The Health Foundation
Inspiring Improvement

What is the science of improvement?

Mary Dixon-Woods, Peter Pronovost and Martin Marshall

May 2013

Objectives

- Explain the key principles of the science of improvement and its application to safety
- Identify recent applications of the science of improvement to large scale safety interventions
- Identify approaches to evaluating patient safety improvement projects

Key principles of the science of improvement

The need to secure improvement in healthcare is clear: patients remain at risk of preventable harm, and care often falls short of what the evidence shows to be best practice. Equally clear is the need for systematic study of how improvement can best be secured.[1] Yet, ironically, improvement research is a field beset by quality problems, including widespread use of study designs that limit the confidence with which change can reliably be attributed to the intervention, use of poorly operationalised measures, poor quality of data collection, limited use of theory, reluctance to search for unintended consequences or determine the cost-effectiveness of interventions, and making inferences about improvements that are overly optimistic (or can not be supported by design and data).[2,3]

Studies of quality improvement are often remarkably poor in describing exactly what the intervention comprises, and often fail to characterise the intervention and its activities in such a way that it can easily be reproduced. Such studies are equally poor at describing the theoretical basis of their interventions (what is the means by which this intervention might reasonably be expected to achieve the hoped-for effects?) and of their evaluation; attempts to update theories in response to the findings of empirical studies also remain rare in quality improvement, so that theory evolution remains stunted.[4] The field has only recently been open to self-reflection; for too long the benefits of improvement efforts were assumed to be self-evident. Enthusiasm for well-meaning interventions often continues to exceed the available evidence; a science approach to improvement is vital in moving the field forwards for the benefit of patients.

The field lacks a single, widely agreed upon definition for the science of improvement. Attempts to produce one can provoke fierce debate, and to date the term has escaped consensus.[5] This is perhaps not surprising, as the field is inherently interdisciplinary, relatively new, and advancing rapidly. At its most fundamental, however, the science is best characterised by its substantive field of interest, by its applied nature, and by commitment to learning that can be applied in practice. We propose a number of principles on which it might be founded:

1 It is guided by a strong set of values based on a commitment to improving the quality and safety of care for patients and the science of improvement

2 It is distinguished by a body of *content* or *'how to' knowledge* (formal and more tacit) that can make a timely difference to patient care

3 It uses high quality, established methods from across the medical and social sciences to investigate research questions relating to improvement of healthcare, and where appropriate also contributes to methodological innovation

4 As applied research, it generally seeks to solve problems rather than puzzles

5 It draws on theories ranging from the 'small theories' of individual programmes and interventions to the 'big theories' operating at a more abstract level

6 It requires authentic partnerships between researchers and those involved in regulating, managing and delivering care.

Applications of the science of improvement to large scale safety interventions

Recent examples of published studies provide vivid demonstrations of just how great a difference the science of improvement can make:

Reducing central line infections

Central line infections in intensive care units (ICUs) 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 US state of Michigan 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 could result in a large and sustained reduction in rates of catheter-related bloodstream infections and mortality.[6] A later study found that these improvements were sustained in participating ICUs compared with controls.[7] Follow-up work undertaken by members of the 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.[4]

Increasing identification and referral of victims of domestic violence

A cluster randomised controlled trial of a complex multifaceted intervention to improve identification and referral of women experiencing domestic violence was undertaken in 48 primary care clinics in two large cities in the UK.[8] One year after the intervention, a dramatic increase in rates of referral to advocacy services in the intervention practices was observed, together with a smaller increase in recorded disclosures of incidents of violence. This evidence is now being used to guide the decisions of people commissioning services in over 20 locations across the UK.

Reducing door to balloon time for patients with myocardial infarction

A cohort study of 29,222 patients established that longer time taken to get from the door of the emergency department to cardiac catheterization 'balloon' was associated with increased in-hospital mortality.[9] A qualitative study using interviews with 122 hospital staff at 11 high performing hospitals was able to identify the features of organisational settings that were implicated in shorter transits between door and catheterisation laboratory.[10] A mixed-method study involving in-depth interviews, surveys and modeling techniques identified and quantified the six most effective strategies for improving the proportion of patients getting to "balloon" on time from 50% to 75%.[11] Using four of these strategies resulted in a door to balloon

time of 79 minutes in comparison with 110 minutes when none were used.[11] A later longitudinal study[12] of 831 hospitals reported 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 primary percutaneous coronary interventions (PPCI) for patients with ST-elevation myocardial infarction (STEMI). By March 2008, more than 75% of patients had door-to-balloon times of 90 minutes or less, compared with only 50% of patients within 90 minutes in April 2005. However, in a rare investigation of unintended consequences of intervention, these efforts were also associated with increased risk of false positives for cardiac catherisation.

These three examples demonstrate some of the characteristics of the science of improvement. First, they all contribute to the development of *content* knowledge: making it 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 both local improvement occurs and large scale improvement interventions work. Finally, all three demonstrate a genuine *ethical commitment*, both to patient benefit and to learning.[13]

Evaluating patient safety improvement projects

A major activity of the science of improvement centres on the design, deployment, and evaluation of complex, multi-faceted interventions, often involving a number of different stakeholders. The studies use the whole range of study designs from the evaluative and clinical sciences: trials, observational studies, qualitative studies, and so on. Yet many quality improvement projects at present rely on single time-period pre- and post-design without concurrent controls, and non-standardised, non-verified data to make judgements about their effectiveness.[14] In so doing, they may be prone to bias, and they frequently demonstrate effect sizes that prove hard to replicate in more objective evaluations. When controlled study designs are used, with independent data collection, it is often more difficult to demonstrate that any improvement effects can confidently be attributed to the interventions studied (rather than secular trends, for example).[15,16] Many quality improvement interventions remain black boxes that are so poorly described they are difficult to reproduce in new contexts.[4] In contrast, a science of improvement is characterised by its commitment to rigorous evaluative methods and high quality data collection and interpretation, and by its attempt to explicate underlying mechanisms of change (or 'small theories'[17] of programmes) with practical intent.

The door-to-balloon studies show how combining multiple data sources – qualitative and quantitative – in a theoretically informed way can provide clear, practical guidance on where organisations need to target their efforts. 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. Moreover, unlike clinical studies of drugs or devices that provide support for data collection at participating sites, improvement studies generally rely on voluntary data collection by clinicians, many of whom lack training in data quality. In the Michigan ICU study, for example, the decision to use a cohort design rather than a more robust cluster randomised trial design reflected a recognition that randomisation would not have been acceptable to those involved, who primarily saw themselves as participants in an improvement project 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. This reflects how the goal of improvement work is generally to optimise by using multifaceted interventions rather than to test the individual impact of a single intervention.

Surfacing the theories, or hypothesised mechanisms of change, for each intervention or programme is an important task for the science of improvement. In each of the three examples above, 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.

Finding new ways to evaluate studies

Each of the three studies presented represents an example of large scale evaluation. Neither the funding nor the academic capacity exist to evaluate all improvement projects in this way. This leaves many improvement projects which are not evaluated in any way that would be recognisable to the research community. The consequence is that we have little unbiased understanding of whether or not they make a difference, and, if they do, then how. In addition, the opportunity for creating sharable learning for others is lost. One of the challenges for pragmatically-oriented improvement scientists, therefore, is to design new approaches to evaluation which are rigorous but can be used at smaller scale, and which engage service teams in self-evaluation. Examples of such approaches include embedding researchers within service delivery teams to advise on how best to interpret research evidence and to design evaluations,[18] or building rapid response evaluation teams comprising mixed evaluation expertise to undertake focused process oriented evaluations at low cost and over a relatively short timescale. Such approaches should not compromise on scientific integrity but can help to define a different kind of relationship between decision-makers in the health services and academics.

Finding new ways of studying improvement will also find new ways of organising science. It is now evident that no single academic discipline alone can provide

the theories and the evaluative techniques needed for this science: the list of those who might 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. Part of the challenge of the science of improvement will be building collaborations between these disciplines, but also ensuring that the science remains rooted in the real-world concerns of patients and those at the sharp end who care for them. Diverse teams, which can be extremely innovative and productive, are the source of further challenges: they are difficult to create, manage, and sustain. The time and skill required to manage relationships increases exponentially with the number of different disciplines. The prize, however, may be great: the prospect of securing better care for patients.

References

1 Shekelle PG, Pronovost PJ, Wachter RM, Taylor SL, Dy SM, Foy R, et al. Advancing the science of patient safety. *Ann Intern Med* 2011, May 17, 154 (10): 693-6.

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

3 Ting HH, Shojania KG, Montori VM, Bradley EH. Quality Improvement: Science and Action. *Circulation* 2009, April 14, 119 (14), 1962-74.

4 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, Jun, 89 (2),167-205.

5 Marshall M, Pronovost P, Dixon-Woods M. Promotion of improvement as a science. *Lancet* 2013, Feb 2, 381 (9864), 419-21.

6 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.

7 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, Jan 28, 342, d219.

8 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, Nov 19, 378 (9805), 1788-95.

9 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, Jun 6, 47 (11), 2180-6.

10 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, Feb 28, 113 (8), 1079-85.

11 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, Nov 30, 355 (22), 2308-20.

12 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, Dec 15, 54 (25), 2423-9.

13 Faden RR, Beauchamp TL, Kass NE. Learning health care systems and justice. *Hastings Cent Rep* 2011, Jul-Aug, 41 (4), 3.

14 Pham JC, Frick KD, Pronovost PJ. Why Don't We Know Whether Care Is Safe? *Am J Med Qual* 2013, Mar 24.

15 Benning A, Dixon-Woods M, Nwulu U, Ghaleb M, Dawson J, Barber N, et al. Multiple component patient safety intervention in English hospitals: controlled evaluation of second phase. *BMJ* 2011, Feb 3, 342, d199.

16 Benning A, Ghaleb M, Suokas A, Dixon-Woods M, Dawson J, Barber N, et al. Large scale organisational intervention to improve patient safety in four UK hospitals: mixed method evaluation. *BMJ* 2011, Feb 3, 342, d195.

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

18 Van de Ven AH. *Engaged scholarship: a guide for organizational and social research.* Oxford: Oxford University Press; 2007.