Journal of Forensic Radiology and Imaging 4 (2016) 2–6

Contents lists available at ScienceDirect

Journal of Forensic Radiology and Imaging

journal homepage: www.elsevier.com/locate/jofri

How does post-mortem imaging compare to autopsy, is this a relevant question?

Bruno Morgan [a,*], Guy N. Rutty [b]

[a] Imaging Department, University Hospitals of Leicester, Leicester Royal Infirmary, Leicester, LE1 5WW, United Kingdom
[b] East Midlands Forensic Pathology Unit, University of Leicester, Robert Kilpatrick Building, Leicester, LE2 7LX, United Kingdom

ARTICLE INFO

Article history:
Received 11 September 2015
Received in revised form
26 October 2015
Accepted 7 November 2015
Available online 10 November 2015

ABSTRACT

The rapid development of imaging techniques used for the investigation of death invites the question "how does imaging compare with autopsy". This paper is based on an invited talk given to the International Society of Forensic Radiology and Imaging in 2015 and attempts to show that this question cannot be answered, as it is in fact several questions depending on the circumstances of death. Review of the literature is fraught with difficulties due to the rapid evolution of technology, the varied circumstances of death investigation and the multiple possible post-mortem investigations that imaging can be compared with. The article focuses on the assessment of adult sudden natural death and suggests the questions that need to be answered.

1. Introduction

Medical imaging is becoming increasingly common for the investigation of death. There is great debate about the relative qualities of imaging as compared with the traditional autopsy and many conclusions have already been made. However, this is not a straightforward debate and we believe it is complicated by a lack of comparative studies, vested interests, resistance to change [1] and probable misjudgements, based on misinterpretation of available data [2].

When a new imaging technique is developed it undergoes a series of testing (Health Technology Assessment, HTA), which can be summarised by the concepts of efficacy, effectiveness and efficiency. Efficacy and effectiveness are the extent to which a health technology influences a favourable outcome and are often used interchangeably, but are crucially different. Efficacy is a more technical term – does the test actually work, whereas effectiveness concerns the extent to which this "efficacy" actually brings about the desired effect in the real world. For example a Magnetic Resonance Imaging (MRI) scan may be able to make a diagnosis accurately, but this has little impact or effectiveness if the diagnosis is untreatable or unimportant. An evaluative framework therefore must decide whether the test identifies the abnormalities, makes a diagnosis, displaces or improves upon other tests, contributes to service delivery and improves on outcomes for the whole population. Not surprisingly, most diagnostic tests have good evidence for their "efficacy" but often very little on their impact or effectiveness [3]. Focusing purely on technical performance can lead to failure to address the bigger picture and to incorrect assumptions about the new test. Therefore, before critically assessing the literature there must be a very clear concept of what overall effect or impact is required. Put simply, if we ask the wrong questions we will often get the wrong answer!

The purpose of this paper is to review what post-mortem imaging can do now and to try and frame the questions that need to be asked, both of previous studies and in new study design. However, after a review of the literature, the reader can come away with the impression that post-mortem (PM) imaging for the investigation of natural death is either very poor, giving a cause of death in less than 10% cases [4,5], or extremely good giving a cause of death in 90% [6]. Similar discrepancies are also seen in unnatural death. All these papers are peer-reviewed and satisfy scientific scrutiny but they ask different questions, which have very different answers. This is the effect of having the same technical performance but with different measures as to the required effect.

We will show how by a combination of changing context, different biases, and inconsistent study comparators and end-points that it is very difficult to compare and even apply study findings to work in different areas. We will focus mainly on the investigation of adult sudden death, but also use information from other post-mortem investigation scenarios as examples.

2. Context

There are many reasons for a PM investigation, and these reasons dictate the questions to be asked. These include whether the

* Corresponding author. Fax: +44 116 252 3274.
E-mail address: bm11@le.ac.uk (B. Morgan).

http://dx.doi.org/10.1016/j.jofri.2015.11.003
2212-4780/

death is of natural or unnatural cause, single or multiple, witnessed or un-witnessed, and also the age and identification of the dead. This opens up a vast range of possible scenarios where the key questions are very different, from medical cause of death, mechanism of death to victim identification. The purpose of the investigation will also vary depending on who has requested it. The police or the local coroner may have very specific questions that need to be asked, but investigation may also be requested for medical care reasons (the so-called hospital PM). This investigation is not only aimed to provide the principle cause of death, but also find all clinically relevant pathology to audit and educate clinical care.

It is also crucial to consider whether the medical imaging is being used as an adjunct to the standard PM investigation or a replacement. Use as an adjunct places less pressure on a test, as false negatives and false positives can be identified by and challenged by the traditional tests. Both tests together will normally have greater sensitivity [7,8], and will therefore become the new gold standard. There is also growing interest in using medical imaging as a replacement for standard autopsy, both for natural and unnatural death. This is particularly true in England and Wales, due to a relatively high autopsy rate of 20% of natural deaths [9]. Although other countries do not have such high autopsy rates, there is also interest in providing PM imaging instead of autopsy internationally to increase the number of investigations (e.g. in Japan) and also to provide audit of medical care and education [10].

Another question is, "what competitor is imaging actually replacing"? This really should only be internal aspect of the traditional autopsy. Imaging cannot compete with the external examination of a body [11], and although this can be performed by imaging [12] it rarely needs to outside specific circumstances. Furthermore toxicology and even histology via targeted biopsy can be routinely added to an imaging examination [13,14].

It is clear that when the context of the study changes then the prevalence of different diseases will also change. This pre-test probability will have a dramatic effect on the assessment of accuracy using the test. Many studies have different recruitment. For example, it is difficult to compare the results of excellent studies of hospital ITU deaths [7] compared with those from the community [15]. Likewise large studies and systematic reviews may have a very heterogeneous population, including both natural and unnatural deaths [4,16]. These statistical considerations are discussed in the next section.

3. Bias in research methods and findings

An on-line search reveals a multitude of potential biases that can affect the outcome of clinical trials. Autopsy and PM imaging studies are no different, with most studies showing "… marked variability in the methodology and the reporting quality, thus hindering … the generalisability of the results [17]". Significant biases in research data are not always obvious and may be subtle. It is for this reason that a paper entitled "Why Most Published Research Findings Are False" has been heavily cited in the scientific literature and press [18].

It is natural at this point to call for more stringent peer-review and vetting prior to publication of scientific studies, but many scientists and editors no longer believe this to be either possible or appropriate. However well meaning, every form of review leads to a form of scientific censorship, and "claimed research findings may often be simply accurate measures of the prevailing bias [18]". Peer-review discourages publication of negative studies and also novel, small and eccentric studies. It often fails to identify falsehood and significant errors in experimental design and statistical analysis [19,20]. Furthermore, the electronic age is bringing greater capacity, with reducing costs for online only journals. Getting manuscripts accepted is therefore becoming easier and the ease of electronic searching and retrieval means that the brand-name of the journal is no longer crucial to disseminating information. This means that regardless of our view of peer-review, the reader should mistrust it and rely on our own "post-publication" peer-review, not just from "letters to the editor" but on a comprehensive discussion of the "market of ideas" [20]. Ideally this would be supported by high quality systematic reviews, but these are only possible when there is enough data on a single circumstance of death with similar questions to study, for example traumatic death [21].

It is now therefore up to the individual to decide whether a study is scientifically accurate and also relevant to their own practice. A few examples are chosen based on the scientific content of the annual meeting of the International Society of Forensic Radiology and Imaging (ISFRI) 2015. None of the examples implies that the research method or the findings are wrong, simply that the reader should be vigilant.

3.1. Conference bias

There is no doubt that conferences benefit from submitted articles. Unlike many journals, the conference will often directly financially benefit from accepting a manuscript, by increased attendance. We served as the scientific review panel for (ISFRI) 2015, and can vouch that all papers submitted were peer-reviewed. However, this was not a censorship process and nor was it strict quality control. It was a check to make sure submissions were intelligible, suitable for the meeting and for what session they were allocated to. To ascribe acceptance of a poster for the meeting as approval from the ISFRI itself would be incorrect.

Beware the invited review: when an expert is invited to review their field the peer-review process may not be particularly vigorous. To make sense of rapidly expanding knowledge most of us rely on these expert reviews to try to put some perspective on a subject and too much scrutiny will simply discourage experts from doing it. This does not imply that the review is false; simply that the experts' opinions and knowledge is often taken on trust. However, scientific reviews can propagate previous weak or even false findings so that they become part of perceived wisdom. Trying to find the origin of questionable dogma can sometimes be difficult and has been coined a "Woozle hunt" after A.A. Milne's book, Winnie the Pooh, where Pooh Bear and Piglet follow a Woozle's tracks, finding more and more tracks, until they realise they are only following themselves! Even when the "Woozle" is found, the problem often does not lie with the original story, but with the retelling. For example, we still hear people telling us that PM computed tomography (PMCT) is not very good for fractures, particularly those of the ribs and flat bones of the skull. We do not believe this as pathologists have to work hard at autopsy to even come close to detecting the amount of fractures seen on CT. In fact this is a good example of a simple question actually being quite difficult to answer, relying heavily on context, technique and experience.

3.2. Technical bias

If clinical CT scan reports from older CT scanners are compared with full dissection at autopsy, fractures are missed, particularly in the ribs and skull [22,23]. However, this does not necessarily compare to modern PMCT practice, which uses higher radiation-dose scans, thinner cuts and a different reporting approach [24]. For example, rib fractures normally hurt, and therefore "diagnose themselves" in life. These are therefore not routinely screened for