Elsevier Editorial System(tm) for The Lancet
Manuscript Draft

Manuscript Number:

Title: The science of improvement

Article Type: Viewpoint

Corresponding Author: Professor Martin Marshall,

Corresponding Author's Institution: University College London

First Author: Martin Marshall, MD

Order of Authors: Martin Marshall, MD; Martin Marshall; Peter Pronovost, PhD

The science of improvement

Martin Marshall MD (corresponding author)

Professor of Healthcare Improvement

University College London

3rd Floor, 170 Tottenham Court Road, London W1T 7HA

martin.marshall@islondon.org Tel: 020 3108 2335

Peter Pronovost PhD

Professor, Departments of Anaesthesiology, Critical Care Medicine and Surgery

The Johns Hopkins University School of Medicine

The Bloomberg School of Public Health, 1909 Thames Street, 2nd Floor, Baltimore, MD 21231

Mary Dixon Woods PhD

Professor of Medical Sociology

University of Leicester

Department of Health Sciences, Adrian Building, University Road, Leicester, LE1 7RH

Summary

Deficiencies in the organisation and delivery of care in all health systems that have been studied are now well recognised and recorded in the scientific literature. Rectifying these deficiencies has proved challenging. In this paper the authors argue that one reason that patients so often receive inadequate care is because efforts to improve services have been insufficiently informed by science. They propose that the science of improvement needs to be developed, promoted and embedded within health services through a partnership between clinicians, managers, academics and those who use health services. Using three examples of published improvement studies that have been subjected to rigorous evaluation, they describe the characteristics of this science and propose a strategic approach to optimise the impact of the science of improvement in the future.

(130 words)

Introduction

The remarkable achievements of modern health care make the deficiencies associated with the delivery of care all the more striking. Patients are routinely exposed to wide and inexplicable variations in quality of care, such as an eight-fold variation across England in the number of patients who receive angioplasty for severe myocardial infarction.[1] Healthcare is now a major source of harm, with about one in ten patients admitted to hospital suffering an adverse event, [2] with there is little evidence of improvement over time.[3] On average, about half of the clinical processes that would generally be accepted as good clinical practice are actually implemented. [4] Problems of variations in care, preventable patient harm and poor quality are common to all health systems, irrespective of how they are organised or financed. These problems have proved remarkably difficult to address, despite good intentions, ambitious programmes of improvement and the investment of significant resources. [5] In this article, we propose that one reason that patients are often failed by healthcare, in major and minor ways, is that improvement efforts have been insufficiently informed by science. We outline the characteristics of a science of improvement, and identify what needs to happen for healthcare to benefit from the application of such a science.

Different stances have been typically been adopted by those attempting to understand and address the challenge of improvement. Some, particularly at policy level, see the problem as primarily a managerial one, and frame the challenge pragmatically as one of ‘delivery’, ‘governance’ and having the right financial incentives in place.[6] Others adopt a more technical stance, focused on how scientific evidence should inform practice, and frame the challenge as one of ‘implementation’. [7] Yet others frame the challenge as one of “quality improvement”, borrowing from approaches used in manufacturing and other industries, and using specific techniques to try to engineer change. [8] Attention to this third position has gathered increasing momentum in recent years, so that a quality improvement "movement" is now said to exist.

This movement has made a valuable contribution, not least in challenging overly technocratic, managerialist, or regulatory-focused approaches to change. However, the urge to action can easily overwhelm the need for evidence, to the extent that much quality improvement work is *unscientific* in its orientation: it is neither based on high quality evidence nor subject to rigorous evaluation to determine its effectiveness, costs, and risks. The irony is that such work risks producing exactly the opposite of improvement: resources can be wasted, energy and enthusiasm dissipated, the side-effects of intervention ignored, and in the end little demonstrable positive change may be seen. The body of systematic, accessible and rigorous knowledge about best how to improve remains meagre. Adopting a more scientific approach to improvement has the potential to enhance the ability of health systems and those who work in them to provide high quality care and to use resources more efficiently and effectively.

Defining the science of improvement

As is often the case in what Kuhn[9] refers to as a 'pre-paradigm' phase of the emergence of a new discipline, no single agreed definition of the science of improvement exists. Attempts to produce one can provoke fierce debate, and to date the term has escaped consensus. [10] We suggest that *healthcare improvement science* has four distinguishing characteristics: (i) a body of content knowledge focused on improving the delivery of care to patients; (ii) a set of methods for investigating research questions relating to improvement of healthcare; (iii) a set of theories operating on different levels, including the "small theories" of individual programmes and the "big theories" of social phenomena (iv) a set of values that underpin the ethical commitments of the science.

The science of improvement draws on a rich heritage of contributing disciplines. As a consequence, it utilises a range of different theories, methodologies and research methods. Using three examples, we explore the value of treating improvement as a science, and use these examples to identify some of the characteristics of good improvement science.

Example 1: Reducing central line infections in Intensive Care Units

The first example (Box A) aimed to reduce central line infections in intensive care units (ICUs) in the United States. Such infections represent a significant clinical problem with often fatal consequences, and are very costly to treat. A cohort study conducted in over 100 ICUs in the state of Michigan in 2004/5 showed that a multi-component programme comprising evidence-based technical interventions, adaptive interventions targeted on culture and systems, and a centralised data collection and feedback system resulted in a large reduction in rates of catheter-related bloodstream infections. A later study found that these improvements were sustained in participating ICUs compared with controls. Follow up work undertaken by members of original study, in partnership with social scientists, explored the mechanisms through which the programme worked, and generated a theory of change that could inform and be tested in subsequent iterations of the programme.

[Insert Box 1]

Example 2: Increasing the identification and referral of victims of domestic violence

The second example (Box B) aimed to increase, in UK primary care, rates of identification and referral to specialist advocacy services of women subjected to domestic violence. Again, the study addressed a common and important challenge for health systems, and one that is often inadequately managed despite the major health and social implications of unidentified and unmanaged abuse. A cluster randomised controlled trial was undertaken in 51 primary care clinics in 2 large cities in the UK in 2007/8, utilising a range of evidence-based interventions. It showed an increase in recorded disclosures of domestic violence and in referral to specialist services in the intervention group.

[Insert Box 2]

Example 3: Reducing door to balloon time for patients with myocardial infarction

The third example (Box C) aimed to reduce "door to balloon" time for patients with ST segment elevation myocardial infarction (STEMI) in hospitals in the United States. There was good evidence at the outset that prompt reperfusion treatment reduces the mortality and morbidity of STEMI patients ("time is muscle"). However, few hospitals met the standard of 90 minutes or less from entering the emergency department (door) to receiving treatment in the cardiac catheterisation laboratory (balloon). A mixed method approach involving in-depth interviews, surveys and modelling techniques identified and quantified the most effective strategies and helped improve the proportion of patients getting to "balloon" on time from 50% to 75%.

[Insert Box 3]

More than anything, perhaps, these studies underline the extent to which improvement interventions and their causal mechanisms are deeply and inescapably complex. As Lipsey describes: *"They involve multidimensional interactions that are often extended over time, complex multistep causal processes in which different individuals may react differently, and uncertain and potentially wide-ranging outcomes, not all necessarily desirable"*.[11] Such complex interventions need a sophisticated science to match.

The three examples demonstrate some of the characteristics of that science. They all contribute to the development of *content* knowledge: it is now much clearer “what works” in reducing infection in central lines, improving the support of women at risk of domestic abuse, and enhancing the chances of a successful outcome for patients with a severe myocardial infarction. They also make use of a set of *evaluative methods* that are well established and enjoy high credibility among practising clinicians and decision-makers, but are adapted for the specifics of the area in which they are applied. In their own way, each represents a methodological innovation. These three studies also make a contribution to the *theoretical development* of the field, by helping to reveal the mechanisms through which improvement interventions work. Finally, all three demonstrate a genuine *ethical commitment* to both patient benefit and to learning.[12]

Characteristics of a science of improvement

Analysis of these examples highlights four characteristics of a science of improvement in healthcare. None is unique to this science, but together they help to explain what it is about:

Improvement science attempts to generate practical learning that can make a timely difference to patient care. Each of the examples we discuss has resulted in practical learning to help health services to improve patient outcomes. Improvement science focuses on identifying and helping to solve problems concerned with quality or safety of care. It is therefore characterised by its substantive field of interest, by its applied nature, and by its commitment to learning that can be applied in practice. In this, it recognises and integrates multiple contributions, much as engineering science uses scientific knowledge and theories to address real-life problems. This helps to address the challenge that those making decisions about patients do not always see traditional health services research as “useful”.

Improvement science aims to produce generalisable or transferable knowledge, ***and it utilises robust, well established research methods applied in highly pragmatic ways.*** For each of the programmes we describe, the improvement activities were conducted in a way that enabled local improvement but produced knowledge with external validity, which was then published as an enduring historical record in well-respected journals that ensured both critical peer review and wide dissemination**.** Developing local knowledge based on a strong sense of ownership and a willingness to adapt improvement activities is clearly important. But wherever possible, there is an obligation, based on both moral and practical imperatives, to learn systematically, and to produce generalisable or transferable knowledge about how to improve.[12]However, many quality improvement projects at present rely on contemporaneous, non-standardised, non-verified data to make judgements about their effectiveness. In so doing, they may be prone to bias, and they frequently demonstrate effect sizes that prove hard to replicate in more objective evaluations. In contrast, a science of improvement is characterised by its commitment to rigorous evaluative methods and high quality data collection and interpretation. The choice of methods in improvement science is, however, often guided by pragmatism and in particular by the formidable realities of intervening and collecting data in messy, highly heterogeneous real-life clinical situations outside of the clean world of the laboratory or the randomised controlled trial of a medicinal product. Thus, the decision to use a cohort design in the central line infection study rather than an RCT reflected a recognition that randomisation would not have been acceptable to those involved, who primarily saw themselves as participants in an improvement rather than a research project. It might also have been useful to have known which elements of the complex intervention were most effective, but this would have greatly increased the burden of data collection for the teams, and posed a risk to subject recruitment and retention that the authors did not consider worth taking. The door-to-balloon studies showed how combining multiple data sources – qualitative and quantitative – can provide clear, practical guidance on where organisations need to target their efforts.

Improvement science requires a genuine partnership between academics and front line practitioners. Researchers bring scepticism, scientific rigour and methodological technical expertise. Practitioners bring content knowledge and a deep understanding of working contexts. Both are needed for improvement science, as our three examples illustrate. For the central venous catheter work, the study was initiated by a doctor and a nurse, working with health service researchers, statisticians, and others. Additional insights into the mechanisms of action were provided subsequently through a partnership with social scientists. The domestic violence study involved clinicians, community social workers, statisticians and economists. The door-to-balloon study involved clinicians, public health specialists, epidemiologists, social scientists and modellers. In all cases, the academic and the service partners were both actively involved in designing, undertaking and interpreting the work, creating a synergy across boundaries that moves the learning beyond that gained from traditional research or quality improvement work. Partnership between researchers and practitioners, or between different disciplines, is not a passive or rigidly compartmentalised one, but rather one of mutual support, respect and healthy challenge. Further, they are partnerships between equals, not one where one discipline or group is seen as subordinate or the servant of the other.

Improvement science draws on, and aims to contribute to, clear and explicit theories of how change happens*.* A major activity of improvement science centres on the design, deployment, and evaluation of complex, multi-faceted interventions. Improvement efforts should be based on good theory, but too often interventions are designed without the benefit of previous learning or explicating the assumptions about how and why change is likely to occur. There is a rapidly developing evidence base to guide effective improvement practice, describing what works, to what extent and the unintended consequences. Too often improvement projects ignore this evidence base and are less effective as a consequence. As a result, many quality improvement interventions remain black boxes that are difficult to reproduce in new contexts. It is important that the underlying theories, or hypothesised mechanisms of change, are surfaced for each intervention or programme. In each of the three examples, the interventions were chosen on the basis of published evidence and each of the studies built on this evidence in an incremental way, but also allowed room for innovation. For the door-to-balloon, study the authors hypothesised that clear managerial strategies and specific practical guidance would change the behaviour of those planning care for acute myocardial infarction. The authors of the domestic violence study placed a strong emphasis on advocacy-led education and training and feedback as a vehicle for change. The central line infection team hypothesised that multiple interventions at all levels of the system were required and that a delicate balance between facilitative and coercive approaches was needed. The central line study also demonstrates the benefits of a flexible and evolutionary approach to theory development.

Discussion

We have attempted to describe the science of improvement in healthcare and make the case for its relevance to those responsible for improving the quality of care provided for patients. Science-guided improvement has the potential to more effectively address the gap between what is done and what is known should be done, to identify what should be done where that is unknown, and to accelerate progress by prioritising systematic learning as a feature of health services. [13] Front-line practitioners deserve high quality evidence on which to base their efforts to serve patients; good intentions are not enough. By integrating science with clinical priorities, by bringing together those who do the work of caring for patients with those who study that work, and by committing all those involved to the same values, the science of improvement is rich in its potential to benefit patients.

A strategic approach is required to optimise the impact of improvement science. First, the policy, service and academic communities need to learn about improvement science, the synergistic relationship between this science and the more familiar biomedical and clinical sciences, and how the science can be used to make a difference to patient care. This will likely require academia to break out of rigid and unhelpful departmental silos and work in interprofessional research teams, and for incentives to be changed to encourage this. It is clear that no single academic discipline can address the needs of improvement science: the list of those who contribute includes (but is not limited to) operations research and industrial engineering, clinical science, health and behavioural economics, management studies, sociology and anthropology, psychology, statistics and mathematics, epidemiology, policy analysis, philosophy and ethics and human factors and systems engineering. A growing number of institutions around the world are demonstrating the benefits of such wider collaboration.

Second, there is a need to build capacity and capability in the field at all levels from improvement science leaders, such as the recently launched Health Foundation Improvement Science Fellowship scheme in the UK and similar schemes which are being planned in Sweden and the Netherlands. Third, the counter-cultural nature of improvement science for traditional service and academic sectors needs to be recognised and space provided to develop new ways of thinking and acting. This will require a redefinition of some of the success criteria in both the academic sector (where the desire to make a difference is rarely rewarded) and the service sector (where the development and application of scientific evidence is not always highly regarded). Some leading institutions are starting to recognise the need for change. For example, Johns Hopkins University in the USA has recently developed promotion criteria for faculty doing improvement work, based on measurable improvements for patients and the development of transferable learning.

Fourth, the creation of an international network or body to support the emerging science and assist improvement scientists in sharing learning would help to ensure that the development of the science is an outward looking endeavour. Finally, funding is required to establish and support centres for improvement science, bringing together diverse academic and clinical disciplines within discrete health economies and to promote high quality improvement projects and capacity building activities.

If this strategy is successful, within the next decade the science of improvement will be embraced as a mainstream body of knowledge and set of activities that will be an essential component of providing high quality care for patients.

2821 words

No funding was received to write this paper

MM had the original idea for the paper and all authors developed the content together. MM wrote the first draft of the paper and PP and MDW adapted and revised this draft. MM is the guarantor.

The author statement form has been completed and will be faxed separately.

Box 1

Pronovost et al NEJM (2006) [14] A cohort study conducted in 103 ICUs in the state of Michigan, this programme aimed to reduce catheter-related blood stream infections using a complex intervention with some evidence-based and some innovative components. The median rate of infections per 1000 catheter days decreased from 2.7 infections at baseline to 0 at 3 months after the implementation of the intervention and the mean rate per 1000 catheter days decreased from 7.7 at baseline to 1.4 at 16 to 18 months of follow-up (a 66% reduction).

Lipitz-Snyderman et al BMJ (2010)[15] Using a retrospective, observational study design, this study compared mortality in patients aged over 65 admitted to ICUs 95 Michigan hospitals with a control group of 364 hospitals in 11 surrounding states. It found significantly decreased mortality in Michigan hospitals compared with controls: the chance of dying in an ICU reduced by 24% in Michigan compared with a reduction of 16% in other mid-western hospitals.

Dixon-Woods et al Milbank Quarterly (2011)[16] This article attempted to explain the success of the Michigan programme using a combination of social science theory and the accounts of the programme designers. It suggested that important factors included the redefinition of central line infections as a social problem requiring concerted and collective action rather than a simple technical problem; the role of peer pressure and importance of social networks in facilitating change; the judicious use of “hard edges” or managerial approaches to changing behaviour, alongside more facilitative profession approaches; and the creation of a clinical community focused on solving the problem.

Box 2

Feder et al, The Lancet, 2011 [17] A cluster-randomised trial conducted in 48 primary care practices in the UK that aimed to improve the identification and referral of women experiencing domestic violence. The trial tested a complex, multifaceted intervention including multidisciplinary practice-based training sessions with significant lay input, prompts in the medical record system, continuous a pathway that could be offered to women to access resources, cards and posters in surgeries, continuous reinforcement and informal contact, and feedback on referrals. 1 year after the intervention, a dramatic increase in rates of referral to advocacy services in the intervention practices was observed (adjusted intervention rate ratio of 22.1, 95% confidence interval 11.5-42.4), together with a smaller increase in recorded disclosures of incidents of violence (adjusted intervention rate ratio 3.1, 95% confidence intervals 2.2-4.3).

Box 3

McNamara et al, Journal of the American College of Cardiology 2006 [18]

Using a cohort study design of 29,222 patients, this study established that longer door-to-balloon time was associated with increased in-hospital mortality and is relevant regardless of time from symptom onset and baseline risk of mortality.

Bradley et al, NEJM 2006 [19]

This study sought to identify and quantify operational and clinical processes to reduce door-to-balloon times. Using a survey methodology and modelling techniques it found six strategies were positively associated with faster door to balloon times: emergency department doctors activating the catheterization laboratory; having a single call to a central page to mobilise the laboratory; activating the laboratory while the patient was en route to hospital; expecting staff to arrive in the laboratory within 20 minutes of being paged; having a consultant (attending) physician permanently on site; and real-time data feedback to staff in the emergency department and the catheter lab. The temporal impact of each of these strategies were calculated and using four of them were found to result in a door to balloon time of 79 minutes in comparison with 110 minutes when none of them were used.

Bradley et al, Circulation, 2006 [20]

A qualitative study using depth interviews with 122 hospital staff at 11 hospitals that had median door-to-balloon times of 90 minutes or less, this work showed that setting explicit goals, having visible support of senior management and uncompromising clinical leaders, standardised clinical protocols but flexibility in implementation, collaborative interdisciplinary teams, data feedback to monitor progress and identify successes or problems, and positive organizational culture were all implicated in shorter transits between door and catheterization laboratory.

Bradley et al, Journal of American Journal of Cardiology, 2009 [21]

This study showed, using a longitudinal study of 831 hospitals, that participation in a national programme (the D2B Alliance) that recommended use of the strategies identified by the earlier studies was associated with marked improvement in practice and performance in the delivery of PPCI for patients with STEMI. By March 2008, more than 75% of patients had D2B times of 90 minutes or less, compared with only 50% of patients within 90 minutes in April 2005.

References

1. QIPP/Right Care. NHS Atlas of Variation in Healthcare (www.rightcare.nhs.uk). 2011; .

2. de Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. *Qual Saf Health Care* 2008; 17(3): 216-23.

3. Landrigan CP, Parry GJ, Bones CB, Hackbarth AD, Goldmann DA, Sharek PJ. Temporal trends in rates of patient harm resulting from medical care. *N Engl J Med* 2010; 363(22): 2124-34.

4. McGlynn EA, Asch SM, Adams J, Keesey J, Hicks J, DeCristofaro A, et al. The quality of health care delivered to adults in the United States. *N Engl J Med* 2003; 348(26): 2635-45.

5. Wachter RM. Patient safety at ten: unmistakable progress, troubling gaps. *Health Aff (Millwood)* 2010; 29(1): 165-73.

6. Department of Health. Equity and excellence: liberating the NHS. 2010; .

7. Eccles MP, Armstrong D, Baker R, Cleary K, Davies H, Davies S, et al. An implementation research agenda. *Implementation Science* 2009; 4(1).

8. Varkey P, Reller MK, Resar RK. Basics of quality improvement in health care. *Mayo Clin Proc* 2007; 82(6): 735-9.

9. Kuhn TS. The structure of scientific revolutions. University of Chicago Press; 1962.

10. The Health Foundation. Improvement Science Research Scan. 2011; /www.health.org.uk/publications/improvement-science/

11. Lipsey M. Theory as method: small theories of treatments. In: Sechrest P, Perrin E, Bunker J, editors. Research methodology: strengthening causal interpretations of non-experimental data. Washington, DC: Agency for Healthcare Policy and Research; 1990.

12. Faden RR, Beauchamp TL, Kass NE. Learning health care systems and justice. *Hastings Cent Rep* 2011; 41(4): 3.

13. Shojania KG, Grimshaw JM. Evidence-based quality improvement: The state of the science. *Health Aff* 2005; 24(1): 138-50.

14. Pronovost P, Needham D, Berenholtz S, Sinopoli D, Chu H, Cosgrove S, et al. An intervention to decrease catheter-related bloodstream infections in the ICU. *N Engl J Med* 2006; 355(26): 2725-32.

15. Lipitz-Snyderman A, Steinwachs D, Needham DM, Colantuoni E, Morlock LL, Pronovost PJ. Impact of a statewide intensive care unit quality improvement initiative on hospital mortality and length of stay: retrospective comparative analysis. *BMJ* 2011; 342: d219.

16. Dixon-Woods M, Bosk CL, Aveling EL, Goeschel CA, Pronovost PJ. Explaining michigan: developing an ex post theory of a quality improvement program. *Milbank Q* 2011; 89(2): 167-205.

17. Feder G, Davies RA, Baird K, Dunne D, Eldridge S, Griffiths C, et al. Identification and Referral to Improve Safety (IRIS) of women experiencing domestic violence with a primary care training and support programme: a cluster randomised controlled trial. *Lancet* 2011; 378(9805): 1788-95.

18. McNamara RL, Wang Y, Herrin J, Curtis JP, Bradley EH, Magid DJ, et al. Effect of door-to-balloon time on mortality in patients with ST-segment elevation myocardial infarction. *J Am Coll Cardiol* 2006; 47(11): 2180-6.

19. Bradley EH, Herrin J, Wang Y, Barton BA, Webster TR, Mattera JA, et al. Strategies for reducing the door-to-balloon time in acute myocardial infarction. *N Engl J Med* 2006; 355(22): 2308-20.

20. Bradley EH, Curry LA, Webster TR, Mattera JA, Roumanis SA, Radford MJ, et al. Achieving Rapid Door-To-Balloon Times: How Top Hospitals Improve Complex Clinical Systems. *Circulation* 2006; 113(8): 1079-85.

21. Bradley EH, Nallamothu BK, Herrin J, Ting HH, Stern AF, Nembhard IM, et al. National efforts to improve door-to-balloon time results from the Door-to-Balloon Alliance. *J Am Coll Cardiol* 2009; 54(25): 2423-9.