of good practice nationally and impacts long term on the effectiveness of the effectiveness of advice on how forces should adapt to escalating demand. It also reflects the reactive dynamics of more established forensic support services, where the capacity to think about the contribution of such services more wholistically is missing (Horsman, 2017; Houck, 2020; Tully, 2020).

Our analysis illustrates how an updated understanding of managing risk is needed to address the demand for DF analysis more effectively. This should not be linked to performance management (Houck, 2020; Morris and Walley, 2022) and encourage the forecasting of internal demand collaboratively. It should break the siloed approach between the technical and operational sides of investigation (Rappert, Wheat and Wilson-Kovacs, 2021; Wilson-Kovacs, 2019, 2022), and be responsive to changes in policing priorities and flexible enough to adapt national guidance to the demand experienced by local forces in order to ensure an equitable access to justice for defendants.

Future research on how to address more effectively the demand for DF examination needs to provide a more systematic understanding of such demand in the first place. Establishing how to capture internal demand for DF examination locally,

regionally, and nationally, the needs of operational units for DF examination, how such needs fluctuate, and the difficulties presented by different types of requests, are some of the most important questions raised by our analysis. Answers to them would provide both a solid basis for building centralized risk management tools in DF and identifying good practices. Disseminating examples of the latter and assisting forces with their struggles to manage demand for DF examination, through centralized, national forums such as the Forensic Capability Network, can help individual forces scrutinize their own internal demand and reflect on how to organize better their own resources. Such initiatives require extensive dialogue, cooperation, and foresight not only between DFUs nationally but also between DFUs and the forces they are part of, and more insight into the challenges of this endeavour is also needed.

Acknowledgements

We thank our participants for their insight, time, and support and to our reviewers for their feedback. For the purpose of Open Access, the author has applied a 'Creative Commons Attribution (CC BY)' licence to any Author Accepted Manuscript version arising. The support of the Economic and Social Research Council (Research Grant ES/R00742X/1) is gratefully acknowledged.

References

Alawadhi, I. (2019). *Methods and Factors Affecting Digital Forensic Case Management, Allocation and Completion*. Doctoral Thesis, University of Central Lancashire. http://clok.uclan.ac.uk/30744/1/30744%20Alawadhi%20Ibtesam%20Final%20e-Thesis%20%28Master%20Copy%29.pdf (accessed March 2021).

Braun, V. and Clarke, V. (2006). 'Using Thematic Analysis in Psychology.' *Qualitative Research in Psychology* **3**(2): 77–101.

Brayne, S. (2020). *Predict and Surveil: Data, Discretion, and the Future of Policing*. New York, USA: Oxford University Press.

Casey, E., Ferraro, M., and Nguyen, L. (2009). 'Investigation Delayed Is Justice Denied: Proposals for Expediting Forensic Examinations of Digital Evidence.' *Journal of Forensic Sciences* **54**(6): 1353–1364.

de Braekt, R. I., Le-Khac, N., Farina, J., Scanlon M., and Kechadi, T. (2016). 'Increasing Digital Investigator Availability Through Efficient Workflow Management and Automation.' 4th International Symposium on Digital Forensic and Security (ISDFS), pp. 68–73.

Horsman, G. (2017). 'Can We Continue to Effectively Police Digital Crime?' *Science & Justice* **57**(6): 448–454.

Houck, M. M. (2020). 'Backlogs Are a Dynamic System, Not a Warehousing Problem.' *Forensic Science International: Synergy* **2**: 317–324.

Gundhus, H. (2021). 'Shaping Migrants as Threats: Multilayered Discretion, Criminalization, and Risk Assessment Tools.' *International Journal for Crime, Justice and Social Democracy* **10**(3): 56–71.

James, J. I. (2014). 'Multi-Stakeholder Case Prioritization in Digital Investigations.' *Journal of Digital Forensics, Security and Law* **9**(2): 59–72.

National Police Chiefs Council (NPCC). (2020). *Digital Forensic Science Strategy*. https://www.npcc.police.uk/Digital%20Forensic%20Science%20Strategy%202020.pdf (accessed February 2021).

Laufs, J., Bowers, K., Birks, D., and Johnson, S. D. (2021). 'Understanding the Concept of 'Demand' in Policing.' *Policing and Society* **31**(8): 895–918.

Morris, G. and Walley, P. (2022). 'Implementing Failure Demand Reduction as Part of a Demand Management Strategy.' *Public Money & Management* **42**(1): 22–31.

Parsonage, H. and Parsonage, H. (2010). *Computer Forensics Case Assessment and Triage – Some Ideas for Discussion*. http://computerforensics.parsonage.co.uk/triage/ComputerForensicsCaseAssessmentAndTriageDiscussionPaper.pdf (accessed March 2021).

Rappert, B., Wheat, H., and Wilson-Kovacs, D. (2021). 'Rationing Bytes: Managing Demand for Digital Forensic Examinations.' *Policing and Society* **31**(1): 52–65.

Tully, G. (2020). *Forensic Science Regulator Annual Report 17 Nov 2018 - 16 Nov 2019*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/877607/20200225_FSR_Annual_Report_2019_Final.pdf (accessed February 2021).

Walley, P. and Adams, M. M. (2019). *An Evaluation of Demand Management Practices in UK Police Forces*. Milton Keynes: Centre for Policing Research and Learning.

Wilson-Kovacs, D. (2019). 'Effective Resource Management in Digital Forensics: An Exploratory Analysis of Triage Practices in Four English Constabularies.' *Policing: An International Journal* **43**(1): 77–90.

Wilson-Kovacs, D. (2022). 'Digital Media Investigators: Challenges and Opportunities in the Use of Digital Forensics in Police Investigations in England and Wales.' *Policing: an International Journal of Police Strategies and Management* **44**(4): 669–682.