

Harveian Oration 2018: Improving quality and safety in healthcare

Author: Mary Dixon-Woods^A

Introduction

Challenges in providing consistently high-quality, safe care are common to health systems worldwide. In this lecture I will propose that, while the NHS is not unique in experiencing these challenges, it has a unique opportunity to address them in an evidence-based way, and I will identify an important role for the healthcare professions in doing this.

Let me make two points straight away. First, the simplest argument to be made for improving health services is that they need to be better resourced. And that is absolutely correct. Funding, staffing, estates, equipment and so on are the structural preconditions for delivery of high-quality, safe care. Economists such as the Health Foundation's Anita Charlesworth have shown vividly that the NHS is significantly underfunded relative to the mission it is expected to deliver.¹ But equally clear is that resources on their own, while necessary, are not sufficient. How to get the best from what is available is a key question, and one that requires high-quality evidence to answer it.

Second, I do not want to suggest that everything in healthcare is terrible and that nothing has ever got better. Many things have. The NHS should be rightly proud of its achievements in relation to infection prevention and control, for example.² Many disease areas have seen steady improvement over time.^{3–6} The safety and outcomes of most surgical procedures have improved, with vastly reduced rates of mortality and complications.

But many areas have stalled. Variations in intervention rates, outcomes and evidence-based practices have proven remarkably persistent. Some indicators show worsening performance (for example, but not only, in relation to access). Serious safety incidents and failures of respect continue to recur, sometimes affecting the most vulnerable groups: we can hear in today's crises about particular units, for example, the grim echoes of the Ely hospital scandal involving neglect and abuse of vulnerable people, even though it happened more than 50 years ago. There has been no shortage of other catastrophes since then, and individual units and organisations are still susceptible to disaster and scandal. Failures of learning and recurrence of the same problems are, sadly,

common: the National Confidential Inquiries often report more or less the same problem and make the same recommendations, and then the same thing happens again. The same incidents get reported to the National Incident Reporting and Learning System year on year. And at the same time as it is difficult to implement some new practices, some innovations are far too quick to enjoy uptake, and de-implementing ineffective practices is often very hard. Problems of overdiagnosis, underdiagnosis, and wrong diagnosis are rife. Though many patients have excellent experiences of care, some do not, and for some those experiences are deeply distressing and may be traumatising.

The NHS is now spending a growing – and frightening – proportion of its budget on payments for negligence claims, each one of them representing avoidable harm caused to patients and their families. Staff and patients alike are asked to use systems and processes that are sub-optimally designed for the goals of care, that are hazard-rich, and that cause waste and inefficiency. So many challenges remain.

In this Oration, I will explore some of the things that need to happen to address these challenges. They are by no means comprehensive, but they are each important.

Stop admiring the problems and start investing in evidence to solve them

Both healthcare and research about healthcare have had a bad habit of admiring problems: describing them beautifully, but not actually solving them. Bad things happen, or good things don't happen, and too often the next action is to describe them again, whether through confidential inquiries, public inquiries, incident reporting, root cause analyses, clinical audits, patient and staff surveys, or any of the many systems for harvesting data about quality and safety across the NHS. Though the costs (financial and human) associated with data collection in multiple different forms to serve multiple purposes and multiple bodies are not fully clear, they are likely to be huge, if evidence from the US is anything to go by.⁷ But the proportion of effort that goes into collecting data about problems and the proportion that goes into addressing the problems is often seriously imbalanced. We too often create cliff-like arcs rather than circles of improvement.

Of course, the quest for improvement has not been (at all) absent. Clinical audit, for example, has a noble history dating back over several decades. With its origins in the professional societies, it is based on a cybernetic model of setting standards, evaluating practice against those standards, and changing practice where

Author: ^AThe Health Foundation professor of healthcare improvement studies and director of The Healthcare Improvement Studies Institute (THIS Institute) in the Department of Public Health and Primary Care, University of Cambridge, UK; professorial fellow, Homerton College, Cambridge, UK

needed to ensure better performance. Alongside the large-scale clinical audits run at national level, small-scale local audit remains widespread, especially as an activity undertaken by doctors (including those in training). Some of the large audits, including some run by the Royal College of Physicians (RCP), have been hugely important in revealing deficits in care and are associated with impressive improvements. A valuable example is the Stroke Audit, which is associated with substantial gains in clinical processes over time that appear to translate into lives saved and disability averted.⁸ Not all audits have been quite so successful, however, and on some audits the indicators show significant deviations from the standards of care time after time, with relatively little or no improvement between cycles.⁹ The yield from local clinical audit is even less clear, in part because there is no national oversight of the various projects and what they produce. More broadly, though the basic principles of audit are sound, in practice it tends to attract criticism for its tendency to perpetuate the problem of spending a lot of energy and resource on data collection while leaving rather obscure what is supposed to happen in the 'change' part of the cycle. At policy level, structural reorganisations intended to produce improvement have been a permanent feature of the NHS, as have various kinds of improvement initiatives.

A large part of the effort to encourage improvement has also traditionally gone into various forms of incentive systems, ranging from pay-for-performance through to public reporting and many others in between. Despite their ubiquity, the results of these schemes are mixed: performance and practice remains variable. Worse still, many incentives and performance management systems create unintended consequences,^{10,11} including complaints from professionals and patients alike that they may reduce care to sets of transactional tasks drained of personal meaning. Further, metric-driven schemes tend to promote effort substitution, goal displacement and gaming, and may be understood and treated as blame allocation machines, even when they are mounted with the best of intentions.¹² Aside from these unwanted effects, what has also become vividly clear is that stimulating organisations to improve is not the same as knowing *how* to do it.

Into this void has come a relatively new field of practice devoted to the *doing* of improvement. One of the most prominent of these (but far from the only) is a distinctive set of practices known as quality improvement (QI), or sometimes as improvement science. I will mostly use the term 'QI' here; the term 'improvement science' is now used to mean so many different things that it does not offer clarity. As with any field of practice, QI has its own approaches, key texts, assumptions, gurus, and ways of working. Several currently popular approaches tend to fall into a number of genres and some have distinctive 'brands' that are promoted by organisations that use or develop them. Many of the better-known have their origins in the work of Edwards Deming, Joseph Juran, Walter Shewhart and others who developed methods for measuring and controlling process variation in industrial settings during the twentieth century. In healthcare, these tools, techniques and principles form the basis of the 'Model for Improvement' developed by Associates in Process Improvement and popularised, along with a range of other methods and ideas, by the Boston-based Institute for Healthcare Improvement (IHI). Many of the methods are technical in character, and include statistical process control (SPC) and 'Plan-Do-Study-Act' (PDSA) cycles to test solutions, and are intended to be complemented by what Deming rather grandly termed a philosophy or 'system of profound knowledge'.¹³ A

second well-known and widely used approach is termed 'Lean'. A distinct package of concepts, methods and tools initially developed as part of the Toyota Production System, it owes some of its basic principles to Deming, but also to others.

These approaches (or methods) to the doing of improvement can be distinguished heuristically from improvement interventions, which may include, for example, specific training programmes, 'bundles' of evidence-based practices, checklists, methods of handover, scoring systems, devices, and many others. QI approaches may be used to implement these interventions: for instance, the Model for Improvement may be used to support the introduction and maintenance of an infection prevention bundle.

QI has been promoted as a way of addressing many of the pressing challenges faced by health systems, including its promise that, by reducing variability and improving process control, it will deliver better efficiency, value, consistency and experience. It is now 30 years since Don Berwick, in the *New England Journal of Medicine*, made a passionate plea for 'understanding and revising the production processes on the basis of the data about the processes themselves... Translated into cultural norms in production systems and made real through sound statistical techniques.'¹⁴ The plea to undertake QI was reiterated again in the Berwick report,¹⁵ commissioned in response to Sir Robert Francis' inquiry into the catastrophic events at Mid Staffordshire NHS Foundation Trust.

It is clear that some change is indeed happening. Some NHS organisations have embarked on ambitious QI programmes. Recent years have seen the emergence of a generation of QI practitioners – people whose roles include the doing of improvement as part of their jobs, and who have often had formal training in one or more improvement methods. Improvement academies are now springing up as interest in the 'how to' of improvement grows. Professional regulators, including the GMC and the royal colleges, are now requiring that many professionals have some exposure to QI through their requirements for accreditation and qualification.

Though many signals of increasing QI activity are evident, it is probably fair to say that uptake is still very uneven. In many ways this is a puzzle. Given the known defects in health systems, it is reasonable to ask: if QI is so good, why don't all healthcare organisations do it. Why is it such a hard sell? And if thousands of professionals are now mandated to do QI or clinical audit, why are the benefits not more visible?

The reasons are multiple. It is hard to build awareness and capacity quickly. Not all organisations know about QI, and those that do have to balance many competing priorities with developing improvement capacity. There are many other explanations. Here, I am going to focus on one important contributor: the weakness of the evidence base supporting QI. It remains much too difficult to answer the question: does quality improvement improve quality?¹⁶ It is, for example, surprisingly difficult to demonstrate the effectiveness of any specific QI approach, making it hard to give a clear answer to the chief executive who asks the very reasonable question of which she should adopt in her organisation. Cost data for QI is scarce, so return on investment for hard-pressed organisations with many competing priorities is difficult to establish. Faced with this, it is perhaps not surprising to find scepticism and reluctance to invest in QI. It is obviously important for this, and for other reasons, that the field of practice benefits from a strong evidence base that shows what works, what doesn't, and why.

Commit to evaluation

A new field of study of improvement is now rapidly emerging, and is an important evolution of the tendency of researchers to make their own contributions to admiring the problem. Dating back to the work of Sir William Petty in the seventeenth century, research has been too often preoccupied with describing problems rather than producing and evaluating the possible solutions. Even the work on variations, which has its modern-day origins in James Allison Glover's work in the 1930s, has too often been content to present the information on the problem without giving much attention to the next stage in the circle.

The field of study of improvement overlaps, of course, with the field of practice of improvement, but the two fields are not quite the same. They tend to co-exist in a somewhat uneasy relationship, informing and shaping each other, but also sometimes in tension, particularly in relation to the nature and quality of evidence needed to make claims of effectiveness. These tensions between the field of practice and research can be productive. How to make that so will be a key challenge in developing both fields, and is especially important in confronting the problem that the literature on improvement is, ironically, in much need of improvement. Though things are getting better, many studies of improvement are poorly done and poorly reported.¹⁷

In part (but not exclusively) this is because many improvement studies are effectively self-evaluations, often of local projects. Measurement and data analysis are not always done well;¹⁸ validated measures, standardised data collection systems, trained data collectors, and methods for minimising missing data are not always used. Reports do not consistently properly report pre- and post-metrics, describe measures adequately, or use statistical tests appropriately to detect pre-post differences.¹⁹ Few projects are set up to identify unintended, perverse, or reverse effects. Fidelity is often compromised. For instance, PDSA cycles are a crucial element of the Model for Improvement approach, yet projects reported in the literature that claim to use PDSA do so properly only 20% of the time.²⁰ Many are constructed essentially as technical projects, with Deming's 'philosophy' little in evidence. When it comes to improvement interventions, some are sound in principle but are corrupted when done inauthentically,²¹ making them very hard to evaluate. The social or technical processes they seek to alter or their theoretical basis (the means by which an intervention might reasonably be expected to achieve the hoped-for effects) are not always clear,²² and descriptions of what the intervention comprises are often so poor that replication and scaling is near-impossible.²³ A failure to curate or cumulate learning from these projects compounds the problems.

The reasons for these failures in producing a high-quality evidence base for improvement are multiple. One has been an ongoing reluctance to fund it. Those commissioning or implementing improvement programmes or projects might be the natural source of support for evaluation of those programmes, but, with some notable exceptions – including the Health Foundation – there is a tendency to perceive evaluation as unnecessary, frivolous, or a diversion from resource that could and should be spent on the improvement. The UK's infection prevention and control programme, for example, extended over a decade with very little evaluation, resulting in regrettable loss of learning.²

Research funders might be the other obvious source of support for studies of improvement. But, though the gains from improved reliability in delivering what we already know works might outstrip

those from new treatments,²⁴ biomedical moonshots have often been more bewitching targets for research investment than the unglamorous groundwork of care. The metaphors alone are revealing: linear models of science, with the service and patients at the far end of a translational pathway waiting to be penetrated, versus cyclical models where each stage of the cycle influences the next, tend to dominate.

A third reason for some of the failures to develop the evidence base concerns an ongoing and sometimes conflicted debate about what should count as evidence for improvement and the methods for generating that evidence. One argument is that improvement needs to be treated like other forms of biomedical intervention, and should seek to build its evidence base through studies that conform to the scientific norms of biomedical research. A recent editorial along these hardcore lines²⁵ argued that QI studies:

[...] should have results that are generalizable; if the problem exists only in 1 center or the intervention can be performed only in a limited number of facilities, the priority for publication would be markedly lower. Studies should document that the intervention actually resulted in better health outcomes, rather than focus only on changes in health care processes, use, or costs. Studies should examine potential adverse outcomes as well as benefits. In addition, we encourage the evaluation of costs, including the costs of the intervention, to enable the determination of value. At a minimum, quality improvement studies should have a concurrent control group to minimize the effect of temporal trends, which can have a substantial association with the quality of health care. Better yet, quality improvement studies should use randomization and blinding.

These standards are exacting. Many – if not most – studies of improvement have to be done in the real world, where the perfect study design is rarely an option. Learning has to be harvested in the pressurised and highly pragmatic context of the real-world of care, where operational realities and clinical priorities crowd the available space for research priorities. Many of the technical challenges of such studies are familiar:²⁶ blinding may be extremely difficult or impossible; determining causality is challenged by the complexity of the causal chains, by multiple confounders and secular trends;²⁷ it is hard to measure many outcomes of improvement efforts reliably and validly, over a reasonable timescale, and in the right places, especially if hard patient outcomes such as mortality are used; and improvement policies and programmes are prone to being abandoned or distorted during the course of implementation in ways that cannot be controlled by researchers.

These challenges mean that studies using robust designs to determine effectiveness of improvement have remained rare.²⁸ They are now increasing in number, and increasingly innovative study designs are being used (including cluster-randomised and stepped-wedge designs, often bundled with process evaluation).

The field of improvement practice and policy has not, however, always been enthusiastic about formal evaluation, in part because of a sense of a urgency (or impatience): there is so much that needs to be done, and the ways of solving it are so self-evidently obvious, that we should just get on with it, goes the argument. It is indeed true that individual improvement projects tend overwhelmingly to report positive results.²⁹ These positive findings are not, however, always reproduced in larger, more robust studies.¹⁶ Evaluations that produce these kinds of disappointing

findings about favoured interventions are sometimes challenged on the basis that traditional epidemiological study designs are ill-suited to the study of improvement.³⁰ The 'lovely baby' syndrome, where people eschew the need for evaluation on the grounds that something looks so good it must work (especially when they have designed it themselves), plays a role in this. But assumptions about effectiveness, however well-intentioned, should not be the basis of practice.

Evaluation is essential to an evidence base for improvement. Without evaluation, patients may be exposed to risk of harm or deprived of benefit, and resources may be wasted on interventions that are ineffective^{31,32} or distribute risks unfairly. Some interventions are well-intentioned, and appear initially plausible, but upon evaluation turn out not to work or even result in harm or deterioration.³³ Some interventions, while they may produce some small change, are just not worth the candle: they require too much effort and opportunity cost to be warranted.³⁴ Evaluation has a particular role in monitoring the side-effects of improvement interventions, even those that appear on the basis of evaluations of those interventions to be effective and safe. Risk management systems themselves generate new risks, for example.³⁵

Sometimes the unintended consequences occur quite distally (in time or place) to the target of intervention, and have to be sought out through careful and holistic analysis. For instance, James Lind's trial of citrus fruits as a treatment for scurvy is rightly held up as an example of how experimental studies can produce evidence that benefits patients. A few interesting wrinkles appear in the story, however.³⁶ First, it is a good example of how something that works can take a very long time to find its way into practice. Lind published *A Treatise of the Scurvy* in 1753, but it was more than 40 years before the admiralty was convinced to approve use of citrus fruits as a preventative measure. William Harvey had similar problems in having this theory of the circulation of the blood accepted, though it only took him 20 years. Second, it is also an example of how things can go wrong if the mechanisms are misunderstood or an intervention is not used in the way intended. Over time, the active ingredient in citrus fruits was forgotten. The British Antarctic Expeditions were a twentieth-century casualty. Dr Reginald Koettlitz, the expedition's physician, was deeply sceptical of 'antiscorbotics', believing instead that scurvy was due to tainted meat. Outbreaks of scurvy were the result. The third wrinkle is an indication that the rise in citrus fruit consumption was implicated in the rise of the mafia in Sicily. This was clearly a consequence unintended by those trying to find a solution to a terrible disease, and could not have been factored in as a secondary outcome in a clinical trial.

Attempts at improvement are similarly not at all immune to unintended consequences, but our current ways of doing improvement do not always lend themselves to their detection and correction. Despite all the rhetoric about 'system-based' approaches and balancing measures, most QI projects are focused on a single condition or pathway, and they do not always consider the whole range of possible effects on whole organisations or systems they might produce. By focusing exclusively on the extent to which the QI intervention produced the desired local effects, the 'toxicities' or iatrogenesis that a QI intervention might produce are rarely given explicit attention or documented. One checklist or sticker might well be a good thing, for example. But too many – the problem known as polyformacy³⁷ – may start to produce unwanted effects. QI projects tend to

focus on single, relatively well-bounded processes, often (though not always) focused on a single condition. Decisions and actions may have distal or delayed effects, may work in the short-term but not the long, or may worsen performance somewhere else or over the long-term.³⁸

Improvement efforts thus need to assess the interactions, impacts and effects on other organisational functions and processes. It is clearly wrong to suggest that improvement should not be studied because it does not need it, cannot be studied or is unworthy of study. It is time to do more to unite the fields of practice and research. It may well be the case that we need some epistemological and methodological shifts to accommodate the specificities of trying to study improvement, and that the hardcore approach is too limiting.³⁹ These shifts in *how* to study improvement must take account of the priorities and realities of the field of practice. As part of that task, the field of study should create scientific assets that are directly useful for the practice of improvement, including, for example, libraries of well-validated measures that assess the areas of care that most matter to patients and staff, methods of analysis and visualisation, and practically oriented theories that help to cumulate knowledge about interventions.

It will remain important to avoid grand claims about the impact of improvement on the basis of no or limited evaluation, or on the basis of study designs that do not permit such conclusions to be drawn.⁴⁰ It is also important that academic evaluations and critiques are not seen as hurtful and damaging to the cause rather than advancing the field of practice. High-quality evidence can help the field of practice by establishing credibility, value and relevance, as well as providing a basis for action, thus optimising the use of precious resources on things likely to deliver benefit and avoid consuming time and energy in things unlikely to work. Much could be achieved using the principle of 'no improvement without evaluation' by refiguring both the field of practice and the field of study so that their goals and practices are much better aligned. A shared commitment to learning and respectful relationships between improvers and researchers is therefore critical.⁴¹

This is why the very large investment that the Health Foundation has made in The Healthcare Improvement Studies (THIS) Institute at the University of Cambridge is an especially exciting one that has come at just the right time. Working with more than 20 outstanding partner organisations, including major patient charities, THIS Institute will develop capacity in the study of improvement through a UK-wide fellowship programme that can be held at any university. It will develop the science behind the study of improvement and it will run large-scale participatory research programmes. I am deeply honoured to be its inaugural director, and privileged to be positioned to help create a scientific asset for the NHS – an institution that I see as the very embodiment of our collective commitments and responsibilities towards each other as humans.

Improve the design and implementation of improvement

Fitting the right kind of solution to the right problem is a key challenge for improvement, but at present our methods for developing and testing improvement interventions and safety solutions are generally sub-optimal. The mechanisms of change are often poorly specified or simplistic, and interventions

selected may be indifferent to the wider environments in which interventions are implemented or the scale needed for optimal system performance. We need to get better at developing or selecting interventions that have a high likelihood of success using the right kinds of expertise, testing them rigorously in different contexts, involving the relevant stakeholders (including patients and staff) at the right times, and considering the potential for scale from the beginning.

An important development will be the recognition that one size of improvement approach does not fit all. Currently popular approaches to improvement may be very well-suited to fixing particular kinds of processes, but may be much less effectual at other challenges. For example, one of the largest bodies of literature on improvement is on audit and feedback, an improvement approach that essentially involves collecting data on performance and sharing it back with practitioners in an effort to encourage individual behavioural change. Systematic reviews do indicate that it can have some small effects (of the order of 4.3% on average).⁴² The range is, however, very broad, and for some areas of practice, audit and feedback does not appear to work at all – probably because behavioural change is not the correct target for intervention. That QI approaches have variable impacts is also evident from an examination of their effects on different processes in the same programme. For instance, a study of a quality improvement collaborative to improve stroke care found that issues that were under the control of a bounded unit and that could be addressed by a better process seemed more tractable to improvement than did problems that required large-scale coordination and negotiation across the hospital.⁴³ Better solution and intervention development is likely to require much more use of prototyping, modelling and simulation, as well as testing in different scenarios and conditions to work out what are the core, non-negotiable elements and what can be locally customised.¹⁶

More broadly, not everything that secures improvement is a formal QI approach or a well-bounded improvement intervention. Not everything needs to be a defined, discrete project with formal PDSA cycles, SPC charts and everything else. For instance, some improvement comes from *noticing*. Some of the people best placed to do the noticing work are those directly engaged in using the systems: patients, relatives, carers, porters, pharmacists, doctors, nurses, healthcare assistants, and many others.^{44,45} But our systems for facilitating noticing are not good. We have systems for making complaints and reporting incidents, but they are not necessarily suitable for the person who notices that patients have to wait an extra hour in clinic because of the order in which the tests are done or who notices that there could be a much better place to put the clock.

Just as importantly, the importance of context for improvement needs to be more fully recognised. When he found excess mortality at L'Hotel Dieu hospital in Paris, Sir William Petty speculated that 'evil administration' of the hospital might be to blame, but also proposed that 'either the Physicians and Chirurgeons of London are better than those of Paris, or that the Air of London is more wholesome'. The searching for explanations of variations in outcomes in structures (facilities, organisational arrangements, workforce and so on), processes (what it is that clinicians and others do) and population differences remains just as relevant today. The success of improvement depends not just on the interventions, but also crucially on environment: improving

processes may take us so far, but if, for example, the basics of structure and resources are not in place, no further.

Better quantitative modelling of contextual variables relevant to improvement is much needed.⁴⁶ Without this, many structural-level influences – such as size, staffing rates, and case mix – may remain unclear, yet are likely to be important risk factors for poor patient outcomes and experiences. We are still not as good as we should be at understanding or taking into account these influences in seeking to do QI, despite evidence of, say, a relationship between staffing levels and infection rates,⁴⁷ and nursing levels and mortality.⁴⁸

Considerations of context must go wider still, and examine the nature of the causal relationship between QI and performance. A distinguishing feature of many high-performing organisations, including many of those currently rated as outstanding by the Care Quality Commission, is that they use structured methods of continuous quality improvement. How much of this success can be attributed to 'branded' improvement methods is much less clear; some organisations, such as the Southmead maternity unit in Bristol, have shown evidence of consistent high performance and safety without subscribing to any of the major brands, for example.⁴⁹ One plausible explanation is that the precise detail of the improvement approach used may not matter *too* much, especially since careful examination shows that most of the major approaches share many similarities, and that the general principles are likely to be more-or-less the same.⁵⁰

More broadly, it is difficult to dissect out the contributions of the improvement approach from the contributions of other generative features of the organisations that adopt them: the kind of place that is to the forefront in initiating and sustaining QI activity is also the kind of place that has all the other characteristics that facilitate quality and safety. On the other side of the coin, organisations that are performing poorly lack not only the infrastructure for QI, but also demonstrate a much broader set of organisational failings – suggesting that the kind of place that does not have a QI programme is also one with other problems.⁵¹ A more satisfying explanation than one that seeks to attribute aliquots of effects to either improvement approaches or contexts is likely to be found in a complex systems approach⁵² that recognises how interventions/approaches and contexts are co-constitutive and mutually emergent.⁵³ Simply put, high-performing organisations are likely to have a structured approach to improvement, but adopting a structured approach is probably not enough in its own right. For QI to work, the right kinds of organisational and institutional features have to be in place. The evidence for this comes not just from healthcare, but from the wider literature on business and economics. Organisations across many industries display the same pattern of variations as healthcare organisations, including large and persistent differences in performance and productivity levels between seemingly similar enterprises.⁵⁴ External environments do play some role, but recent important work, some of it experimental, is beginning to show that what distinguishes them is the quality of their management practices.⁵⁵ These practices include, but are not limited to, continuous quality improvement; they also include skills training, human resources and operational management, for example. Thus, though you wouldn't want to be without QI, attention is also needed to broad organisational strengthening if its full benefits are to be seen, and QI in an

inadequate organisational environment may have limited impact.

Think beyond effectiveness

Important as they are, evaluative studies of individual programmes, approaches and interventions cannot answer all of the questions that may be worth asking (and answering) about improvement. I will be forever grateful to the Wellcome Trust for giving me a senior investigator award in 2012 that allowed me to begin to ask some of these, and I will continue to ask them through THIS Institute. The kinds of questions that go beyond effectiveness include, for example, how structures of power and influence determine who gets to 'own the definition' and determine what counts as safe, quality care; who is able to set rules, standards and thresholds, monitor performance, and take action, and with what authority; what sanctions and incentives are used, and by whom; what institutional and organisational structures are relevant to improvement which 'institutional logics' characterise the field and direct attention towards particular issues and selection of particular solutions; whose interests are served (and disadvantaged) by how fields are organised; and whose expertise is privileged, and who gets to take action on what authority. The study of improvement may also turn its gaze on the research field itself, again questioning the ways in which the field constructs the research problems to be addressed, the nature of the assumptions embodied in the literature, the choice of research methods, and the extent to which it engages with normative questions (the 'ought' questions) as well as evaluative ones.

Thinking about the study of improvement as a wider field of inquiry facilitates examination of assumptions about what is seen as important and relevant and thus counts as a problem, how structures of power come to act on those assumptions, and what gets seen a solution or as a method for generating solutions and why. It helps to make visible and draw into debate the attention given to different topics in improvement as a field of practice, examining how certain problems come into visibility and attract resources and attention that may amplify or diminish over time.^{56–59} We might, for example, interrogate how problems of quality and safety are identified and defined, which problems and deficits in health systems are seen as worthy of attention (both in practice and in research) while others remain obscured or neglected, and which problems are selected as the focus for improvement efforts, on what warrant, for whose benefit. Why, for example, has surgery seen welcome improvements in safety over time, but improving the safety of those with mental illness or learning disability has received rather less attention in practice, policy and research⁶⁰ despite high levels of morbidity and mortality, and evidence of serious harm and failures of organisational learning? The challenge is also broader than the issue of how attentional resources get distributed to particular clinical areas, and extends to why, for example, the topic of social inequities in healthcare improvement has remained so muted.⁶¹

Once we think about improvement in this way, many aspects of improvement can be recognised as having a technical character, but also a normative character – as was beautifully explained in a recent special issue of *Health Care Analysis* edited by Alan Cribb.⁶² For instance, standards are generally very important for quality and safety, but are prone to a number of technical problems, for example relating to the rapidly evolving and often

uncertain nature of clinical evidence base – something William Harvey himself experienced. But technical problems are joined by conflicts over what constitutes the 'right' standards of practice, which groups and alliances get to define those standards, whose priorities are valorised, and the different possibilities for action that are created. Standards, accordingly, are as much about values, judgements and epistemic contestations as they are about technical definition.^{63,64}

Next, we need to examine the influences on our understanding of the right ways of intervening in systems. We might, for example, ask questions about what influences choices of improvement approaches, who gets to make decisions about those choices, on what authority, and with what consequences. Some interventions – like education and training – have a kind of abject status. Despite the evident potential for improvement if done well,⁶⁵ and the vast scale of it across the NHS workforce, training is largely seen as undeserving of evaluation or research, and the risk is that much of it is not done well at all. Some interventions, such as the surgical checklist, have diffused rapidly worldwide despite being supported by an initially very weak evidence base (an uncontrolled before-and-after study in which only three of the eight participating hospitals showed improvement, and two of the three showed no relation between compliance with the checklist and the outcomes).⁶⁶ At the same time, other interventions that appear to have good evidence, including those focused on teamwork,⁶⁷ have been very slow to spread.

Similarly, some 'branded' approaches to improvement (like the Model for Improvement and Lean) have become institutionalised as a form of orthodoxy, but other potentially valuable approaches to improvement have had limited penetration or take-up. They include systems approaches to improving quality and safety, which are based on the premise that that many quality and safety failures are preventable through careful system design. A systems approach draws on a number of key engineering principles, on the experiences of other high-risk sectors, and on the science of human factors/ergonomics. Despite advocacy for systems approaches in healthcare dating back to the late 1950s, and regular exhortations since then, their impact in healthcare remains much less than might be expected.^{68,69} As one example, many systems engineering techniques that are common in other industries remain surprisingly little-used.⁷⁰ Similarly, though there has been a rapid increase in the number of improvement practitioners in the NHS, the same cannot be said of people with human factors and ergonomics expertise. NATS (the air traffic control service), with a staff of 4,500 people, employs 25 human factors practitioners.⁷¹ The number of NHS trusts that directly employs a chartered ergonomics and human factors specialist is one.⁷¹

It is also important that we attend to the assumptions and values embodied in the logics of interventions and approaches. Paul Batalden has, for example, in the current excellent *BMJ* series on quality improvement, called out the 'product dominant' logic that characterises many efforts at healthcare improvement, which assume that one party makes and then conveys a product to a consumer.⁷² He proposes that we need instead a 'service dominant' logic that sees health as co-produced with patients. Another logic embedded in many current methods for controlling risk and improving healthcare is a reliance on individual performance and personal vigilance,⁷³ reflected in the enduring preoccupation in both research and practice with individual

clinician (and particularly doctor) *behaviour*. Yet too much of a focus on individual behaviour risks perpetuating narratives of individualised blame rather than forcing attention onto the defects in systems and processes that may cause or influence the behaviour. Recognising both the importance of individual agency and its limits is critical. Opportunities for workers to behave well are made logistically possible and cultivated culturally both by the organisations they work in and by wider institutional structures. Systems need to be designed and operated to support, cultivate and sustain individual competence and virtue;⁷⁴ a key responsibility both of organisations and accountability systems is to provide the right conditions for individuals to behave as well as possible: as Kant would put it, 'ought implies can'.

Giving attention to how improvement is organised institutionally allows recognition of further issues. Healthcare has thousands upon thousands of processes, pathways and product lines. Just one small organisation, like a general practice, will have hundreds. Many of these interact with others in more or less loosely coupled ways. Yet QI is often conducted as a highly local activity, with each organisation painstakingly working out its own solution or intervention for each problem. Though sometimes top-down solutions and interventions are imposed when what is needed is local innovation or customisation, it is also the case that sometimes interventions and solutions are left up to local innovation when they need to be standardised, or at a minimum harmonised, across the NHS. The standardisation of the crash call system is an example.

The localised and often very small nature of much QI means that the solutions that are developed do not always benefit from enough of the right kinds of expertise, such as the technical skill in human factors/ergonomics necessary to engineer a process or devise a safety solution, and they sometimes end up being something more akin to a bodge. Having hundreds of organisations all trying to do their own thing means a very high level of waste, and the reinvention of many wheels introduces inefficiencies and risks, since people moving around the system have to learn new processes every time. The localised nature of QI further results in an intra-organisational approach to improvement, often focused on specific service lines or patient groups. Accordingly, it is typically implemented in ways that are insufficiently attentive to the *interorganisational* nature of health service provision.

These problems are compounded by the current norms, organisation and incentives of the field, which is characterised by the problem of many hands. First described by the political philosopher Dennis Thompson,⁷⁵ this arises in contexts where multiple actors – organisations, individuals, or groups – contribute to the performance seen at the system level, but no single actor can be held responsible for the overall outcome. As a result, system weaknesses develop and unresolved issues accumulate over long periods of time; responsibility and authority for coordinating action to rectify structural deficiencies is diffused, confused or absent; and a profusion of localised practices and components erode the integrity and functioning of the system as a whole.⁷⁶

In healthcare, one consequence of this problem (and there are many) is poor coordination of improvement activity and the creation of incentives for particularisation rather than generalisation. Some of this arises because of competitive pressures built into the NHS that may suppress sharing. Some of it arises because the legal and institutional infrastructure

does not facilitate a national approach to many issues. But a further problem is that a lot of QI is currently carried out by doctors for purposes of fulfilling their ARCP (Annual Review of Competence Progression, for doctors in training) or revalidation requirements. Some of this work is very valuable. But, as has also happened with clinical audit,⁷⁷ some of it is perhaps not. It is patchy, uncoordinated, and, because of its nature, focused on the short-term and on small, often bitty projects. It rewards individual claims of achievement that can be recorded formally. A particularly pernicious feature of the current system is that it rewards 'leadership' and initiation of small individual projects rather than participation in large-scale programmes or replication of previously successful interventions. The result is likely to be an erosion of the collective endeavour likely to be most useful to improvement overall.

A better approach to intervention development is going to require solving the problem of many hands, not least to find ways of agreeing on the kinds of challenges for which full standardisation and interoperability are needed across the sector, which solutions can be agreed at the level of principle and left up to local customisation at implementation, and which should be entirely locally developed. It is also going to need aggregation rather than disaggregation of effort. The medical profession has a crucial role to play in all of this.

Respect the role of professions

The founding of the RCP by Henry VIII 500 years ago marked the beginning of the ascent of the medical profession as one of the most enduring and influential public institutions. For a very long period, 'professionalism' had two senses. One was the ability to exercise discretion and act autonomously without direct supervision or coercion by those outside the medical collegium. It meant that individual members of the profession could, to a large extent, set the standards governing their work and claim immunity from external scrutiny and supervision. The second sense of professionalism is as a normative value,⁷⁸ describing both exceptional kinds of expertise (technical knowledge and skills) and ethical commitments (a public interest orientation and a fiduciary relationship with clients).^{79,80}

The last half-century has seen both of these claimed characteristics of the professions come under scrutiny and challenge, under the weight of growing public and policy-level scepticism, changes in ideological dogma, evaporating deference to the professions, academic critique, evolving models of financing and organisation of healthcare, and incident and scandal.⁸¹ The result has been large-scale transfer of responsibility for defining, controlling and supervising quality and safety of care into a new 'polycentric'⁸² regime involving multiple agencies and actors that includes regulators, commissioners, insurers, academics, consultancy organisations, charities, and patients and their advocates. Unprecedented external regulation, accountability, oversight and surveillance are now the norm, displacing 'doctors know best, and can be trusted to do their best without interference' as the prevailing philosophy. Rather than spelling the end of professionalism, however, the new era of quality and safety has become an opportunity to renew professionalism, a means of retaining its importance in the world and its moral stature.⁸¹ The profession has regrouped and reconfigured itself in response, and continues to be a hugely

important institutional force for good, as even sociologists who were once its critics have recognised.⁸³

The new role for the profession in quality and safety is not the individual-level autonomy of old, controlled through club-like structures, but professional autonomy at a collective level.⁸¹ The royal colleges, including the RCP, have taken on the task of driving quality in health systems through multiple mechanisms, including their traditional education and training functions, but also through improvement programmes. Their efforts to provide high-quality training in QI are an example of the commitment of the colleges to improving care, and these efforts can yield huge benefits, especially if some of the challenges identified earlier regarding the current organisation of QI can be addressed.

One especially important advantage of the medical profession and the royal colleges in securing improvement is that they form social structures known technically as 'networks'. In contrast to hierarchical relationships within formal organisations, contracts, or market forces, networks are distinctive for being held together and functioning through cooperative social connections.⁸⁴ In networks, exchanges occur through reciprocal, mutually supportive relationships, where aspects of trust, interdependence, and reputation are key to ensuring speedy and efficient exchange of 'know-how'. The networks in which people participate shape the norms and values that guide their decisions and actions, the opportunities available to them, the constraints on what they do, and the activities they undertake.⁸⁵ They can influence behaviour through 'economies of regard';⁸⁶ social controls, including the threat of social sanction from peers, may be uniquely powerful in directing professionals' behaviour.⁸⁷ Networks are thus highly effective and efficient ways of getting things done. Given the strength of networks, it is thus perhaps not surprising that some of the most impressive achievements of recent years in improving quality have been led by the professions themselves. This is important, because professionally-led initiatives, as a general principle, are much more likely to succeed than those that are hierarchically imposed.

Further roles that may be especially helpful in building further momentum for improvement from within the professions include convening, coordinating, and advocacy. Allowing a thousand flowers of quality interventions to bloom is not a sensible or efficient way of going about fixing healthcare. Many of the quality and safety challenges that confront healthcare need to be solved at the level of entire systems, not organisation by organisation. The royal colleges have an important role in putting the 'national' back into the National Health Service by convening and coordinating the responses to challenges, ensuring that procedures and systems are designed with the right expertise, tested properly, implemented with professional leadership at the core, and remain open to innovation. The right structures for enabling this will need to be designed, but they will need to be properly inclusive, involving patients, carers, healthcare and multiple professional and academic disciplines who can work together to agree on solutions that are satisfying, workable, informed by professional values, clinical expertise and the priorities and needs of patients, and are capable of being customised for specific situations where needed. To achieve this, it may be necessary to plan in terms of long-term programmes of work that are coordinated through some central hub, and that doctors-in-training and others work on for particular periods of time as part of a contribution to a bigger effort. The

incentives for improvement work in terms of career rewards and satisfaction need to be addressed as part of this, and aligning the goals of improvement and research are likely to be especially helpful in doing this.

Finally, the royal colleges, including the RCP, have a unique and respected voice. They can take on important roles in political advocacy, not least by forming important alliances with patients and other stakeholders, and by advocating for action on problems where the responsibility lies outside healthcare itself – for example, in the ongoing failure to address issues of alarm fatigue, incompatible devices, or drug-naming and packaging practices.

Conclusions

Healthcare has many quality and safety challenges. A lot has been learned about how to address them; an awful lot still needs to happen. Much can be achieved, I propose, by ceasing to admire problems and putting the effort into solving them in an evidence-based way; by committing to evaluation in ways that combine the goals of improvement and research and the interests of all; by improving the design and implementation of improvement; by going beyond considerations of effectiveness; and by recognising and respecting the role of the health professions in achieving the goals of improvement. ■

References

- Charlesworth AJ, Johnson P (eds). *Securing the future: funding health and social care to the 2030s*. London: Institute for Fiscal Studies and the Health Foundation, 2018.
- Holmes A. Infection prevention and control: lessons from acute care in England. *Towards a whole health economy approach*. London: Health Foundation, 2015.
- Neuburger J, Currie C, Wakeman R *et al*. The impact of a national clinician-led audit initiative on care and mortality after hip fracture in England: an external evaluation using time trends in non-audit data. *Medical Care* 2015;53:686–91.
- Brown JR, Solomon RJ, Sarnak MJ *et al*. Reducing contrast-induced acute kidney injury using a regional multicenter quality improvement intervention. *Circ Cardiovasc Qual Outcomes* 2014;7:693–700.
- Cloud G, Hoffman A, Rudd A; on behalf of the Intercollegiate Stroke Working Party. National Sentinel Stroke Audit 1998–2011. *Clin Med* 2013;13:444–8.
- Montroy J, Breau RH, Cnossen S *et al*. Change in adverse events after enrollment in the national surgical quality improvement program: a systematic review and meta-analysis. *PLOS ONE* 2016;11:e0146254.
- Blanchfield BB, Demehin AA, Cummings CT *et al*. The cost of quality: an academic health center's annual costs for its quality and patient safety infrastructure. *Jt Comm J Qual Patient Saf* 2018;44:583–9.
- Rudd AG, Hoffman A, Paley L *et al*. 20 years of researching stroke through audit. *Clin Rehabil* 2018;32:997–1006.
- Boyle DA, Keep DJ. Clinical audit does not work, is quality improvement any better? *Br J Hosp Med* 2018;79:508–10.
- Himmelstein DU, Ariely D, Woolhandler S. Pay-for-performance: toxic to quality? Insights from behavioral economics. *Int J Health Serv* 2014;44:203–14.
- Woolhandler S, Ariely D, Himmelstein DU. Why pay for performance may be incompatible with quality improvement. *BMJ* 2012;345:e5015.
- Armstrong N, Brewster L, Tarrant C *et al*. Taking the heat or taking the temperature? A qualitative study of a large-scale exercise in

- seeking to measure for improvement, not blame. *Social Sci & Med* 2018;198:157–64.
- 13 Deming WE. *The new economics: for industry, government, education*. MIT Press, 2000.
 - 14 Berwick DM. Continuous improvement as an ideal in health care. *N Engl J Med* 1989;320:53–6.
 - 15 Department of H. *A promise to learn: a commitment to act. Improving the safety of patients in England*. London: Department of Health, 2013.
 - 16 Dixon-Woods M, Martin GP. Does quality improvement improve quality? *Future Hosp J* 2016;3:191–4.
 - 17 Alexander JA, Heard LR. What can we learn from quality improvement research? A critical review of research methods. *Med Care Res Rev* 2009;66:235–71.
 - 18 Benn J, Burnett S, Parand A *et al*. Studying large-scale programmes to improve patient safety in whole care systems: challenges for research. *Soc Sci Med* 2009;69:1767–76.
 - 19 Holden RJ. Lean thinking in emergency departments: a critical review. *Ann Emerg Med* 2011;57:265–78.
 - 20 Taylor MJ, McNicholas C, Nicolay C *et al*. Systematic review of the application of the plando-study-act method to improve quality in healthcare. *BMJ Quality Saf* 2014;23:290–8.
 - 21 Martin G, Ozieranski P, Willars J *et al*. Walkrounds in practice: corrupting or enhancing a quality improvement intervention? A qualitative study. *Jt Comm J Qual Patient Saf* 2014;40:303–10.
 - 22 Scott I. What are the most effective strategies for improving quality and safety of health care? *Int Med J* 2009;39:389–400.
 - 23 Jones EL, Lees N, Martin G *et al*. How well is quality improvement described in the perioperative care literature? A systematic review. *Jt Comm J Qual Patient Saf* 2016;42:196–206.
 - 24 Woolf SH. The meaning of translational research and why it matters. *JAMA* 2008;299:211–13.
 - 25 Grady D, Redberg RF, O'Malley PG. Quality improvement for quality improvement studies. *JAMA Int Med* 2018;178:187.
 - 26 Basu S, Meghani A, Siddiqi A. Evaluating the Health Impact of Large-Scale Public Policy Changes: Classical and Novel Approaches. *Annual review of public health* 2017;38:351–70.
 - 27 Chen YF, Hemming K, Stevens AJ *et al*. Secular trends and evaluation of complex interventions: the rising tide phenomenon. *BMJ Qual Saf* 2016;25:303–10.
 - 28 Marshall M, Pronovost P, Dixon-Woods M. Promotion of improvement as a science. *Lancet* 2013;381:419–21.
 - 29 Nicolay CR, Purkayastha S, Greenhalgh A *et al*. Systematic review of the application of quality improvement methodologies from the manufacturing industry to surgical healthcare. *Br J Surg* 2012;99:324–35.
 - 30 Berwick DM. The science of improvement. *JAMA* 2008;299:1182–84.
 - 31 Benning A, Dixon-Woods M, Nwulu U *et al*. Multiple component patient safety intervention in English hospitals: controlled evaluation of second phase. *BMJ Clin Res Ed* 2011;342:d199.
 - 32 Benning A, Ghaleb M, Suokas A *et al*. Large scale organisational intervention to improve patient safety in four UK hospitals: mixed method evaluation. *BMJ Clin Res Ed* 2011;342:d195.
 - 33 Morgan DJ, Diekema DJ, Sepkowitz K *et al*. Adverse outcomes associated with Contact Precautions: a review of the literature. *Am J Infect Control* 2009;37:85–93.
 - 34 Westbrook JJ, Li L, Hooper TD *et al*. Effectiveness of a 'Do not interrupt' bundled intervention to reduce interruptions during medication administration: a cluster randomised controlled feasibility study. *BMJ Qual Saf* 2017;6:734–42.
 - 35 Power M. *Organized uncertainty: Designing a world of risk management*. Oxford: Oxford University Press, 2008.
 - 36 Leigh A. *Randomistas: how radical researchers are changing our world*. New Haven: Yale University Press, 2018.
 - 37 Allen D. From polyformacy to formacology. *BMJ Qual Saf* 2017;26:695–7.
 - 38 Sterman J, Oliva R, Linderman K *et al*. System dynamics perspectives and modelling opportunities for research in operations management. *J Oper Manag* 2015;39–40:1–5.
 - 39 Bindman AB, Pronovost PJ, Asch DA. Funding innovation in a learning health care system. *JAMA* 2018;319:119–20.
 - 40 Wachter RM, Pronovost PJ. The 100,000 lives campaign: a scientific and policy review. *Jt Comm J Qual Patient Saf* 2006;32:621–7.
 - 41 Brewster L, Aveling E-L, Martin G *et al*. What to expect when you're evaluating healthcare improvement: a concordat approach to managing collaboration and uncomfortable realities. *BMJ Qual Saf* 2015;24:318–24.
 - 42 Ivers N, Jamtvedt G, Flottorp S *et al*. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012;6:CD000259.
 - 43 Power M, Tyrrell PJ, Rudd AG *et al*. Did a quality improvement collaborative make stroke care better? A cluster randomized trial. *Implementation Sci* 2014;9:40.
 - 44 Martin GP, McKee L, Dixon-Woods M. Beyond metrics? Utilizing 'soft intelligence' for healthcare quality and safety. *Social Sci Med* 2015;142:19–26.
 - 45 Martin GP, Aveling E-L, Campbell A *et al*. Making soft intelligence hard: a multi-site qualitative study of challenges relating to voice about safety concerns. *BMJ Qual Saf* 2018;27:710–7.
 - 46 Shekelle PG, Pronovost PJ, Wachter RM *et al*. *Assessing the evidence for context-sensitive effectiveness and safety of patient safety practices: developing criteria*. Rockville, MD: Agency for Healthcare Research and Quality, 2010.
 - 47 Mitchell BG, Gardner A, Stone PW *et al*. Hospital staffing and health care-associated infections: a systematic review of the literature. *Jt Comm J Qual Patient Saf* 2018;44:613–22.
 - 48 Aiken LH, Sloane D, Griffiths P *et al*. Nursing skill mix in European hospitals: cross-sectional study of the association with mortality, patient ratings, and quality of care. *BMJ Qual Saf* 2017;26:559–68.
 - 49 Draycott T. *Practical obstetric multi-professional training – PROMPT*. London: The Health Foundation, 2013.
 - 50 Powell A, Rushmer RK, Davies H. *A systematic narrative review of quality improvement models in healthcare*. Edinburgh: NHS Quality Improvement Scotland, 2009.
 - 51 Vaughn VM, Saint S, Krein SL *et al*. Characteristics of healthcare organisations struggling to improve quality: results from a systematic review of qualitative studies. *BMJ Qual Saf* 2018, in press.
 - 52 Rutter H, Savona N, Glonti K *et al*. The need for a complex systems model of evidence for public health. *Lancet* 2017;390:2602–4.
 - 53 Wells M, Williams B, Treweek S *et al*. Intervention description is not enough: evidence from an in-depth multiple case study on the untold role and impact of context in randomised controlled trials of seven complex interventions. *Trials* 2012;13:95.
 - 54 Syverson C. What determines productivity? *J Econ Lit* 2011;49:326–65.
 - 55 Bloom N, Eifert B, Mahajan A *et al*. Does management matter? Evidence from India. *Quar J Econ* 2013;128:1–51.
 - 56 Kingdon JW, Thurber JA. *Agendas, alternatives, and public policies*. Boston: Little, Brown, 1984.
 - 57 Schneider JW. Social problems theory: the constructionist view. *Ann Rev Sociol* 1985;11:209–29.
 - 58 Stone DA. Causal stories and the formation of policy agendas. *Political Sci Quarterly* 1989;104:281–300.
 - 59 Pfohl S. Toward a sociological deconstruction of social problems. *Social Problems* 1985;32:228–32.
 - 60 D'Lima D, Archer S, Thibaut BI *et al*. A systematic review of patient safety in mental health: a protocol based on the inpatient setting. *Syst Rev* 2016;5:203.
 - 61 Boozary AS, Shojania KG. Pathology of poverty: the need for quality improvement efforts to address social determinants of health. *BMJ Qual Saf* 2018;27:421–4.

- 62 Cribb A. Improvement science meets improvement scholarship: reframing research for better healthcare. *Health Care Anal* 2018;2:109–23.
- 63 Carter SM, Williams J, Parker L *et al*. Screening for cervical, prostate, and breast cancer: interpreting the evidence. *Am J Prevent Med* 2015;49:274–85.
- 64 Carter SM. Valuing healthcare improvement: implicit norms, explicit normativity, and human agency. *Health Care Anal* 2018;26:189–205.
- 65 Draycott T, Sibanda T, Owen L *et al*. Does training in obstetric emergencies improve neonatal outcome? *BJOG* 2006;113:177–82.
- 66 Haynes AB, Weiser TG, Berry WR *et al*. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2009;360:491–9.
- 67 Neily J, Mills PD, Young-Xu Y *et al*. Association between implementation of a medical team training program and surgical mortality. *JAMA* 2010;304:1693–700.
- 68 Clarkson PJ, Bogle D, Dean J *et al*. *Engineering better care*. London: Royal Academy of Engineering, 2017.
- 69 Dean J. Systems engineering for healthcare improvement 101. *Future Hosp J* 2017;4:s19.
- 70 Sujan MA, Embrey D, Huang H. On the application of human reliability analysis in healthcare: opportunities and challenges. *Reliability Eng Syst Saf* 2018, in press.
- 71 Shorrock S. The loneliest profession in healthcare. *Humanistic Systems* 2018. <https://humanisticsystemscom/2018/05/11/the-loneliest-profession-in-healthcare>.
- 72 Batalden P. Getting more health from healthcare: quality improvement must acknowledge patient coproduction. *BMJ* 2018;362:k3617.
- 73 Dlugacz YD, Spath PL. High reliability and patient safety. In: Spath PL (ed). *Error reduction in health care: a systems approach to improving patient safety*. San Francisco: Jossey-Bass, 2011.
- 74 Aveling EL, Parker M, Dixon-Woods M. What is the role of individual accountability in patient safety? A multi-site ethnographic study. *Social Health Illness* 2015;38:216–32.
- 75 Thompson DF. Moral responsibility of public officials: The problem of many hands. *Am Polit Sci Rev* 1980;74:905–16.
- 76 Dixon-Woods M, Pronovost PJ. Patient safety and the problem of many hands. *BMJ Qual Saf* 2016;25:485–8.
- 77 Bowie P, Bradley NA, Rushmer R. Clinical audit and quality improvement – time for a rethink? *J Eval Clin Pract* 2012;18:42–8.
- 78 Evetts J. The sociological analysis of professionalism. *Int Sociol* 2003;18:395–415.
- 79 Yeung K, Dixon-Woods M. Design-based regulation and patient safety: a regulatory studies perspective. *Social Sci Med* 2010;71:502–9.
- 80 Parsons T. The professions and social structure. *Social Forces* 1939;17:457–67.
- 81 Martin G, Armstrong N, Dixon-Woods M *et al*. Professionalism redundant, reshaped, or reinvigorated? realizing the ‘third logic’ in contemporary healthcare. *J Health Social Behav* 2015; 56:378–97.
- 82 Black J. Constructing and contesting legitimacy and accountability in polycentric regulatory regimes. *Regul Governance* 2008;2:137–64.
- 83 Freidson E. *Professionalism, the third logic: on the practice of knowledge*. Chicago: University of Chicago Press, 2001.
- 84 Powell WW. Neither market nor hierarchy: Network forms of organisation. *Res Org Behaviour* 1990;12:295–336.
- 85 Granovetter M. Economic action and social structure: the problem of embeddedness. *Am J Sociol* 1985;91:481–510.
- 86 Offer A. Between the gift and the market: The economy of regard. *Econ Hist Rev* 1997;50:450–76.
- 87 Tarrant C, Leslie M, Bion J *et al*. A qualitative study of speaking out about patient safety concerns in intensive care units. *Social Sci Med* 2017;193(Supplement C):8–15.

Address for correspondence: Professor Mary Dixon-Woods, The Healthcare Improvement Studies Institute (THIS Institute), Clifford Allbutt Building, Cambridge Biomedical Campus, Cambridge CB2 0AH, UK.
thisinstitute.cam.ac.uk
Email: director@thisinstitute.cam.ac.uk