**HEALTH ECONOMICS,
POLICY and LAW**

CrossMark

ARTICLE

Clinical negligence cases in the English NHS: uncertainty in evidence as a driver of settlement costs and societal outcomes

Alexander W. Carter[1]* iD, Elias Mossialos[1,2] iD, Julian Redhead[3] and Vassilios Papalois[3,4]

[1]Department of Health Policy, London School of Economics & Political Science, London, UK, [2]Institute of Global Health Innovation, Imperial College London, London, UK, [3]Imperial College Healthcare NHS Trust, London, UK and [4]Department of Surgery & Cancer, Faculty of Medicine, Imperial College London, London, UK
*Corresponding author. Email: a.w.carter@lse.ac.uk

Abstract

The cost of clinical negligence claims continues to rise, despite efforts to reduce this now ageing burden to the National Health Service (NHS) in England. From a welfarist perspective, reforms are needed to reduce avoidable harm to patients and to settle claims fairly for both claimants and society. Uncertainty in the estimation of quanta of damages, better known as financial settlements, is an important yet poorly characterised driver of societal outcomes. This reflects wider limitations to evidence informing clinical negligence policy, which has been discussed in recent literature. There is an acute need for practicable, evidence-based solutions that address clinical negligence issues, and these should complement long-standing efforts to improve patient safety. Using 15 claim cases from one NHS Trust between 2004 and 2016, the quality of evidence informing claims was appraised using methods from evidence-based medicine. Most of the evidence informing clinical negligence claims was found to be the lowest quality possible (expert opinion). The extent to which the quality of evidence represents a normative deviance from scientific standards is discussed. To address concerns about the level of uncertainty involved in deriving quanta, we provide five recommendations for medico-legal stakeholders that are designed to reduce avoidable bias and correct potential market failures.

Key words: Jurisprudence; Medical error; Patient safety; Medical negligence

1. Introduction

High quality, evidence-based medicine (EBM) is central to health system strengthening and an important part of this agenda in high-income settings is clinical negligence reform. Expenditure on clinical negligence is a growing portion of the total National Health Service (NHS) spending (NHS Resolution, 2019), reflecting the mix of changing consumer behaviours and growth in medico-legal activity. This is occurring at a time when health and social care must contend with rising demand and consequent funding challenges (Licchetta and Stelmach, 2016). Clinical negligence research spans three decades (Kravitz et al., 1991), originating from growth in medical litigation in the 1970s in the United States (USA). To address growth in claims and find solutions, economists began characterising these legal interactions in the 1980s (Danzon, 1983). Whilst much research attention has been given to medico-legal issues in the USA, practical solutions that reduce the growing burden of claims against care providers are lacking in the English system (National Audit Office, 2017). Specifically, there is a lack of detailed observational studies that isolate and address the most influential factors of claim outcomes.

New, systematically and scientifically derived solutions are needed to what has now become an ageing problem for the NHS in England (Fenn *et al.*, 2002).

Recent research on the economics of medical negligence tends to evaluate the economic burden of harm, by specialism (Markides and Newman, 2013; White *et al.*, 2015) rather than developing evidence on the (cost-) effectiveness of policy or process reforms. Evidence on the latter tends to be published in the USA and the literature remains under-developed, given the breadth of market failures that need to be addressed (Kachalia *et al.*, 2010; Mello *et al.* (2017); Pettker *et al.*, 2014). Therefore, novel evaluations are needed to ameliorate the growing financial risk to the English NHS, balancing the need for socially optimum outcomes, which is where harm is compensated at an appropriate rate for both society (encompassing health systems) and individual claimants. The lack of solutions to a now stabilising annual claim rate, but growing average settlement costs (NHS Resolution, 2019), shows that new approaches are needed to address the underlying drivers. Insufficient attention has been given to sources of bias and inefficiency in the legal processes governing outcomes, with reforms instead focusing on high-level changes to legal costs (Rimmer, 2018), streamlining national processes (National Audit Office, 2017), building larger insurance pools (NHS Resolution, 2019), banning legal firms from advertising their services in the NHS (NHS England, 2018), and reducing the personal injury discount rate to minus 0.25% (Ministry of Justice, 2019). The effects of these measures require formal evaluations, once they have matured; however, continued growth in negligence costs suggest that solutions aimed at reducing these costs are needed. In the case of the personal injury discount rate, which is the rate at which compensation for loss of future earnings is reduced, the direct effect is that settlement costs increase, which may partly explain why the average cost per settlement in England has increased (NHS Resolution, 2019).

This article presents an exploratory study of medical negligence claims that aimed to appraise the quality of evidence informing outcomes. We posited that evidence quality in these cases is a fundamental driver of claim outcomes and that the most appropriate means of appraising that evidence was by adopting the same methodological standards as used in health care decision making. This research was motivated by the lack of evidence on an issue that has significant implications for medical negligence systems around the world.

2. Identifying underlying drivers of medico-legal outcomes

Under tort law, which provides the legal underpinnings for negligence claims in England, a claimant (patients or their family) must evidence four things to successfully sue the NHS (McBride and Bagshaw, 2012):

(1) that the provider had a duty of care to the claimant,
(2) that a duty of care was breached by the provider,
(3) the breach caused the claimant some form of loss or harm, and
(4) at least one of the losses caused by the provider's breach is actionable.

Probabilistic judgements are made at each of these four stages, whereby clinical and legal events align to produce an outcome for society. If liability is established, financial settlements attempt to correct the market failure, using a quantum calculation that captures special (economic) and general (non-economic) damages. In a welfarist view, social welfare is restored by the quantum of damages. 'Perfect' quanta reflect a monetary settlement where the social marginal benefit (SMB) equals the social marginal cost (SMC) caused by the breach of care. Given the evolving landscape of both clinical harm and its governance, research-driven solutions are needed to continually optimise medico-legal processes. On average, financial compensation must restore the long-run social welfare equilibrium by, for example, incentivising a provider to improve its care using explicit punitive damages. If the size of the difference in the differences between

claimant marginal costs (CMC) of harm and claimant marginal benefits (CMB), and external marginal costs (EMC) and external marginal benefits (EMB) is unknown, then the socially efficient position cannot be estimated. This is a fundamental research and policy challenge in clinical negligence that is hindered by inconsistent data across the NHS, thereby slowing our understanding of the root causes of negligence (National Audit Office, 2017). This extends to the problem of developing economic indicators to estimate the difference between SMC and SMB from claims.

In addition to research and measurement problems, tort law does not require the production of scientifically rigorous estimates of causality to derive quanta. There are several reasons for this. Torts are adjudicated under Common law, whereby judges decide causation and settlements within the loose interpretation of 'greater than 50 per cent probability' that a breach of care led to private losses (McBride and Bagshaw, 2012). This approach is desirable if it achieves social optimums, allowing for errors due to chance, but if on average the social optimum position is outside an agreed margin of error, then welfare losses to external parties may fall disproportionately on one side, amounting to system failure. To isolate the cause(s) of these failures, researchers need to investigate a complicated series of interactions between healthcare providers, payers, patients and legal firms, using both top-down and bottom-up methods.

Legal documentation from each claim case contains detailed information on claims of harm, plus experts' statements on causality, the Courts' judgements, legal costs, and the size of final settlements; these are a bottom-up data source. Claim cases provide a means of exploring the drivers of quanta calculations and they have potential to reveal predictors of claim outcomes. In isolation, cases do not explain whether settlements are socially efficient or not. In the least, there may be signals as to the level of uncertainty in liability/causation judgements and, therefore, the likelihood of system failures. In claim cases, expert witnesses advise the Courts' and if expert reports are biased then there is potential for uncertainty, which may drive social inefficiency. In science, bias can be avoidable (systematic bias) (Mahtani *et al.*, 2018) or unavoidable (epistemological bias caused by limitations in the way we measure phenomena). We discuss what constitutes avoidable or unavoidable bias as part of a conceptual framework for this paper, which depicts social outcomes as a product of several sources of bias in claims (see below).

To extend the theoretical and empirical background of clinical negligence, this paper aimed to estimate the quality of evidence (the primary outcome) informing financial settlements in claims against NHS provider organisations. Clinical evidence in this context is generated by expert witnesses whose reasonable judgement draws from evidence-based guidelines, professional norms, or direct examination of the claimant (Deitschel *et al.*, 2002). We did not test the effect of an intervention. The research is a direct response to reports that call for collaboration between healthcare providers and payers in England to address the threat that clinical negligence presents to the sustainability of the NHS (NHS Resolution, 2019). For this reason, a practicable and scalable approach was developed by aligning our methods to clinical competencies and standards.

3. Conceptual framework

We were unable to find explicit methodological guidance in the literature for data extraction from claim cases. Therefore, a conceptual framework was designed to inform data collection. This framework was designed using reported taxonomies of common market failures in healthcare (at the clinical level) (Gandjour, 2016) and sources of bias in biomedical evidence (expert-witness level) (Evidence Based Medicine Working Group, 1992; Montori and Guyatt, 2008; Higgins *et al.*, 2011) that are predictors of social outcomes in medical negligence claims. Theoretical outcomes that relate the predictors based on conditional statements were derived e.g. if low-quality evidence is used to inform a judgement about the negligence of a provider or providers, then uncertainty about the severity of claimant harm is higher, which raises the likelihood of adverse quanta calculations and a suboptimal long-run social outcome.
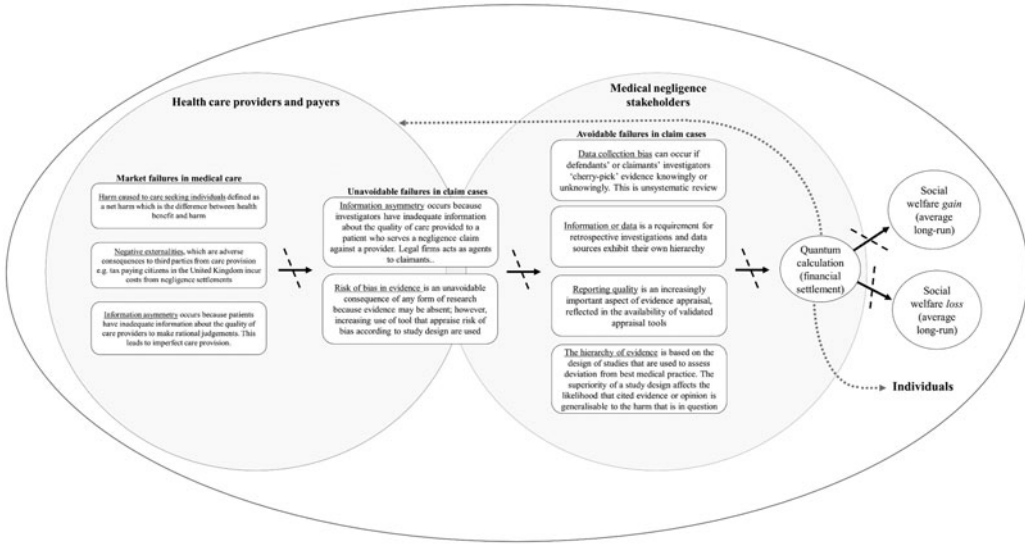
Figure 1. A conceptual framework for analysis of medical negligence markets, from a welfarist perspective. Society encompasses health care providers (individuals and organisations) and medical negligence stakeholders (collectively the 'market'). Market failures in medical care occur, manifesting as harm to care seeking individuals (left side of the graphic), causing negative externalities in the form of adverse consequences to third parties. Harm occurs due to information asymmetries between patients and providers and in the sphere of medical negligence, patients' understanding of whether harm occurred or not is another information asymmetry that predicts the pursuance of a legal claim against a provider. Two sets of potential unavoidable and avoidable failures in evidence that is used to inform quanta calculations are presented. These amount to risk of bias in both the short-term calculation of damages and long-run social welfare outcomes, some of which can be corrected.

In the framework, society encompasses health care providers (individuals and organisations) and medical negligence stakeholders (collectively the 'system'). Market failures in general medical care occur, manifesting as harm to care seeking individuals (left side of Figure 1), that also cause negative externalities in the form of acute adverse consequences to third parties. Harm occurs for several reasons, a central cause is information asymmetries between patients and providers, as described by Gandjour (2016). In general, these market failures are addressed by patient safety policy, which is a global health policy priority (WHO, 2020). Downstream, in the sphere of medico-legal stakeholders that includes the Courts, legal firms and expert witnesses, patients' understanding of whether harm occurred or not is another information asymmetry that predicts the pursuance of a legal claim against a provider. As an example, this was shown by rising negligence claims following legislative changes to the advertisement of legal services in the United Kingdom (UK) (Birks *et al.*, 2018). Claim cases are then subject to unavoidable market failures in the form of information asymmetry between legal agents (expert witnesses acting on behalf of defendant or claimant legal teams) and providers or patients. Similarly, the information acquired is unavoidably at risk of bias due to unobserved or unobservable predictors of the outcome in a claim case. This manifests as a measurement error in the legal process.

Firmly within the medical negligence market, there is a set of avoidable market failures in claim cases. Avoidable failures include unsystematic (rather than systematic) reviews of available evidence – we summarise these as a set of four potential predictors: first, unsystematically collected evidence that is used to inform causation risks biasing the interpretation of whether a defendant is negligent or not. This can mean that harmed patients do not receive damages or wrongly receive damages and / or adversely calculated settlements. The next predictor relates to the nature of the retrospective data used to construct a story of the care given to a claimant.

Irrespective of the availability of information, in pursuit of objective fact some data sources are at lesser risk of bias than others. For example, witness statements are subject to recall bias and may be inferior to electronic medical records, which provide objective records of the care administered to patients. The next avoidable failure in claim cases relates to the hierarchy of evidence. The scientific community increasingly emphasises the role of reporting quality as a means of appraising scientific claims. The absence of ingredients that are necessary for peers to critically appraise a scientific judgement risks reporting bias, creating additional measurement error in claim cases. Of the evidence that is reported, studies exhibit a hierarchy that relates to the likelihood of bias and, consequently, the certainty that an observation is epistemologically true. Systematic reviews of homogenous randomised trials are at the top of the hierarchy and the likelihood of bias and error increases with evidence below that, culminating in expert opinion. The extent to which market failures are avoidable changes with methodological advancements and growth in data availability. Put another way, we must bear in mind that what we define as harm or negligence at any one time can change due to new technologies, processes or policies that move the target (Vincent and Amalberti, 2015).

As convened in our framework, based on greater than 50% probability that a provider's action was negligent and caused actionable harm, these are the elements of clinical negligence investigations that inform quanta of damages. If a court decides that the probability of negligence is less than 50% then the quantum is zero. Conversely, if a court decides that this probability is greater than 50% then the claimant receives compensation. The impact on society of the settlements chosen may confer a long-run average gain or loss; however, this has not been systematically measured and a central data repository for estimating this impact does not exist. The conceptual framework facilitates our hypothesis that avoidable failures in claim cases, characterised as normative deviance from scientific standards, leads to suboptimal societal outcomes i.e. a net loss to society.

4. Methods

A sample of 15 high-value, non-consecutive clinical negligence claims were drawn from a repository of paper-based claims that settled through litigation for at least £100,000 between 2004 and 2016 in one large NHS Trust. This period was chosen because it captures the period of growth in clinical negligence costs to the NHS (Financial Times, 2017). A non-consecutive sample was drawn to capture several specialties and enhance generalisability of findings. The inclusion criteria were complete(d) claim cases expected to settle or resolved for at least £100,000, because these represent high-priority incidents for patients, providers and payers. A threshold of £100,000 was selected because these settlements represent the costliest proportion of all claims in the Trust. The representativeness of our case series was estimated visually using NHS Resolution's most recent summary of the proportion, by specialty, of all cases that it handled in 2018/19 (NHS Resolution, 2019); supplement one provides this comparison. A data frame was derived to analyse the 15 cases selected. This data frame is an original contribution to medical negligence research because, in England, a centralised database containing the variables collected to perform this study does not exist. As such, it is rare for researchers to have access to end-to-end claims data, which is partly due to the length of these legal cases (often several years). As such, these samples take many more years to accumulate in a single organisation, which explains why our sample of 15 cases is of high scientific value. The clinical exposures were identified in the analysis and the nature of harm caused was defined by two reviewers using the typology of errors provided by the Agency for Healthcare Research and Quality (2019) and James (2013). Episodes of care may be exposed to greater uncertainty if recording errors are more prevalent in longer or shorter durations of care, therefore a sequence analysis was performed to count and linearize the key events in each clinical case. The care pathways constructed from these notes are valuable to NHS Trusts because they also offer the most complete picture for managing risk in patient care,

when curated electronically. Similarly, claim cases include indicators that are either unrecorded or subject to coding errors in routine administrative data, particularly in smaller hospitals (Holt *et al.*, 2012), which means they are an alternative, reliable data source for resource constrained providers. This analysis allowed cases to be categorised by clinical specialty, the root cause and type of harm and final health state of the claimant (final health states are not reported due to their sensitivity).

Detailed patient-level information is not reported and approval for provider-level information was granted by our NHS organisation. The following clinically-oriented variables were collected by the lead author: the year that the claim was made, the primary specialty against whom the claim was made, the number of provider organisations involved, the number of days between admission and discharge during which alleged negligent action occurred (expressed categorically rather than continuously), the number of events in the causal chain of care that led to harm, and the type of harm caused. The number of expert witness statements, and the number of bibliographic references informing claim cases was documented (Table 1).

Expert witness evidence was measured as a combination of the number of citations reported and a code indicating the study design(s) of the article(s) cited. Citations refer to published or unpublished evidence from bibliographic sources that inform the arguments made by the expert. The quality of evidence was calculated as a total score for each case, which was derived by scoring cited articles on a scale from one to ten (highest quality) and adding these scores together. The scale was chosen because it aligns with the ten levels of evidence given by the Oxford Centre for Evidence-Based Medicine's (OCEBM) levels of evidence (Ball *et al.*, 2009), providing an informal ordinal scale for measuring risk of bias. A score of one was given to 'Level 5' evidence and ten to 'Level 1a' evidence. These scores were linked to each negligence claim and, in the absence of any citations, expert reports were themselves given a score of one for each report attached to a claim. The expert witness statements were scored in this way as they are a level of evidence informing the case. A numerical form of analysis was chosen because a qualitative analysis of these cases, which would be of value only if specific evidence was described with respect to specific elements of a case, posed a risk to patient confidentiality.

The primary outcome was evidence quality, which was measured by the lead author who is trained in evidence-assessment and was blind to the amount that claims settled or were expected to settle for. Level 1a evidence is assigned to systematic reviews with homogeneity of prospective cohort studies and Level 5 evidence to expert opinion without explicit critical appraisal, or informed by physiology, bench research or 'first principles' in medicine. This outcome was assessed using the evidence sources informing the final judgement. Two reviewers (AC and VP) subsequently resolved disagreement in the Level assigned, by discussion. An OCEBM grade of recommendation for the whole sample was determined by the two reviewers. This was decided by ascertaining the most prevalent level of evidence cited or presented in the cases.

5. Analysis

The primary outcome was summarised as a quality of evidence score, calculated as the sum of scores by claim and in aggregate ($n = 15$). The low sample size necessarily restricted our analysis to descriptive forms of interpretation. Descriptive statistics were used to present the cases and care durations were summarised to the approximate year or month. An exploratory multivariate analysis was used to examine associations between the primary outcome (evidence quality score), the duration of care, and the number of events identified in the causal chain of care that led to a claim (expressed as continuous variables), using Spearman's (non-parametric) correlation coefficient. This exploratory analysis was performed to explore whether temporal aspects of claims were associated with evidence quality, i.e. whether longer care durations give more high-quality explanatory evidence. Spearman's correlation coefficient was also calculated for the number of

Table 1. Summary of analysis of 15 cases, including the main liable specialty, number of provider organisations involved in the care given, the duration of care given (approximated by year), the number of clinical events associated with harm, the description of harm entered by the claimant, the category of harm using definition from the Agency for Health Research and Quality (2019) and James (2013), number of expert reports provided in each case, the total number of citations presented for each case (within the expert reports provided) and the quality of evidence scores derived by trained reviewers

Main specialty	Provider organisations involved	Care duration (year category)	Number of events identified	Category of harm	Type of error	Number of expert reports	Total citations presented[1]	Quality of evidence score
Oncology	2	>5	5	Missed or delayed diagnosis	Omission	0	0	0
Obstetrics & gynaecology	8	>5	39	Surgical error	Commission	4	0	4
Orthopaedics	2	<2	16	Surgical error	Communication	3	5	13
Cardiology	3	2–5	32	Device error	Commission	1	0	1
Ear, nose and throat	3	2–5	9	Surgical error, consent error	Context	1	0	1
Neurosurgery	4	2–5	21	Surgical error, consent error	Commission & Communication	1	0	1
Fertility	3	<1	12	Surgical error	Commission	2	0	2
Ophthalmology	3	<2	8	Missed or delayed diagnosis	Omission	1	0	1
Ophthalmology	2	2–5	13	Infrastructure failure (power)	Commission	4	10	20
Neonatal	3	<1	10	Missed or delayed diagnosis	Communication & Diagnostic error	5	15	41
Acute medicine	4	<1	12	Missed or delayed diagnosis	Diagnostic error	2	17	22
Obstetrics	2	<2	19	Surgical error, drug error	Commission	1	0	1
Gynaecology	3	2–5	14	Medical error, missed or delayed diagnosis, consent error	Communication & Diagnostic error	2	0	2
Orthopaedics	2	<2	12	Failure to diagnose	Omission	1	0	1
Ophthalmology	3	<1	11	Missed or delayed diagnosis	Omission	0	0	0
						28	47	

Quality of evidence scores were calculated by adding the individual scores for each piece of evidence presented in a case. Individual scores range from 1 to 10, with 10 being the highest quality evidence in the OCEBM scheme. The quanta of damages are not given with each case due to their sensitivity; however, all cases settled for more than £100,000. Care duration categories (approximate years) are as follows: <1 denotes less than 1 year; <2 denotes less than 2 years and more than 1 year; 2–5 denotes between 2 and 5 years; >5 denotes more than 5 years. Sum totals are displayed below the relevant columns for 'Number of expert reports' and 'Total citations presented'.

citations given in each case and the evidence quality (score). This was performed to explore whether citing more relevant evidence is associated with better evidence quality in medical negligence cases. If so, this would support the case for better reporting standards in expert witness statements. All analyses were performed using MS Excel®.

For ethical reasons that are linked to patient confidentiality, a detailed qualitative analysis of the cases was not performed.

6. Ethical approval

This research did not involve human subjects, therefore ethical approval was not sought; however, institutional approval was given. We followed the General Medical Council's guidance on handling patient information (GMC, 2017) and used peer-reviewed guidance (El Emam, Rodgers and Malin, 2015) to minimise the risk of patient identification during the research and in the reporting of our findings.

7. Results

The case series represented a full follow-up period in that all negligence investigations had been completed at the time of analysis. Of the 47 unique citations presented in four of the 15 cases, 29 were Level 5 evidence (expert opinion without explicit critical appraisal, or based on physiology, bench research or 'first principles'), nine were Level 4 evidence (case-series (and poor quality cohort and case−control studies)), five were Level 2c evidence ('outcomes' research or ecological studies), four were Level 2b evidence (individual cohort study (including low-quality RCT)), and zero were Level 1a evidence (systematic review (with homogeneity) of randomised control trials). Of the Level 2c evidence identified, three citations were published by professional bodies or the National Institute for Health and Care Excellence (NICE).

With inclusion of the 31 expert reports to the assessment, which were considered individual expert opinion pieces for the analysis, 78 unique sources of evidence could be appraised. Sixty were Level 5 evidence, nine were Level 4 evidence, five were Level 2c evidence, four were Level 2b evidence, and zero were Level 1a evidence.

Two cases did not contain expert reports or citations (Table 1). The distribution of evidence by OCEBM level, with and without inclusion of expert reports as Level 5 evidence, is displayed in Figure 2. The overall grade of the evidence assigned by the reviewers was D.

8. Secondary outcomes

Thirty-six years of care and 233 events in the care pathway were identified from the 15 cases. The median care duration recorded in the investigations was five years (range: less than 1 month to greater than 8 years) and the median number of discrete events in the care pathway identified was 12 (range: 8−39). The median number of provider organisations involved in the cases was 3 (range 2−8) and most errors were primarily those of commission (six out of 15 cases). Errors of omission accounted for four out of 15 cases; errors of communication accounted for three out of 15 cases; errors of context and diagnostic errors accounted for one out of 15 cases, respectively. Diagnostic error was a secondary error in four out of 15 cases and error of communication was a secondary error in one case.

One strongly positive association between the number of citations presented and the quality of evidence in claim investigations was identified ($R^2 = 0.86$). Negligible associations between evidence quality and the duration of care ($R^2 = −0.07$) and the number of significant events identified in the causal chain of care leading to harm ($R^2 = −0.2$) were found.
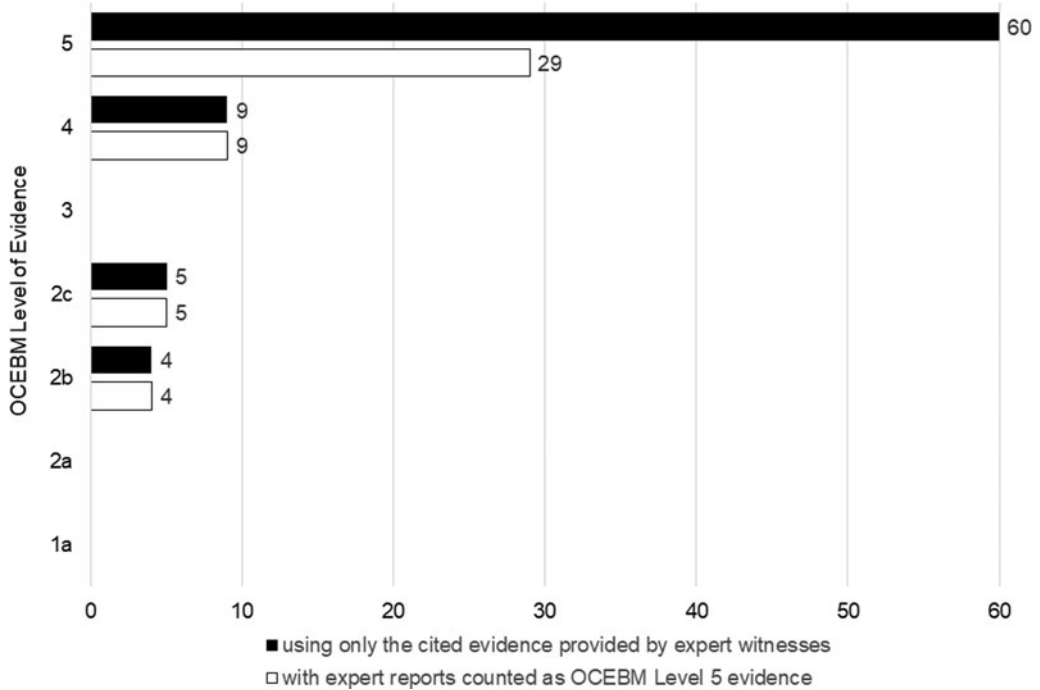
Figure 2. Frequency of OCEBM levels of evidence from 15 medical negligence claim investigations at one large NHS Trust, measured with expert reports counted as Level 5 evidence informing claim cases (dark blue bars) and measured using only cited evidence from expert reports. For the former, 77% of the evidence informing claim cases was of the lowest quality (Level 5). By not treating expert reports as Level 5 evidence, our finding presents more favourably, with 62% of evidence informing claim cases of the lowest quality.

9. Discussion

Claim cases are the record of evidence used to calculate quanta in negligence investigations and, as is common to all scientific reports, they must be peer-reviewed using appropriate standards. The finding that most evidence is Level 5 suggests a high level of uncertainty associated with the formulation of quanta in clinical settlements. Without an extensive clinical systematic review of each of the 15 harms, this finding does not indicate that expert witnesses are themselves the cause of inefficiency. Additionally, our sample did not contain information about depositions given in court, which may have challenged the evidence presented by expert witnesses. This means that the uncertainty in evidence may have been considered in quanta calculations without us accounting for it; however, this lack of transparency about how evidence was treated in court can also be considered a further failure in establishing normative scientific standards in clinical negligence. The other main limitation of our study is the relatively low sample size, which is symptomatic of the duration of time that high value negligence cases take to resolve, as shown by the years across which we derived our sample. Additionally, there is no central repository of cases that collates the variables used in our analysis, which explains in practical terms why the findings from these 15 cases are valuable. In relative terms, given these unavoidable constraints and the novelty of the analysis, which used 'end-to-end' data from closed cases, we consider the sample size to be justified. However, future research is needed to expand the effective sample size to other settings. These are the key limitations of our study, which we were unable to address fully because of the resource-intensive nature of extracting data from these unstructured claim cases. Despite these limitations, in the absence of evidence from a similarly detailed source, we demonstrated the potential for social welfare decisions to suffer complex information

failures across several levels; these are failures that social science can address. Similarly, health policy that subdues these market failures should improve healthcare providers' and payers' ability, partly through motivational incentives, to become learning systems; this is an important aspect of patient safety and medical negligence policy (Yau *et al.*, 2020). Another limitation is that our findings are not generalisable to most claims because the mean settlement cost in the UK is approximately £50,000 (NHS Resolution, 2019), which is lower than the threshold used in this study. This may limit generalisability, however, the extent to which avoidable market failures in claim cases occurs in lower value claims is unknown and is another area for investigation. We prioritised the analysis of high-value claims in one NHS institution because these are significant cost drivers in this setting. Additionally, our exposition on clinical negligence and social welfare is simplified as it deals with a tax-based health system; however, similar outcomes, albeit through differing mechanisms, can be expected in social insurance and private insurance payer systems.

We attempted to avoid measurement bias in our own appraisal of evidence by using validated tools from EBM, which reduces but does not eliminate this limitation. Importantly, as these methods are relatively easy to reproduce, scaling the investigation to other NHS providers is a feasible policy response that would increase the sample size and increase generalisability. If our finding is strongly generalisable, then the existence of a negligent outcome may be due to a (unavoidable) lack of high-quality evidence informing evidence-based practice in medicine, rather than a lack of (avoidable) high-quality reporting of evidence by expert witnesses. To determine what is avoidable or not and the extent to which a claim outcome is bounded by high uncertainty, evidence appraisals are needed. In the least, this requires consistent reported standards in expert witness reports.

Due to varying forms of clinical harm, the differing nature of claims made by claimants and a lack of standards for expert witness reports, a considerable amount of resource is needed to extract data, as compared to that for routinely collected datasets. Our analysis gives cause for significant concern about both scientific and reporting quality of evidence underpinning clinical negligence in the UK. However, the strong association between the number of expert witness citations and evidence quality ($R^2 = 0.86$) suggests that reporting standards could be improved to raise evidence standards, either through better reporting (transparency) and / or by ensuring more systematic methods are used to ensure relevant evidence is retrieved. Deficiencies in both domains are likely to manifest as inefficiencies, which are readily observable in the annual financial reporting for claims in England. Of equal concern, in the long-run, suboptimal settlements could create perverse incentives for the NHS to reduce avoidable harm. Therefore, suboptimal standards in claim cases are at odds with the NHS's ambition to be the 'safest healthcare system in the world' (NHS Improvement, 2017) and these issues must be addressed by policy makers and researchers.

These issues are partly driven by tort law. Tort law does not require the production of scientifically rigorous estimates of causality and quanta; currently, the Bolam Rule (Bolam v Friern Hospital Management Committee, 1957) is used to establish the strength of the link between a breach of care and claimant harm. There are several reasons for this. Torts are adjudicated under Common law, whereby judges decide outcomes and settlements within the loose interpretation of 'greater than 50 per cent probability' that a breach of care led to claimant losses (McBride and Bagshaw, 2012). This approach is desirable if it achieves social optimums, allowing for errors due to chance. The current threshold for establishing causality in claims is substantially less stringent than that used to establish causation in science. Judges reject or accept evidence below the 95% probability that effects are not due to chance, generally used in EBM and EBP. Unlike EBM, the outcome of an unprecedented case is used to estimate quanta for similar claims, without systematic appraisal of the risk of bias in these historical cases. This raises potential for errors and uncertainty in evidence to compound over time, exacerbating long-run social welfare losses. Befitting EBP, policy makers need to monitor average long-run social outcomes achieved

under tort law methods. If, on average, the socially optimum position is outside an agreed margin of error, then welfare losses to external parties may fall disproportionately on one side. Put another way, other NHS patients may inadvertently receive less resources for their care, due to the marginal costs of inequitable and inefficient negligence settlements. On the other hand, if settlements are too low, such that they do not incentivise providers to adapt their care practices, then the likelihood of the same harm repeating is expected to be undiminished, manifesting as a disproportionate loss to society. Inconsistent data across the NHS hinders policy makers' ability to understand the root causes of negligence (Public Accounts Select Committee, 2017), which extends to the problem of measuring economic costs and formulating EBP. For instance, from the limited empirical evidence available, caps on general damages are consistently the most effective way of controlling costs (Viscusi, 2019); however, this outcome does not capture effects on wider welfare and it shows the narrowness of the evidence base available to policy makers.

As a potential remedy, claim cases are a route to a better understanding of the cause(s) of legal outcomes and their impact on social welfare. Claim cases are an important data source for assessing bias in negligence cases because they provide detailed healthcare and medico-legal information are not available elsewhere. The quality of evidence about clinical-level harm is unique in legal cases because of the cross-referenced detail required by the Courts. As such, the series of events that took place when harm occurred is relatively clear if failures in evidence collection are avoided. As a means of establishing scientifically robust estimates of causality between care delivered and claimant outcome(s), NHS organisations have a significant opportunity to learn, using the record of these investigations. Legal processes are also detailed in claim cases, offering a source for appraising failures in investigations that create uncertainty to quanta calculation. These failures may or may not be avoidable, but they can only be estimated if the claim data are used to determine what is correctable.

10. Actions for immediate and long-term impact

From our study, we believe there are several corrections that can be made, and these should be debated by stakeholders. These actions do not target the tort law system *per se* and, instead, as an extension of Norrie's (1985) explanation, they focus on the fundamentals of not just standards of care, but standards of evidence in medical negligence. The recommendations are also an extension of recently repeated calls for the NHS to address clinical negligence at the source, by becoming a safer system (Yau *et al.*, 2020). The UK NHS has led efforts to improve patient safety over the last decade and, yet, clinical negligence claims continue to grow, which indicates that more acute solutions are needed to address the complexities of clinical negligence.

The first recommendation is that expert evidence is appraised using the now widely available tools provided by EBM (Howick *et al.*, 2011). This is a standard requirement for all evidence used in the NHS, exemplified by the scientific peer-review process or that of the National Institute for Health and Care Excellence (NICE); however, this standard is rarely applied to claim cases. The use of OCEBM methods to appraise claim cases is consistent with the principle of high-quality care in the NHS (Darzi, 2008) and, therefore, wider implementation of these standards by medical negligence stakeholders is uncontroversial.

The findings in this analysis invite several recommendations for care practitioners, administrators, and policy makers, which we expand on here. These are the first steps for developing cost-effective solutions in the immediate and long-term to address, we suspect, widespread societal inefficiencies in clinical negligence. It is essential that the NHS begins a process of transformation in the use of claims data, from the bottom-up, so that risk of bias in claim pathways is addressed at the source. This approach will complement existing efforts to reform the negligence process from the top-down. With relevant (ethical) approvals, quality improvement and patient safety practitioners in all NHS organisations can access their own claims data to identify rapidly

actionable solutions that have short- and long-term benefits to patients and the financial position of their organisation.

First, we recommend that the NHS adopts consistent standards of reporting for expert witness reports. The variables presented are those that we considered a high priority to curate and we recommend this framework be replicated by other NHS Trusts. Prospectively, this will improve the efficiency with which data on care and claim pathways can be extracted for routine analysis. Retrospectively, tools from EBM can be applied to existing case records as soon as they are curated in this manner. This would address concerns about the low sample size used in our study, albeit we view the approach presented in this article as a necessary foundation for that data collection effort. Linked to the first recommendation, but presented separately because of the focus on scientific quality, our second recommendation is that these standards draw from those given by scientific journals, whose requirements for article publication are designed to ensure evidence is presented transparently and comprehensively. Also, a template capturing the background and experience of expert witnesses would support the development of measures to ensure that expert witnesses are selected appropriately.

Third, we recommend that clinical negligence teams implement similar analyses in their setting, using recommendations one and two. Moreover, the approach to improvement should be cyclical, with incremental growth in the sample size, given that a comprehensive analysis of all clinical negligence claims would be a resource barrier to organisations achieving immediate benefits. Recommendation four is that NHS organisations, local and national, should commission arm's length peer-review committees, tasked to appraised claim pathway evidence with impartiality. In the case of local organisations, blinded appraisals should be performed by at least two qualified staff members, wherever possible. Nationally, the remit of a committee should initially focus on the claims that disproportionately account for total NHS costs; this remit could therefore start with claims from obstetrics and gynaecology, which account for 50% of total costs, although representing just 10% of all claims, annually.

The case series was constructed using claims data, which we recommend other NHS Trusts and stakeholders curate and analyse for themselves, as a matter of routine; these teams should be responsible for developing and maintaining these datasets. The resource needed to generate these datasets is not insignificant, however, once conceived, data can be used to invoke organisational changes that benefit all stakeholders. Whilst the relatively small sample size of our study does not lend to generalisable findings, it does initiate the research and regulatory processes for predicting and optimising medico-legal outcomes and reducing their concerning cost to organisations. In the medium- to long-run and with the wider system in-mind, pooling of these samples should reveal the scale of bias in clinical negligence claims nationwide. A larger sample size can provide regulators with representative frontline data to accelerate the delivery of solutions. Therefore, our fifth recommendation is that a national registry is developed, populated by observations mirroring those contained in our research. Longitudinal data with representative sampling will facilitate the delivery of cost-effective solutions to the burden of negligence claims. In a manner, we are mirroring the main recommendation made in Towse and Danzon's (1999) analysis 20 years ago, which called for a claim database in the NHS. We recommend the inclusion of contemporary variables to the national dataset that indicate the scientific quality of medico-legal processes. Solutions are easier to identify when the root causes and variation at the frontline are known and, currently, these remain largely unobserved.

In our institution, this study vastly improved our understanding of evidence bias in expert witness reports that informed the claims and outcomes experienced. These lessons offer practicable solutions to an ongoing policy problem, which we summarised as recommendations in Table 2. Expert witness evidence is used to inform financial settlements, which represent 2% of the total NHS budget (Yau *et al.*, 2020). To our knowledge, this is the first time that market failures driven by suboptimal scientific and reporting standards in expert witness reports have been formally appraised. In the absence of similar evidence, the case series we present sheds light on a

Table 2. Five core recommendations for practitioners and policy makers of medical negligence, from a bottom-up perspective, with suggestions for resourcing

Level	Recommendation	Description	Resourcing
Bottom top	(1) Standardise expert witness reports	In the long-run, an evidence-based reporting standard must be developed and endorsed by medical negligence bodies, led by NHS Resolution. In the short-term, NHS provider organisations can develop templates that are designed to reduce the burden of extracting information from negligence case records	Local quality improvement and medical negligence teams
	(1) Develop scientific standards for expert witness reports	The content of expert witness reports requires higher standards of transparency and reproducibility. Citations of relevant research that informs experts' position must be produced to a scientific standard. To this end, editorial guidelines and standards from peer-reviewed journals can be implemented immediately	NHS Resolution and national experts
	(1) Extract data from local medical negligence cases and apply quality improvement methodology	Sampling from historical claim cases and applying the method presented in this paper will initiate micro cycles of improvement in medical negligence evidence. Routine data extraction and sampling, albeit with more lenient requirements than scientific evidence, is needed to reduce avoidable market failure	Local quality improvement and medical negligence teams
	(1) Commission arm's length peer-review committees	An impartial expert scientific committee should be resourced to provide risk of bias assessments for high-value expert witness reports. These assessments and any recommendations should be provided to the Courts	NHS Resolution and national experts
	(1) Develop a national registry for care and claim pathway data	Continual improvement in medical negligence can only be achieved if complementary bottom-up and top-down datasets are developed. Information extracted from our 15 cases provides a template for this type of negligence database. This should be designed to address the root causes of harm that lead to claims against the NHS. Similarly, indicators for avoidable failures in claim cases can be captured with the intention of measuring the magnitude of uncertainty in evidence informing quanta calculations. These data can be used to correct system-wide information failures, using cost-effective approaches	NHS Resolution, NHS Digital, NHS England, national experts

These recommendations are consistent with and in addition to recent recommendations provided by the Medical Defence Union (2020) and NHS Resolution (2019).

previously unexplored area of claims cases in England. This analysis is needed because detailed data on claims processes is not centralised, which is a hindrance to improvement. Future research is essential to establish the extent to which our findings are generalisable to other NHS settings and other countries. Globally, we believe the findings are of interest to other health systems, particularly those seeking to strengthen their medical malpractice functions (Kinga Bączyk-Rozwadowska, 2011; Wang *et al*., 2017) using methods rooted in EBM.

11. Conclusions

This observational study shows that clinical negligence claims in the English NHS are exposed to measurable uncertainty, some of which can be managed by stakeholders. The association between claim case characteristics and quality of evidence should be systematically investigated to estimate the extent of widespread uncertainty in evidence used to settle claims. In the pursuit of fairer outcomes, NHS providers can assess the quality of expert witness evidence, using the methodological approach presented in this paper. These methods intentionally align jurisprudence and the principles of EBM. Some evidence may be unavoidably biased, but avoidable failures can be readily corrected to reduce uncertainty and potential for unfair and inefficient outcomes to patients and society from clinical negligence claims.[1]

References

Agency for Healthcare Research and Quality (AHRQ). (2019) Adverse events, near misses, and errors. Available at https://psnet.ahrq.gov/primer/adverse-events-near-misses-and-errors [20 March 2020].

Ball C, Sackett D, Phillips B, Haynes B and Straus S (2009) Levels of evidence and grades of recommendations. In: EBM (Web site of the Oxford Centre for Evidence-Based Medicine). Available at www.cebm.net/levels_of_evidence.asp; https://www.cebm.net/2009/06/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/ [26 October 2018].

Birks Y, Aspinal F and Bloor K (2018) *Understanding the Drivers of Litigation in Health Services*. University of York and The King's Fund. Available at https://www.york.ac.uk/media/healthsciences/images/research/prepare/UnderstandingDriversOfLitigationInHealthServices.pdf [20 March 2020].

Bolam v Friern Hospital Management Committee (1957) 2 All ER 118 is an alternative citation for this medical negligence case. This is how legal cases are cited, I gather. For example: https://doi.org/10.1177/0025802416657262

Danzon PM (1983) An economic analysis of the medical malpractice system. *Behavioral Studies and the Law* **1**, 39–54.

Darzi A (2008) Evidence-based medicine and the NHS: a commentary. *Journal of the Royal Society of Medicine* **101**, 342–344.

Deitschel CH, Buckley JM, Evans G, Fraser JJ, McAbee GN, Reynolds SL, Berger JE, Berman JI, Craft WH, Myers H and Ake JK (2002) Guidelines for expert witness testimony in medical malpractice litigation. *Pediatrics* **109**, 974–979

El Emam K, Rodgers S and Malin B (2015) Anonymising and sharing individual patient data. *British Medical Journal* **350**, h1139.

Evidence Based Medicine Working Group (1992) Evidence-based medicine: a new approach to teaching the practice of medicine. *Journal of the American Medical Association* **268**, 2420–2425.

Fenn P, Diacon S, Gray A, Hodges R and Rickman N (2002) Current cost of medical negligence in NHS hospitals: analysis of claims database. *British Medical Journal* **320**, 1567–1571.

---

[1]Citations refer to published or unpublished evidence from bibliographic sources that inform the arguments made by the expert. In the absence of bibliographic citations, we counted the expert report as OCEBM Level 5 expert opinion informing the claim case, which was counted and scored in addition to the appraisal of bibliographic citations.

Financial Times (2017) Cost of NHS negligence claims quadruples to £1.6bn in decade. Available at https://www.ft.com/content/9a7c010a-9307-11e7-a9e6-11d2f0ebb7f0 [20th March 2020].

Gandjour A (2016) Prioritizing health services research: an economic perspective. *European Journal of Health Economics* **17**, 375–377.

General Medical Council (2017) Confidentiality: good practice in handling patient information. Available at https://www.gmc-uk.org/-/media/documents/gmc-guidance-for-doctors---confidentiality-good-practice-in-handling-patient-information----70080105.pdf?la=en&hash=08E96AC70CEE25912CE2EA98E5AA3303EADB5D88 [28th September 2020].

Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L and Sterne JA; Cochrane Bias Methods Group; Cochrane Statistical Methods Group. (2011) The Cochrane collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal* **343**, d5928.

Holt PJE, Poloniecki JD and Thompson MM (2012) Multicentre study of the quality of a large administrative data set and implications for comparing death rates. *British Journal of Surgery* **99**, 58–65.

Howick J, Chalmers I, Glasziou P, Greenhalgh T, Heneghan C, Liberati A, Moschetti I, Phillips B and Thornton H (2011) The 2011 Oxford CEBM Evidence Levels of Evidence (Introductory Document). Oxford Centre for Evidence-Based Medicine. Available at https://www.cebm.net/wp-content/uploads/2014/06/CEBM-Levels-of-Evidence-Introduction-2.1.pdf [20 March 2020].

James JTA (2013) A new, evidence-based estimate of patient harms associated with hospital care. *Journal of Patient Safety* **9**, 122–128.

Kachalia A, Kaufman SR, Boothman R, Anderson S, Welch K, Saint S and Rogers MA (2010) Liability claims and costs before and after implementation of a medical error disclosure program. *Annals of Internal Medicine* **153**, 213–221.

Kinga Bączyk-Rozwadowska (2011) Medical Malpractice and Compensation in Poland, 86 Chicago-Kent Law Review. 1217. Available at https://scholarship.kentlaw.iit.edu/cklawreview/vol86/iss3/9.

Kravitz R, Rolph J and Mcguigan K (1991) Malpractice claims data as a quality improvement tool: epidemiology of errors in four specialties. *Journal of the American Medical Association* **266**, 2087–2092.

Licchetta M and Stelmach M (2016) Fiscal sustainability analytical paper: Fiscal sustainability and public spending on health. Available at https://obr.uk/docs/dlm_uploads/Health-FSAP.pdf [Accessed 16 October 2019].

Mahtani K, Spencer EA, Brassey J and Heneghan C (2018) Catalogue of bias: observer bias. *BMJ Evidence-Based Medicine* **23**, 23–24.

Markides G and Newman C (2013) Medical malpractice claims in relation to colorectal malignancy in the National Health Service. *Colorectal Disease* **16**, 48–56.

McBride NJ and Bagshaw R (2012) *Tort Law*, 4th Edn, Harlow, UK: Longman.

Medical Defence Union (2020) The campaign. Available at https://www.themdu.com/about-mdu/fair-compensation/the-campaign [24th April 2020].

Mello MM, Kachalia A and Studdert DM (2017) Medical liability – prospects for federal reform. *New England Journal of Medicine* **376**, 1806–1808.

Ministry of Justice (2019) Lord Chancellor announces new discount rate for personal injury claims. Available at https://www.gov.uk/government/news/lord-chancellor-announces-new-discount-rate-for-personal-injury-claims [20 th March 2020].

Montori VM and Guyatt GH (2008) Progress in evidence-based medicine. *Journal of the American Medical Association* **300**, 1814–1816.

National Audit Office (2017) Managing the costs of clinical negligence in trusts. Available at https://www.nao.org.uk/wp-content/uploads/2017/09/Managing-the-costs-of-clinical-negligence-in-trusts.pdf [16th October 2019].

NHS England (2018) Lawyers who target the NHS banned from advertising or setting up shop in hospitals. Available at https://www.england.nhs.uk/2018/01/lawyers-who-target-the-nhs-banned-from-advertising-or-setting-up-shop-in-hospitals/ [20th March 2020].

NHS Improvement (2017) Our approach to patient safety: NHS Improvement's focus in 2017/18. Available at https://improvement.nhs.uk/documents/1869/NHSI_approach_to_patient_safety_FINAL_24_Oct_2017_MCWC0eH.pdf [22nd January 2020].

NHS Resolution (2019) *Annual report and accounts* 2017_18. Available at https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/824330/NHS_resolution_annual_report_and_accounts_-_web_pdf.pdf [16th October 2019].

Norrie KM (1985) Medical negligence: who sets the standard? *Journal of Medical Ethics* **11**, 135–137.

Pettker CM, Thung SF, Lipkind HS, Illuzzi JL, Buhimschi CS, Raab CA, Copel JA, Lockwood CJ and Funai EF (2014) A comprehensive obstetric patient safety program reduces liability claims and payments. *American Journal of Obstetrics and Gynecology* **211**, 319–325.

Public Accounts Select Committee (2017) Managing the costs of clinical negligence in hospital trusts. Available at https://publications.parliament.uk/pa/cm201719/cmselect/cmpubacc/397/39705.htm [16th October 2019].

Rimmer A (2018) Government will cap legal costs of clinical negligence claims. *British Medical Journal* **360**, k763.

Towse A and Danzon P (1999) Medical negligence and the NHS: an economic analysis. *Health Economics* **8**, 93–101.

Vincent C and Amalberti R (2015) Safety in healthcare is a moving target. *BMJ Quality & Safety* **24**, 539–540.

Viscusi W (2019) Medical malpractice reform: what works and what doesn't. *Denver Law Review* **96**, 775–792.

Wang Z, Li N, Jiang M, Dear K and Hsieha CR (2017) Records of medical malpractice litigation: a potential indicator of health-care quality in China. *Bulletin of the World Health Organization* **95**, 430–436.

White V, Nath A and Stansby G (2015) Litigation claims relating to venous thromboembolism in the NHS. *Phlebology* **30**, 389–396.

World Health Organisation (2020) Patient Safety. Available at https://www.who.int/patientsafety/en/ [20th March 2020].

Yau CWH, Leigh B, Liberati E, Punch D, Dixon-Woods M and Draycott T (2020) Clinical negligence costs: taking action to safeguard NHS sustainability. *British Medical Journal* **368**, m552.