Title: Quality Standards for Digital Forensics: Learning from Experience in England & Wales

Authors:

Author	Affiliation
Gillian Tully	Forensic Science Regulator
Neil Cohen	Defence Science and Technology Laboratory
David Compton	United Kingdom Accreditation Service (UKAS)
Gareth Davies	University of South Wales
Roy Isbell	University of Warwick
Tim Watson	University of Warwick

Corresponding author: Gillian Tully

Gillian.tully@forensicscienceregulator.gov.uk
Forensic Science Regulation Unit
5 St Philip's Place
Colmore Row
Birmingham
B3 2PW

Quality Standards for Digital Forensics: Learning from Experience in England & Wales

Abstract

The Forensic Science Regulator has the role of setting quality standards for forensic science in the Criminal Justice System (CJS) in England and Wales. The current requirement is for organisations carrying out digital forensics to gain accreditation to the international standard ISO/IEC 17025 and the Forensic Science Regulator's Codes of Practice and Conduct. The aim of this requirement is to embed a systematic approach to quality, including understanding methods, validating software and systems, understanding risks, ensuring that all involved in the crime scene to court process have the skills and competence they need and the appropriate equipment and environment for the work, and providing ongoing assurance of quality through audit and proficiency tests. However, the challenge of implementing the standards in digital forensics should not be underestimated, particularly in an environment where there is insufficient capacity to meet a growing demand for services in an area of increasing complexity and fragmented delivery. It is therefore timely to review available data to determine the extent to which accreditation to ISO/IEC 17025 is addressing quality issues in digital forensics and consider what changes and resources could be made available to assist with implementation of quality systems.

Keywords quality assurance; quality standards; accreditation; skills; competence; validity; regulation

1. Introduction

Digital forensics is the process by which information is extracted from digital systems or data storage media, rendered into a useable form, processed and interpreted for the purpose of obtaining intelligence for use in investigations, or evidence for use in criminal proceedings. The scope includes, but is not restricted to, aspects such as remote storage and systems associated with computing, imaging, image comparison, video processing and enhancement (including CCTV), audio analysis, satellite navigation, communications; emerging technologies will also form part of the scope. Digital forensic methods will typically include evaluation of the approach to be taken, choice of tool(s), quality checks and production of reports. Ensuring that digital forensics, like all forms of forensic science, is delivered to the appropriate level of quality for its use in a CJS is not in itself contentious. Casey (2019) gave the stark warning that "as more criminal investigations involve digital traces in increasing amounts and complexity, the quality of digital forensic results is decreasing and comprehension of cybercrime is diminishing" and Jones and Vidalis (2019) warned of increasing concerns with regard to the veracity of commercial tools relied on by digital forensics practitioners; this concern underlines the need to ensure that

methods used in the CJS, of which tools are a part, are validated and their limitations understood.

Introduction of new types of scientific or technical evidence in a criminal justice context without sufficient scrutiny has led to assumptions of validity that were unjustified and to some practitioners giving opinion in a range of physical forensic science disciplines over many years, using methods lacking scientific rigour (e.g. Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council, 2009; President's Council of Advisors on Science and Technology, 2016; Ward *et al.,* 2017). A particularly extreme example, where comparison microscopy of hairs was used to reach conclusive opinions on identity has led to a review of every such case carried out by the Federal Bureau of Investigation (Federal Bureau of Investigation / Department of Justice, 2015). The impact of such failures to apply scientifically robust principles is far-reaching in criminal justice terms, in personal terms for those directly impacted and in financial terms, given the vast cost of retrospective reviews.

However, the means by which quality of digital forensics should be assured has been the topic of intense debate (e.g. Casey, 2006; evidence to the House of Lords inquiry into forensic science, 2018-2019; Jones and Vidalis, 2019; Marshall and Paige, 2018; Page *et al*., 2019; Sommer, 2018). To provide context to the debate, Figure 1 illustrates the types of standards and guidance discussed.

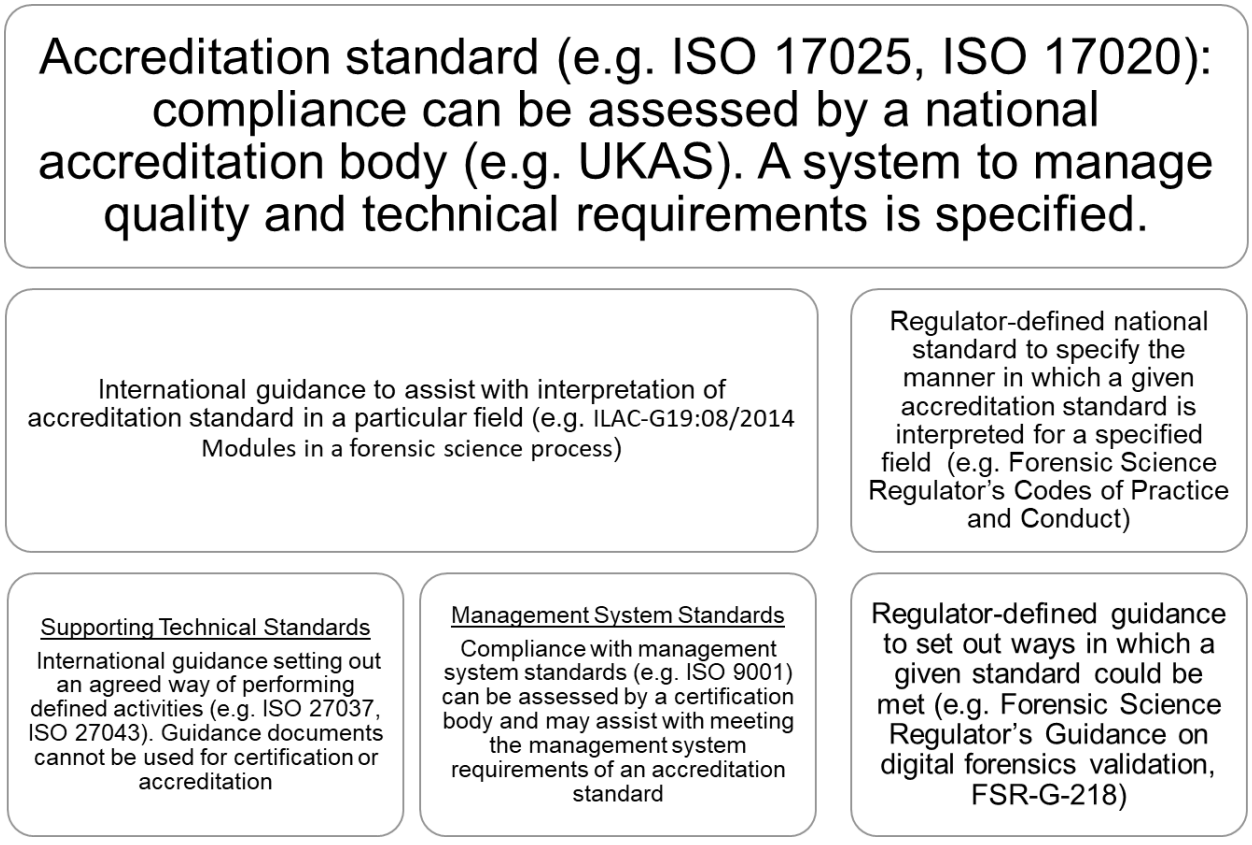

Figure 1: Types of standard. The accreditation standard ISO/IEC 17025 sets out what must be achieved, and not a detailed description of how it must be done. Setting a standard at this level

enables organisations to innovate and find the way of achieving the standard that suits them. Organisations may choose to follow additional guidance (such as that in ISO 27037) to enable them to meet some of the high-level technical requirements in ISO/IEC 17025 or may choose to define their own manner of operating. ISO/IEC 17025 covers the activities not just of practitioners at the bench, but also the system in which they work, with requirements for "top management" to be accountable for quality; audit, review and improvement are all expected, A full description of the ISO 27000 series is given by Cusack (2019).

Sommer (2018) argued that a "messy" combination of a range of standards for data acquisition, together with a state-sponsored certification scheme for individuals, case by case scrutiny in courts and "informed informal recommendations" may be the best approach. His concerns related to the applicability of accreditation to the international standard ISO/IEC 17025[1], included costs and the unique features of digital forensics, for example the rate of change and need to deploy novel techniques rapidly. We note that the Forensic Science Regulator's Codes of Practice and Conduct (2017; hereafter referred to as "the Codes") allow for use of novel methods prior to accreditation provided that the customer understands the extent to which it is validated prior to commissioning. This allowance is not intended as an alternative to the usual requirement for accreditation, but reflects the need on occasion, particularly in the field of digital forensics, for novel methods to be introduced more quickly than accreditation can be obtained. Jones and Vidalis (2019) stated that ISO/IEC 17025 is not fit for purpose in relation to digital forensics but gives no specific justification for this assertion. Whilst noting that external evaluation is not an end in itself, Casey (2006) argues against "trading justice for cost savings", given the potential impact of digital evidence on individuals' freedom. Page et al (2019) note that as among the newest forensic science disciplines, digital forensics could have built on learning from established disciplines, but arguably has the least robust quality management procedures; they conclude in favour of integrating additional quality measures and meeting the requirements of ISO/IEC 17025 and the Codes. Marshall and Paige (2018) argue that, provided clear technical specifications are set, verification and validation requirements in ISO/IEC 17025:2005 and ISO/IEC 17025:2017 are achievable for digital forensics. Their preferred approach, where such specifications are publicly available, has several advantages including the potential for users of tools to influence more effectively the development of such tools, the potential for verification and validation of tools and methods to be simplified and effort shared between organisations and more equality of arms for defence review of evidence produced by digital forensics units instructed by the prosecution. Casey (2016) proposed that differentiating between technical processes and scientific processes helps determine what knowledge, training and other elements of quality assurance are fit for purpose. In such a model, technical processes such as making forensic copies of digital evidence, extracting all active and deleted files, observing data and running presumptive tests such as automatically checking for potential child pornography, may need a quality assurance regime which is different

[1] ISO/IEC 17025 General requirements for the competence of testing and calibration laboratories.

from that required for evaluation of digital evidence. Others, (e.g. Sommer, 2018 and one of the authors of this paper (GD)) also point to differences between the more routine activities, such as evidence acquisition and preservation, where accreditation may have a place, and the more complex digital forensic activities such as interpretation and evaluation, where accreditation may not fit. England (2018) argued that ISO/IEC 17025 is ill suited to digital forensics and that the Regulator's validation requirements are placing a "near unmanageable" burden on digital forensic providers. A number of Accreditation Bodies have opted to assess forensic science laboratory-based activities against the requirements of ISO/IEC 17020[2] as it has more emphasis on the use of professional judgement. The different approaches adopted by National Accreditation Bodies is recognised in the international guidance ILAC G19[3] which emphasises the requirements for interpretation, quality assurance and validation irrespective of which standard is used. In addition, ILAC G27:06/2017 – "*Guidance on measurements performed as part of an inspection process"* provides guidance on additional requirements e.g. validation, for inspection activities where relevant. In the UK, ISO/IEC 17020 will be required for screening, capture and preservation or analysis of data from a device conducted at scene (including but not limited to Servers and Routers) from October 2020, but no UK-based organisations have yet been accredited for these activities. The workload associated with validation can perhaps only be tackled by national and/or international collaboration as the scope and complexity of digital forensics grows.

The authors believe that it is important to move the debate from theoretical considerations of the applicability of certain standards, extrapolation from small datasets and/or anecdotal observation to an examination of the data. In this paper, therefore, we seek to contribute to the debate on the basis of a significant data set, assessing the impact, value and costs of accreditation to ISO/IEC 17025 and the Codes, based on data. Two primary sources of data are included: findings from 61 initial assessments in 30 organisations and 29 surveillance visits to 29 accredited units[4] by the United Kingdom Accreditation Service (UKAS) from 2015 to 2019, and 48 quality referrals to the Forensic Science Regulator (the Regulator) between 2012 and 2019.

2. Methods

2.1 UKAS Assessments

From 2015 to the end of August 2018, UKAS undertook approximately 61 initial assessments or first extensions to scope of Digital Forensic Units (DFU) for the

[2] ISO/IEC 17020 Conformity assessment — Requirements for the operation of various types of bodies performing inspection
[3] ILAC-G19:08/2014 Modules in a forensic science process
[4] Several different units were within one legal entity, but each unit held accreditation independently from the others.

different digital disciplines (e.g. computer, phones, video). The findings raised during these initial assessments were reviewed, with a focus on the findings raised. In addition, feedback from UKAS Technical Assessors was collated to identify general trends identified on visits; a summary of these data was submitted by UKAS as further supplementary evidence to the House of Lords inquiry into forensic science (UKAS, 2019). Subsequently, findings from 29 surveillance assessment visits by UKAS to accredited DFUs between September 2018 and April 2019 were collated. All data from assessments were anonymised, to maintain the confidentiality of the assessment process.

2.2 Referrals to the Regulator

The Codes require escalation to the Regulator of issues that have potential to attract adverse public interest or lead to a miscarriage of justice. Organisations which hold accreditation to the Codes are therefore assessed against this requirement and over time, will embed appropriate escalation requirements. As quality standards have been introduced across forensic science disciplines, there has been an increase in reporting of problems for each. This trend is reproducible across disciplines and suggests that implementation of quality standards increases reporting and dealing with problems. Other than self-referral, problems can come to the attention of the Regulator through expert review of the evidence in a case, by concerns raised by a judge or other trial participant or reports by concerned third parties.

Referrals from April 2012 (the first digital forensics referral) until the end of August 2019 were collated. Data from the referrals were anonymised, since the purpose of the referrals system is to ensure identification of root causes of problems and implementation of actions to reduce the risk of recurrence; it is not to attribute blame.

3. Results

3.1 Data from UKAS Assessments

The number of organisations holding accreditation for digital forensics activities has risen. As of 31 July 2018, 30 legal entities held accreditation for one or more digital forensic activities at one or more locations. The accreditations covered 75 different physical locations and were held by 6 commercial companies, 3 government organisations and 21 police forces. As of April 2019, the number of legal entities holding accreditation for one or more digital forensic activities at one or more locations had risen to 47, spread over 101 physical locations. These were held by 11 commercial companies, 3 government organisations and 33 police forces. A number of the police forces which hold the accreditation additionally take responsibility for work undertaken at DFUs situated within other police forces, governed by collaborative agreements.

During 2017, two commercial organisations had their accreditation suspended temporarily; both had their accreditation reinstated when improvements had been made to demonstrate compliance to the applicable requirements.

Between 2000 and 2010, UKAS accredited a further 3 commercial organisations (not included in the numbers quoted for July 2018 or April 2019), which subsequently resigned their accreditations when they ceased to offer digital forensic services.

The scopes of the current accreditations are shown in table 1.

Digital Forensic Activities	Number of legal entities accredited (August 2018)	Number of legal entities accredited (April 2019)
Computer – Triage	0	3
Computer Imaging	28	43
Computer analysis	8	16
CCTV	3	5
Phones	16	19
Sat Nav	2	3

Table 1: Accredited digital forensics activities

During the 61 initial assessments carried out between 2015 and the end of July 2018, 3,083 findings were raised in relation to adherence to ISO/IEC 17025 and ILAC G19 requirements with 2,972 being mandatory findings (a non-conformity requiring action to be taken to become compliant) and 111 recommended findings (suggested improvement action). A breakdown of average numbers of mandatory findings per discipline is provided in Table 2.

Digital Forensic Discipline (No of assessments)	Average number of mandatory findings per visit	No. of assessments resulting in no offer or a restricted scope.	% of assessment requiring extra visit	Average length from application to grant (months)
Computer (36) (Imaging and analysis combined)	50	7	53	20
Phones (20)	47	7	55	14
Video/CCTV) (5)	44	0	100	21
Overall	49	14	56	19

Table 2: Outcome of initial assessments

Table 3 details the findings raised against the different areas of ISO/IEC 17025 per discipline and as a whole for the digital forensic assessments.

Mandatory Findings (related area)	% of computer findings	% of phone findings	% of video/CCTV findings	% of total Mandatory findings raised
QUALITY RELATED				
Organisation and Management	2.5	3.3	2.5	2.7
Quality Management System Documentation/Records	20.5	16.0	2.0	17.6
Sub-contracting	1.3	1.2	0.4	1.2
Service/Suppliers	2	1.9	2.1	1.9
Non-conforming work and complaints	1.5	1.9	2.2	1.7

Audits	4.4	5.1	2.5	4.4
Management Review	0.5	0.4	0	0.4
TECHNICAL RELATED				
Contract Review / Customer Requirements	3.2	2.6	5.9	3.2
Training	4.1	4.4	3.4	4.1
Competency	1.9	2.4	4.7	2.3
Procedures (lack of detail)	17.1	17.0	17.8	17.1
Practice (Poor or no following procedures)	2.6	3.6	3.4	3.0
Validation	12.2	16.7	17.4	14.0
Ongoing Quality Assurance	4.5	7	8.9	5.6
Technical Records	6.4	5	9.3	6.1
Equipment	6.7	4.4	6.4	5.9
Accommodation	1	0.9	0	0.9
Reference Material	2.9	2.7	2.1	2.8
Exhibit Handling/Continuity	3.6	1.8	8.9	3.4
Reporting	1.1	1.7	2.5	1.4

Table 3: Breakdown of findings raised by area of ISO/IEC 17025 requirement

In addition to the numerical analysis of findings, a qualitative view on the areas where findings of the highest significance were identified was collated from feedback from technical assessors. A summary of this qualitative analysis is given in Table 4. Any example findings have been anonymised as the intention is to illustrate key points learned during the assessment process and not to criticise individual DFUs. In some instances, anonymisation has required minor changes to wording but in no instance has the meaning of the words been altered.

Area in which findings raised	**Summary of Issue**	**Anonymised example non-conformance findings**
Technical procedures	Technical procedure documents were either missing or contained insufficient detail to ensure consistent application or effective direction to staff on what should be undertaken on a routine basis. When variation between forensic staff was observed during an assessment, it was unclear what the expected procedure was.	SOP-X does not describe in sufficient detail the procedures to follow to cover all stages of the extraction and examination process e.g. guidance on how to handle exhibit, order of examination, which tools to use, what equipment to use, what settings should be used for Imager, how to verify output.
Technical practice	There was often variation in practices being employed within the same digital forensic	The x locations under this application adopt a different approach to verification of the data extracted from mobile devices.

	units. The main findings raised related to staff not following the unit's documented procedures, others related to good practices which were not being shared within the unit and on a few occasions poor practices were witnessed. While this was not one of areas where a large number of findings were raised those raised do highlight inconsistent practice within units and lack of previous standardised approaches.	An inappropriate functionality test of a laptop was performed after imaging which could have made changes to the evidential media. The performance of this test (or otherwise) and other elements of re-assembling devices was not covered in the Standard Operating Procedure (SOP). During a witnessed test of imaging an [computer] the laboratory staff member did not use the correct key to control the start-up process and this caused it to boot-up on three occasions. The laboratory was unable to demonstrate the required level of knowledge and competency in the use of this equipment.
Technical records	The technical records being made and retained by a number of DFUs were weak. Instances of poor photography were witnessed along with a lack of detail in associated notes such that work could be repeated, or critical decisions and findings identified. Electronic notes were not always traceable to the individual making them, nor was there the ability to identify subsequent changes to the notes, made either intentionally or unintentionally. There were issues identified with the back-up processes for electronic information and security of IT equipment.	The notes produced for the witnessed activities would not enable another expert to identify any critical issues with the process and identify any trends at a later date. It would further not allow for another expert to examine if an appropriate course of action had been followed during the examination. Imaging notes are recorded contemporaneously but are not protected in a manner which prevents alteration. They are created and held in a digital format within a word document. This is available in a mutable format within shared directory structure for the case. Changes are not identifiable at the time of creation of notes. Record #AA shows that Person X imaged a hard drive when the image logfile shows it was done by Person Y. During the witnessing of 1234/17 a laptop exhibit fell from the table, and whilst no physical damage was observed the event was not recorded. There is no procedure or documented policy for the backing up and archiving of case data. Case data is solely stored on the workstation of the examiner who completed the examination. A review of Person B showed that the "backup" drives contain n cases (c.300) which are not backed up or archived.

Training and competency	A common finding raised was that DFUs did not have objective evidence to demonstrate the competency of their staff, other than relying on staff attending courses and having x years of experience. Another common finding related to DFUs not having a mechanism to demonstrate on-going competence once individuals had been initially deemed competent.	During the witnessed test of extraction and processing of data method using the software tool X, the staff member verbally expressed unfamiliarity and a lack of training with it and then struggled to navigate through some settings and demonstrate competence in the process. Staff have been authorised as competent to perform tasks in the digital forensic unit; however, the manager that has conducted this evaluation is not technically competent to do this. The current requirements for evidence of training and competency consists of one test, this is not sufficient to demonstrate competence given the number of variables that can be encountered.
Validation	The issues with validation were common across the different digital forensic units and related to the fact that the initial validation focused on tool verification and not overall method validation. In addition, the tests undertaken did not cover the significant risks in the process or appropriately stress test the method (appropriate test data or devices), there was no identification of main uncertainty in the method and the lack of evidence to demonstrate that the method is repeatable within units.	The equipment used in Method Validation of [x type of computer] imaging is faulty and not fit or purpose. The results of Method Validation work have not been communicated to staff engaged forensic processes in scope, therefore they are not aware of the details of success and or limitations. The Digital unit have accepted misleading CRC errors in two of the verification tests for Software Y. If these errors occurred in live case work this may result in evidence being lost. Two errors were identified within the validation log spreadsheet relating to 'altered hash', these errors had not been picked up and the validation had been signed off. The validation testing proved the software write blocking system used in the method is not totally fit for purpose. No other tools were tested or is being used. Assessments of whether software updates would, or would not, trigger further validation are inadequate and not implemented.

		The submitted validation material did not provide any assurance that the overall method was fit for purpose as it focused on tool capabilities.
Ongoing quality assurance	Assessments identified that the DFUs had often not implemented a robust on-going quality assurance mechanism. Some staff were undertaking informal dip checking of their own work but there was no structure or consistency. None of the existing quality checks involved a robust assessment of the technical validity of the work so none would provide assurance that any significant amount of data had not been missed or that tools had been used appropriately.	The quality assurance mechanism at present does not provide on-going assurance on the reliability of work being delivered out of the unit. The QA process on each case provides assurance that notes are appropriate but there is nothing in place to provide assurance that data has not been missed or the technical work has been undertaken appropriately. The Quality Assurance SOP is not clearly understood by lab staff. The QA process has not been implemented. There is no clear guidance specifying non-conformance standards. As a result, lab staff are unclear when non-conformance in imaging processes should be raised. For the ILC imaging trial undertaken in 2016, differences in outcome achieved by the two sites have yet to be subject of effective investigation and formally recorded within the QMS. ILC Round 2 – Two sets of test material provided. One fail due to hash mismatch with the expected hash. Investigation to date has not identified the route cause and therefore cannot provide sufficient assurance that the error has not occurred before and does not continue to occur. The comments in the peer review folder are not clear and some are irrelevant. There is no evidence of any follow up action on the comments made nor is there consistency on case notes recording that the review has happened. Dip sampling of mobile device cases has only been performed for a relatively short period. A relatively high number of fail and remedial results appeared to be identified. The reviews have not yet supplied enough assurance that the required quality is being achieved.

		There is no actual policy in place for the dip sampling and thus no triggers for additional actions on significant levels of negative results.
Reporting of results	Results are provided to customers of DFUs in a multitude of formats. Some outputs involved the customer receiving multiple extractions of data generated via different examination tools, containing similar but not identical information, with no indication which is the best one to use. DFUs rarely produce statements or attend court. Frequently, staff have not received appropriate court awareness training and do not have access to template statements if a request was received. In addition, knowledge of the criminal procedure rules was mixed.	The current output from the Unit could include two extractions with different output for a specific subset. This output is ambiguous, and it is not clear to the customer the reasoning's for this process and the potential differences between the outputs. SOP-XYZ states that reports are produced as a PDF and an .xls 97-2003 file. The row limit of a .xls file is 65,536. Reporting to a .xls file without consideration on data set size could lead to large quantities of data being left out of reports. Statements are being issued by staff that are not signed off as competent to complete the work. There is no reference in the statements and contemporaneous notes that the staff member is not signed off or reference to who has taken responsibility for the work completed e.g. the mentor is not referenced.
Exhibit handling	Exhibit handling in the main was appropriate, however, it was common to find poor records relating the chain of custody for an exhibit. DFUs often accepted poorly packaged and labelled exhibits. This was a more significant issue in the video/CCTV field.	The Digital unit does not have control over the integrity/continuity of an exhibit when it is left unattended during the acquisition process. There are situations where exhibits are received by one person or on a different day (e.g. example witnessed received 08/07 but booked in 13/07 by different person) and therefore the records used by staff to demonstrate chain of custody do not correspond with information on the transit document (e.g. date and person receiving item from driver). In addition, there is no retained hard copy signature for items received or returned by hand. The current systems in place do not show a clear chain of custody of item movement within the department.

		There is a lack of records demonstrating the movement of items. There is no procedure in place that covers the preservation of exhibit integrity within the laboratory. During witnessed activities exhibits were received in varying states offering different levels of security and protection to the exhibits. • Exhibits were received in open exhibit bags and no notes were made to record the condition of the item. • One item was received in an open envelope with no exhibit bag, again the condition was not recorded in notes.
Other		The forensic workstations in the unit are connected to the internet. This creates a number of risks from viruses / malware and unauthorised access to case data. There have been insufficient internal audits conducted to demonstrate effective implementation of the quality system and procedures and integration of the Digital team into this. The audits which were reviewed had not highlighted a level of non-compliance commensurate with what was found during the UKAS assessment. It is therefore unclear as to whether the appropriate breadth and depth of reviewing of the processes associated with this ETS have been included in audits. The main server room is an inappropriate lab environment. Server cabinets are unsecured without doors to control access. Windows do not have suitable locks or any type of physical hardening. Various boxes and combustible items including a wooden pallet are stored alongside or near servers representing a fire hazard. There is no fire suppression or fire fighting equipment within the server room. The use of Software X in the laboratory for the method of processing Y related data is unlicensed. There are no confidentiality agreements in place with students undertaking a year in

		industry placement with the Digital Unit and nothing in the QMS to describe the verification and confidentiality requirements required when using staff who are not employed by the Digital Unit.

Table 4: Qualitative assessment of issues raised during initial assessments of DFUs.

During the period of September 2018 to April 2019 UKAS performed 29 surveillance visits to accredited DFUs which resulted in 571 mandatory findings being raised (an average of 20 findings per assessment compared to 49 for initial assessments).

Notable trends observed during the surveillance visits were as follows.

- The number of overall findings (non-conformities) was greatly reduced compared to the number raised at the initial assessments.
- The number of findings relating to the quality management system reduced compared with the initial assessments, with a greater percentage relating to technical issues.
- The percentage of findings relating to the management of audits and non-conforming work increased. At initial assessment these systems were relatively new therefore the first surveillance assessment provides a good reflection of how the DFU is gaining its own assurance on the implementation of its systems and the handling of any quality issues when they arise.
- Findings relating to validation reduced as the methods are embedded into use. However, a number of the validation findings related to failure to review, verify and/or justify changes to the methods, such as software updates.
- There was an increase in the percentage of findings raised in relation to staff not following the documented procedures which can lead to variation of processes being undertaken in DFUs. There were no significant issues relating to poor practice witnessed.
- There was an increase in the percentage of findings raised in relation to technical records. This includes inaccurate or insufficient information being recorded in notes or supporting quality records.
- An increase in exhibit handling issues was witnessed, with examples of exhibits being accepted with inappropriate packaging and incorrect descriptions or reference numbers. In addition, exhibits were not being stored appropriately or in the locations which were recorded in the system.
- In relation to Contract Review and Reporting, findings were raised in relation to incorrect or misleading statements of accreditation status being declared in Service Level Agreements or Reports.
- With more Forensic Units gaining accreditation to the Codes, it was noted that a number of findings were related to requirements which are specific to the Codes and do not have an ISO/IEC 17025 equivalent, such as Business Continuity Planning, Staff Vetting, IT Security, and the format and structure of validation documentation.

Assessment, by its nature, concentrates on documenting non-conformances to the specified standard, so it is more difficult to identify from assessment documentation when good practice has been observed. Nonetheless, in some instances, assessors did note areas of good practice and Table 5 illustrates anonymised examples of the more serious non-conformances raised against the requirements of the standard and good practice noted.

Area in which findings raised	**Anonymised example non-conformance findings**	**Anonymised examples of good practice**
Technical procedures		a Deviation Request Form is used by staff to document deviations from a prescribed method and the justification for doing so, this is approved and countersigned by the unit manager. This process was deemed a good method to record and justify such deviations.
Technical practice		The configuration of software on each of the workstations is managed centrally with a gold build of software kept and copies deployed on all machines with the same configuration. There is good control over software versions and validated versions of software are recorded in the system with appropriate reference in the tool guides.
Technical records	SOP XX section Y.1 provides a table entry which relates to the number of occasions that non-validated software or equipment has be used which is incorrect when it says that this hasn't happened. An unvalidated write blocking device was used during a previous UKAS assessment. A number of hard disk drives are used for copying image files on and off site. There are no records of the management (wiping) of these drives both before use and after use. Validation data files and supporting evidence for the acquisition workstation 1 and 2 are not present in the folder. The records must have previously existed as printouts are present in the laboratory validation folder.	

Training and competency	Staff were not aware of UFED Phone Detective to enable them to identify what UFED can extract from a device. Person x had carried out casework testing yet they had not completed the competency test and were not yet authorised for casework unsupervised. Staff engaged in the imaging of tablet/mobile devices lack current training and evidence of sufficient continued professional development to keep them updated on current trends, methods, opportunities and threats.	It is also clear the Training and Competency has continued to be an important part for the staff working within the unit. Records reviewed show excellent traceability. Overall the training and competency records seen demonstrated that the laboratory is recording, maintaining and reviewing competence to a good standard.
Validation	During witnessed activity for imaging with EnCase staff used software in the acquisition process that has not been tested through Method Validation or approved (updated versions) on case work. The ten test handsets currently in use for validation are not representative of what is seen in the Forensic Unit.	Significant work had been conducted on the validation of processes to ensure that the equipment and method used can obtain a verifiable forensic image. The validation has used a number of reference disks generated internally as well as material from NIST which provides an independent verified source of data. Re-validations of FTK Imager had been conducted since the last visit. These were sufficiently thorough and well documented.
Ongoing quality assurance	Although the current QC process is fit for purpose, the QC is not completed by staff that are signed off as technically competent to perform the activities they are completing the QC for. For the Proficiency Test (123), the positive observations had not been communicated back to the team. An evaluation of the Proficiency Test (resulted in December 2017) has not been conducted or formally recorded and it is now September 2018.	32 cases were seen for 2018 and a selection examined. All reviewed showed a suitable level of QC with notes made where relevant. Issues identified in the QC had then been rectified with supporting records. An evaluation of the latest ILC test was completed by the QM. A professional discussion took place regarding the value of ILC and any non-conformities and learning that can be obtained through its use. The internal quality control methods reviewed at the initial assessment visit were found to be still being undertaken with good records supplying evidence of

		continued compliance. The combination of dip sampling, administrative review and data acquisition repeats has provided evidence of the on-going quality of their work.
Reporting of results	For case XYZ a non-standard method was used to acquire a forensic image of a hard disk drive where standard methods had failed. The SLA includes the laboratory schedule of accreditation but does not mention the use of non-standard methods such as a disk duplicator or forensic boot disk. In this case the customer had not approved the use of this non-standard method and it was not noted in the output to the customer.	
Exhibit handling	Items are being received and accepted with incorrect seal numbers. e.g. TEB: P0570XXXXX. For case XYZ the records state that the 4 phone items were moved from the exhibit store to the operator's workstation on xx/yy/2018. The items were processed on 11/06/2018 and then returned to the store from the workstation on aa/bb/2018. There is no tracking record for location in this three-month period or record that the items were being securely stored.	
Other	The digital forensics unit has not completed/revised the 2018 method witness audit schedule for all its units. The current internal auditing programme requires the completion of in excess of 500 audits per year. This is not a sustainable model for the size of the current auditing resource.	

Table 5: Examples of issues raised during surveillance visits to accredited DFUs

3.2 Data from Referrals to the Regulator

The first quality referral to the Regulator in relation to digital forensics was received in April 2012. Between then and the end of August 2019, 53 referrals were received. There was insufficient information given to evaluate 6 of the referrals. Of the remaining 47, 17 were self-referrals, 28 were referred by a third party and 2 were identified by the Regulator from court judgements. Figure 2 shows the number of referrals by year

and Figure 3 illustrates the source of those referrals and which were self versus third party referrals.

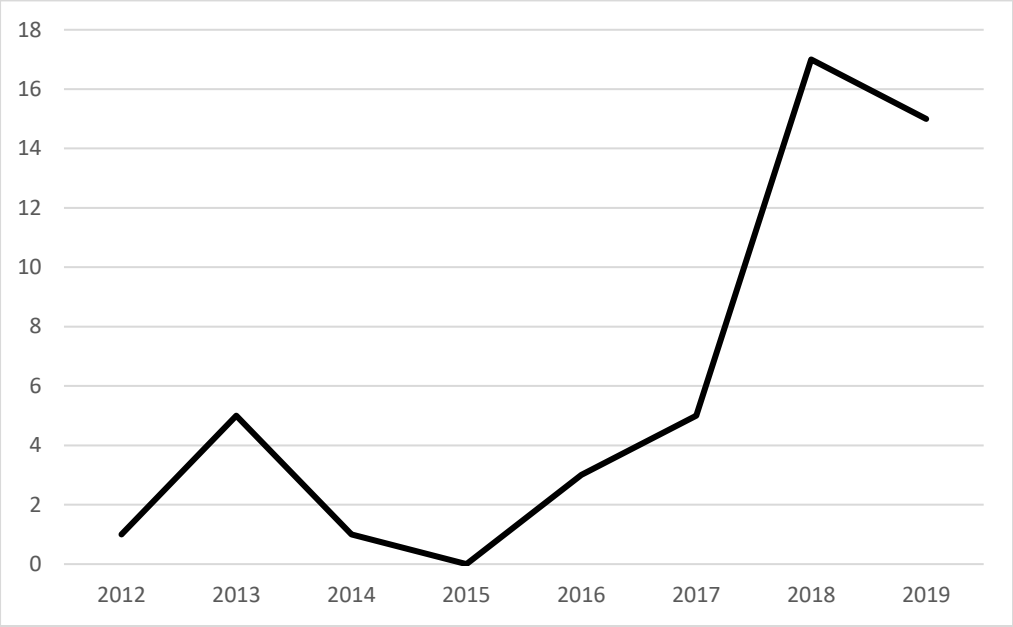

Figure 2: Number of referrals received by the Regulator from 2012 to the end of August 2019

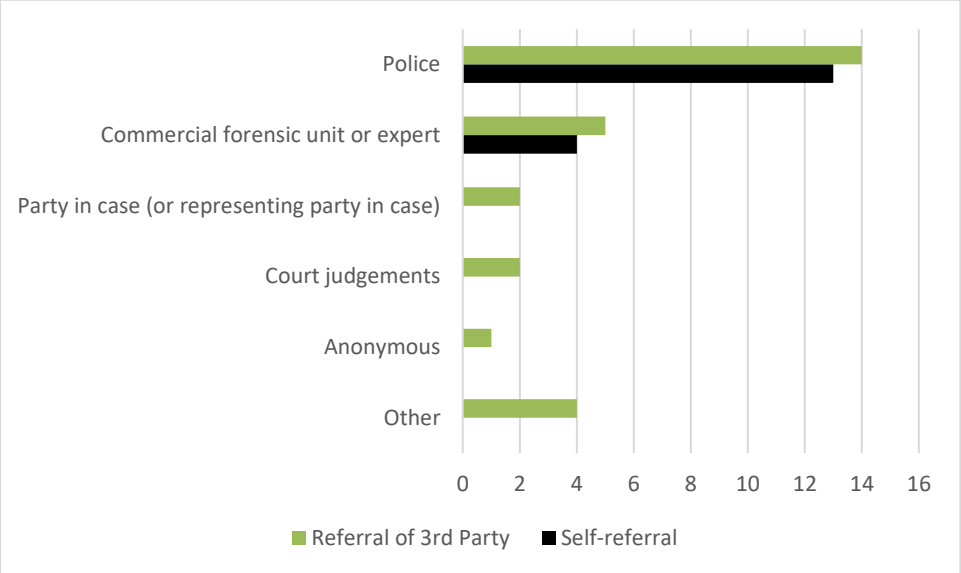

Figure 3: Source of referrals concerning digital forensics to the Regulator

The digital forensics referrals were sub-divided according to discipline. Because of the nature of some referrals (e.g. loss of data), only broad classifications could meaningfully be applied. The distribution of referrals between these classifications is shown in Table 6.

Category of Work	**Number of Referrals**
Analysis of imagery	17
Audio analysis	1
Cell site analysis	2

Other digital forensics	27

Table 6: Classification of referrals

Both of the referrals concerning cell site analysis originated from court judgements, one of which was self-referred to the Regulator by the police force concerned.

The referrals regarding analysis of imagery, with the exception of one police force which self-referred poor timeliness, were all directed at commercial forensic units or individuals, with two companies attracting 9 referrals between them, albeit several referrals related to the same issue(s). The referrals regarding imagery were overwhelmingly associated with lack of competence/expertise (14) and lack of method validity (9); some referrals concerned both.

The referrals regarding general digital forensics covered a broader range of concerns, as illustrated in Figure 4.

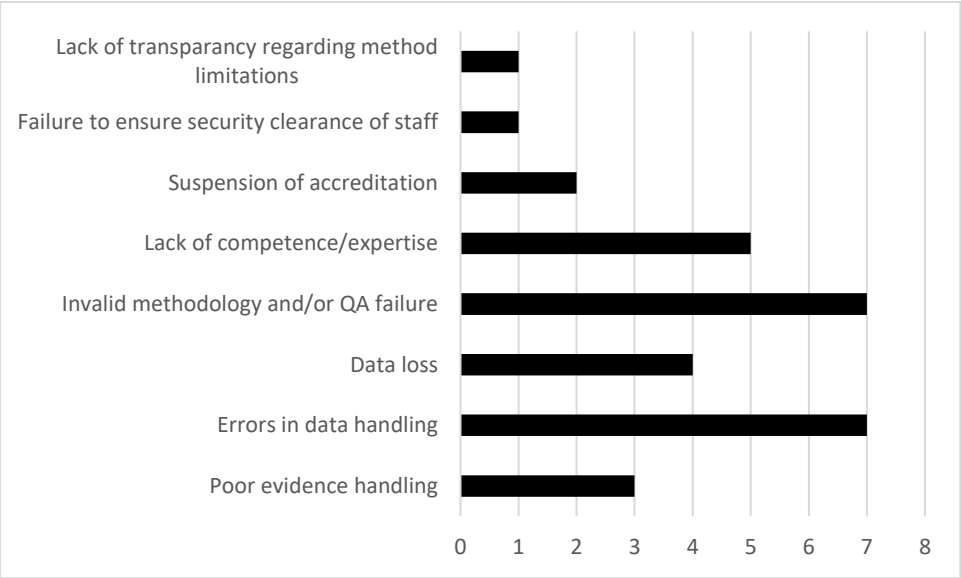

Figure 4: Areas of concern raised to the Regulator in relation to general digital forensics (i.e. excluding imagery analysis, audio analysis and cell site analysis). The number of areas of concern is greater than the number of referrals, since some referrals concerned multiple issues.

3.3 Costs of Compliance

The costs of complying with the standards and gaining third party accreditation to demonstrate compliance can be split broadly into two subsets.

a. The “internal” cost of validating methods, establishing objective evidence of competence and implementing robust quality management procedures.
b. The charges levied by the accreditation body.

We do not have access to reliable estimates for the internal costs. External costs are more easily measured but vary greatly between organisations, based on the organisation’s level of preparedness for the accreditation visit, the size and

complexity of the organisation including number of operational sites and the scope of accreditation sought.

The UK's accreditation body, UKAS, bases its charges on a fixed day rate. Taking the initial assessments described in Table 2, the range of days required was from 4.75 to 19.25, equating to charges between £4.2k and £16.5k. No meaningful average can be calculated, because of the impact of size, complexity and readiness. However, the costs to two organisations of similar size, with similar accreditation scope differed by as much as £8.7k (c.10 assessment days), due to the need for one to have additional visit(s) and extra office time to review evidence that actions raised during the assessments had been effectively closed.

4. Discussion

The findings described here show that significant areas of concern, with the potential to impact negatively on the production of expert reports and the CJS, were identified during the process of assessment for accreditation to ISO/IEC 17025. The evidence therefore supports the need for quality standards in digital forensics. A wide range of technical, administrative and management findings were raised, demonstrating that the accreditation process is addressing the provision of digital forensics services as a system and not concentrating on a single element, such as the technical examination of an item. No matter how skilled the examiner, if he or she is not supported with the requisite equipment, resources and ongoing training, the system will not function as it should. Each finding raised must be addressed, with objective evidence submitted and reviewed for adequacy before accreditation is granted. The process of gaining accreditation therefore leads to measurable improvement and is achievable, with 47 different legal entities now holding accreditation for at least one digital forensics discipline.

Quality failures in accredited organisations were referenced by Sommer (2018) as evidence that standards are not an absolute guarantee of quality. We agree that standards are not an absolute guarantee of quality; accreditation to standards gives external assurance that an organisation has the sustainable competence to produce reliable results in the accredited activity. UKAS technical assessors are drawn from the digital forensics practitioner community. They are trained and assessed as competent prior to their first unsupervised assessment, but each will have gained additional experience of assessment over their time from observing a number of assessments. Technical assessors meet regularly to share learning and minimise variability in approaches to assessment, although it must be recognised that different findings will be raised in different units because assessment is a sampling exercise, to check if an organisation is effectively managing its own quality. Accreditation cannot prevent all error, nor does the quality standard address the financial viability of a company, which was one of the examples quoted to demonstrate its alleged failure. The fact that system cannot achieve 100% success in preventing error is not a reason to discount its effectiveness in improving the quality and reliability of digital

forensics work, as has been demonstrated by the non-conformances raised and, as a consequence, the improvement actions put in place.

The question then arises whether or not ISO/IEC 17025 is the appropriate standard to apply. The core principles of the standard are that an organisation must be structured to support quality improvement, with defined responsibilities and policies, that its staff must be competent to conduct their roles (whatever those roles may be), that the methods used must be fit for the intended purpose as demonstrated through the process of validation, that there must be an appropriate environment, equipped with the required equipment, maintained and calibrated as applicable and that there must be ongoing monitoring of the quality of results prior to their issue to the customer. All are applicable to digital forensics provision at a systemic level. Standards within the ISO 27000 series have been suggested to be more applicable to digital forensics. This series of standards provides useful and detailed guidance on a range of digital forensics elements; ISO 27037[5] provides a framework for meeting some of the technical requirements of ISO/IEC 17025. However, it is not a standard that can be used for accreditation: it is a technical guidance document (Figure 1). England (2018) argues that the Regulator could, essentially, change the "should" advisory language to "shall" mandatory language and adopt this modified version of the standard for digital forensics instead of ISO/IEC 17025. Organisations may choose to use the guidance in ISO 27037 to assist with meeting some of the technical requirements of ISO/IEC 17025, but it is not within the gift of the Regulator to unilaterally change an international standard or convert a guidance document to an accreditation standard. Further, concentrating only on technical requirements does not necessarily ensure improvement in the broader quality system, such as effective review of performance and ongoing improvement. As organisations in the UK begin to adopt ISO/IEC 17020 for their digital forensic activities at crime scenes, it will be possible to monitor any differences in effectiveness or applicability between that standard and ISO/IEC 17025 along with the assessment approach to ensure methods are fit for purpose. If data exist in other jurisdictions regarding the implementation of ISO/IEC 17020 for digital forensic activities, we would encourage publication of such data, to further inform the debate.

The issue of whether or not the validation requirements in ISO/IEC 17025 are too onerous to be achievable in a fast-moving environment such as digital forensics is critical. We start from the position that understanding the strengths and limitations of methods employed in the CJS is essential, in order that investigators and courts know what may not have been found or what artefacts may be present. Marshall and Paige (2018) concluded that the absence of clear requirements statements (and corresponding lack of transparency about those requirements) leads to a break in evidence of correctness for tools and methods. They observed an absence of clear technical requirements within digital forensics service providers and a reluctance to

[5] ISO/IEC 27037:2012(E) Information Security – Security Techniques – Guidelines for identification, collection, acquisition and preservation of digital evidence.

disclose customer requirements by tool providers. They also described a lack of technical requirements in validation plans. This lack of technical requirements contravenes the validation requirements set out in the Codes; we are unable to determine whether the organisations included in the Marshall and Paige study also sought accreditation to the Codes, but it is standard UKAS practice to raise non-conformities if the validation is not in line with the requirements of the Codes. There are, however, practical improvements that could be made to assist with validation and verification.

1. We agree with Marshall and Paige (2018) that greater transparency of technical requirements would be an improvement. If users clearly specified their requirements of tools and methods, tool providers would be better able to prioritise development to meet those requirements and to test the performance of their tools against the requirements. Users of the tools will still need to validate the performance of their end to end method, which includes not only tool(s) but also evaluation of the case, selection of the optimal examination strategy and tool(s), use of those tools and subsequent quality assurance mechanisms, but sharing of specifications and of testing results, in a mature customer-supplier relationship, would bring value to all parties, reducing duplication of effort and hence costs.
2. A widely available resource of ground truth data, which is kept up to date as technology progresses and is accessible to all digital forensics providers in the Criminal Justice System, whether in policing or the private sector, would centralise a significant level of effort and expense, reducing duplication of effort. Specification of what such a resource would contain and how it would be made available is not straightforward but is worth pursuing; it has the potential to reduce the costs of achieving and maintaining the standard and decrease the time spent in validating or verifying updates and new methods. A project under the auspices of the European Network of Forensic Science Institutes (ENFSI) is underway which may assist in this regard (Luck, 2016): it seeks to gather a database of datasets and a database of tool test results. This may assist with understanding of tool performance and hence contribute to understanding of uncertainty in digital forensics methods. A project to propose suitable areas of scope for ground truth databases, assess approaches to database construction and produce recommendations has been commissioned by the Regulator's Digital Forensics Specialist Group, as a first step to establishment of a widely available resource of ground truth data; the project is due to report its recommendations in the Spring of 2020.
3. There has been much learning as organisations have undertaken validation and sharing of this learning in the community should be encouraged. A project to validate the performance of digital "kiosks" for use by front-line officers to extract data from mobile phones in a "level 1" analysis[6] has recently been

[6] Level 1 analysis has been defined within policing in England & Wales as "*Logical Capture of standard data types which a single preconfigured tool can recover (could be limited) from Handset,*

undertaken, following the procedure set out in the Codes; a subsequent paper will detail this work and the learning gained.

4. Central resources for tool testing, such as that at the National Institute for Standards and Technology (NIST), have the potential to reduce duplication of effort (accepting that organisations would still need to validate their end to end methods). Similarly, the provision of central resources for development and validation of standard methods has the potential to substantially reduce, although not eliminate, the validation burden on each organisation.

In practical terms, it is important to note that an organisation's schedule of accreditation does not specify the specific version of software used within a method: UKAS assessments include consideration of an organisation's methodology for upgrading software, assessing risk and conducting revalidation as required. Accredited organisations can thus continue to keep up to date with changes without having to await external assessment of each, providing they demonstrate the competence to do so in a controlled manner. As experience with accreditation of digital forensics activities increases, it will be important to keep under review the manner in which accreditation scope can best be defined; as the range of devices being examined increases (e.g. drones, vehicle systems, routers, smart watches, RAID arrays, Internet of Things devices and so on), a device-based approach to scope may become unmanageable. A technique-based rather than device-based approach, which seeks to identify the common methods involved in examining a broad range of novel devices, may warrant further consideration.

Recognising the multifactorial nature of risk: not responding to a need for rapid method development on one hand and uncontrolled introduction of untested methods on the other, the Codes provides a route to introduce a completely novel method rapidly, with the proviso that the customer must be fully informed prior to commissioning.

The Chartered Society of Forensic Sciences is developing a "Generic Quality Management System"[7], which will deal with non-technical policies and procedures and has the potential to reduce the abstraction time from operational work to produce effective policies which meet the demands of the standard. Central assessment of the policies has the potential to reduce the cost of each individual assessment, since it can then focus more on the implementation of the policies and on the technical procedures. The Regulator is liaising with Government to determine if such a scheme could be subsidised, to reduce the costs to participants (Government, 2019; Tully, 2019).

Tablet, (U)SIM or Memory Card, deployed at a fixed site outside a laboratory environment. (The tool having locked down data recovery methods and control as set out in the Forensic Science Regulators Codes of Practice & Conduct)." (John Beckwith, personal communication)

[7] See https://www.csofs.org/Quality-Competency, accessed 12 November 2019

Availability of and participation in high quality proficiency tests would improve the level of assurance gained during the accreditation process. Review of available schemes is beyond the scope of this paper but increasing the quality and availability of proficiency testing schemes across forensic science is likely to achieve increased focus in the coming years.

No organisation in England and Wales currently holds accreditation for image analysis or comparison. The quality problems in that discipline are the source of a separate publication (Tully and Stockdale, 2019) but method validation has historically been lacking, with conflicting views on the reliability of commonly used methods, poor understanding of uncertainty of measurement, even in measurement-based analyses such as height estimation from CCTV footage.

Similarly, the level of experience in accreditation of more complex areas of digital forensics, including interpretation is limited thus far. We advocate continued integration of digital forensics with other branches of forensic science, where development of scientific approaches to evaluation of evidence have been the subject of many years of research. Ensuring that a scientifically robust approach is adopted across digital forensics will inevitably smooth the adoption of quality standards.

The Regulator's Digital Forensics Specialist Group will continue to monitor the effectiveness of standards as their application widens and will pursue ways in which method validation and verification can be improved to provide optimal assurance whilst minimising the burden on individual digital forensic units.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Acknowledgements

We thank Eoghan Casey for helpful comments on an early draft manuscript and Simon Iveson of the Forensic Science Regulation Unit for input at all stages preparation of the paper.

References

Casey E. Cutting corners: trading Justice for cost savings. Digit Invest 2006; 3:185-186, doi:10.1016/j.diin.2006.10.006.

Casey E. Differentiating the phases of digital investigations. Digit Invest 2016; 19: A1-A3, http://dx.doi.org/10.1016/j/diin.2016.11.001.

Casey E. The Checkered Past and Risky Future of Digital Forensics, Australian Journal of Forensic Sciences 2019; 51: 649-664, doi: 10.1080/00450618.2018.1554090

Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council. Strengthening Forensic Science in the United States: A Path Forward. 2009; https://www.ncjrs.gov/pdffiles1/nij/grants/228091.pdf, accessed 12 November 2019.

Cusack B. Making sense of digital forensic standards. Digital Forensics Magazine 2019; 39: 10-14.

England, G. Written evidence to the House of Lords Inquiry into Forensic Science. 2018; http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee-lords/forensic-science/written/89891.html, accessed 12 November 2019.

Federal Bureau of Investigation / Department of Justice Microscopic Hair Comparison Analysis Review. 2015; https://www.fbi.gov/services/laboratory/scientific-analysis/fbidoj-microscopic-hair-comparison-analysis-review, accessed 12 November 2019.

Forensic Science Regulator's Codes of Practice and Conduct. 2017; 4 https://www.gov.uk/government/publications/forensic-science-providers-codes-of-practice-and-conduct-2017 accessed 12 November 2019.

Government Response to the Lords Science and Technology Select Committee Report: Forensic Science and the Criminal Justice System: A Blueprint For Change. 2019; para 14. https://www.parliament.uk/documents/lords-committees/science-technology/forensic-science/Govt-response-forensic-science.pdf, accessed 12 November 2019.

House of Lords inquiry into forensic science. 2018 – 2019; written and oral evidence available at https://www.parliament.uk/business/committees/committees-a-z/lords-select/science-and-technology-committee/inquiries/parliament-2017/forensic-science/forensic-science-publications/, accessed 12 November 2019.

Jones A and Vidalis S. Rethinking digital forensics. Ann Emerging Tech Computing 2019; 3:41-53, doi 10.33166/AETiC2019.02.005

Luck, J. Challenges and Opportunities for Statistics in Digital Forensics. 2016; https://gateway.newton.ac.uk/sites/default/files/asset/doc/1612/Luck.pdf, accessed 12 November 2019.

Marshall AM and Paige R. Requirement in digital forensics method definition; observations from a UK study. Digit Invest 2018; 27: 23-29, https://doi.org/10.1016/j.diin.2018.09.004

Page H, Horsman G, Sarna A, Foster J. A review of quality procedures in the UK forensic sciences: what can the field of digital forensics learn? Sci & Justice 2019; 59: 83-92, https://doi.org/10.1016/j.scijus.2018.09.006.

President's Council of Advisors on Science and Technology, Forensic Science in Criminal Courts: Ensuring scientific validity of feature comparison methods. 2016; https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf, accessed 12 November 2019.

Sommer P. Accrediting digital forensics: What are the choices? Digit Invest 2018; 25:116-120, https://doi.org/10.1016/j.diin.2018.04.004 1742-2876/

Tully, G. Letter to Chair of House of Lords Science and Technology Committee. 2019; https://www.parliament.uk/documents/lords-committees/science-technology/forensic-science/Forensic-science-regulator-response-forensic-science.pdf, accessed 12 November 2019.

Tully G and Stockdale M. Commentary on: Hak. Evaluation of the Forensic Science Regulator's recommendations regarding image comparison evidence. Forensic Sci Int: Synergy; 2019: 298-301 DOI: 10.1016/j.fsisyn.2019.09.006

UKAS. Further supplementary written evidence to the House of Lords Inquiry into Forensic Science. 2019; http://data.parliament.uk/writtenevidence/committeeevidence.svc/evidencedocument/science-and-technology-committee-lords/forensic-science/written/94683.html accessed 12 November 2019.

Ward T, Edmond G, Martire K and Wortley N. Forensic science, reliability and scientific validity: Advice from America. Criminal Law Review 2017; 5: 357-378. ISSN 0011-135X.