EVIDENCE FROM HEALTHCARE MODELING:
WHAT IS ITS NATURE, AND HOW SHOULD IT BE USED?

Sally Brailsford
Jonathan H. Klein
Terry Young

Southampton Business School
University of Southampton
Southampton, SO17 1BJ, UK

Department of Computer Science
Brunel University London
Kingston Lane, Uxbridge, UB8 3PH, UK

ABSTRACT

This paper summarizes an event which brought together modeling experts, users, potential users and healthcare management practitioners to explore the question of evidence in the field of healthcare: how it is, and how it should be, used. The paper is an abridged version of a full report on the event. The paper highlights the high status accorded to empirical evidence (generated, in particular by randomized controlled trials), and suggests that a more balanced view of evidence, as being composed of empiricist, rationalist, and historicist material, is of value. Modeling and simulation constitute sources of rationalist evidence. Tensions between different types of evidence are identified, and the tension between statistics and stories as evidential forms is explored. The paper concludes on an optimistic note: there are clear signs that within the UK National Health Service, modeling and simulation are being taken more seriously as sources of evidence.

1 INTRODUCTION

The Cumberland Initiative (CI) is a UK collaborative of universities, National Health Service (NHS) organizations and commercial organizations founded in 2010 with the aim of using simulation, modeling and systems thinking to stimulate changes in NHS care delivery which will deliver radically improved quality and cost. One of its key aims in realizing its ambitions is to develop an understanding of how reliable judgements may be made concerning what really works in healthcare: an understanding of the nature and role of evidence in healthcare. Currently the dominant evidential paradigm in healthcare is that of the randomized controlled trial (RCT), often referred to as the gold standard of evidence provision. This dominance, which reflects the medical perspective in healthcare, is clearly embodied in the evidence-based medicine (EBM) movement, as well as the broader evidence-based practice (EBP) movement which developed from EBM. Within the field of medicine, the value of this empiricist spirit is unquestionable. However, in considering the broader issue of healthcare, such as quality of care delivery or patient experience, the ideal of a meaningful RCT study is generally unattainable in practice (as, indeed, it may sometimes be even within medicine), and the issue of how evidence might be generated and assessed to inform sound policy decisions and operational changes to practice is huge and unresolved.

To begin to address the issue of evidence, the CI convened a one-week Festival of Evidence (FoE) conference in October 2014, at Runnymede in Berkshire. The purpose of the conference was to explore the ways in which systems thinking, modeling, operational research (OR) and other tools are used (and might be used) to inform decision-making in healthcare delivery, and to bring together people who have access to evidence from such tools with people who might need to know of it, to facilitate the commissioning of improved services and more effective research. Though the range of tools included in the scope of the conference was deliberately broad, simulation modeling was a dominant context for much of the week. The programme included presentations by several keynote speakers: Mike Carter
A detailed account of the Festival of Evidence, which also includes conceptual and literature-based material not explicitly addressed during the conference, is, at the time of writing of this paper, in preparation (Brailsford and Klein 2015). This paper is, in essence, an abridged version of that report which highlights its main arguments and conclusions.

2 THE NATURE OF EVIDENCE

Within healthcare, evidence is material which supports a proposition or belief concerning healthcare. Cartwright and Hardie (2012) characterize the problem of using evidence as one of getting from knowing, of a particular action or policy, that “it worked there” to having good reason to believe that “it will work here” (p. 14): that a cause-effect relationship is transferable from “there” to “here”. Their three-legged stool metaphor argues that evidence for a proposed cause-effect relationship has three essential components: (1) a demonstration that the relationship works somewhere; (2) an argument to support the proposition that it is valid here too; and (3) an indication that the conditions necessary for the relationship to produce the effect are contextually present here.

Many authorities (see, for example, Evans 2003) emphasize the importance of both internal validity (the definitiveness with which a study establishes a cause-effect relationship between action and outcome in the particular context under study) and external validity (the generalizability of the study beyond the immediate case). Frequently, studies exhibit a trade-off between internal and external validity. A well-conducted RCT is likely to be of high internal validity, demonstrating (but not explaining) a cause-effect relationship between under the particular conditions in which the study has been carried out, but of low external validity, because the existence of the cause-effect relationship would be suspect under different conditions. In Cartwright and Hardie’s (2012) terms, internal and external validity may be associated with, respectively, the trustworthiness of evidence and its relevance. They argue that in many evidence-assessment schemes, trustworthiness tends to be emphasized at the expense of relevance.

Many schemes have been formulated, often in the context of evidence-based practice (EBP), for rating the value of methods of generating evidence (see, for example: Carter 2010; Balshem et al. 2011; Evans 2002). They tend to be hierarchical, with RCTs at or near the top, and other methods lower down. Expert opinion is often placed at or near the bottom of such schemes. Cartwright and Hardie’s (2012) general observation concerning the guidance provided by such schemes is that while they are “actually good at identifying policies that work, that is, policies that work somewhere”, they fail to identify “what works here” (p. 137). We observe, additionally, that they place emphasis on the desirability of empirical results as a source of evidence rather than, say, rational argument, or anecdotal material.

In any case, not all stakeholders may respond in the same way to empirical evidence. While such evidence clearly has strong appeal for clinical investigators, other stakeholders may be more convinced or persuaded by other forms of evidence (Lewith et al. 2002). Walshe and Rundall (2001) point to the difference in cultural assumptions between clinicians and healthcare managers, observing the former to be “highly professionalized, with a formal body of knowledge that is shared by all members of the profession”, while the latter “are a highly diverse group drawn from different professional and disciplinary backgrounds” and “personal experience and self-generated knowledge play a much larger part in determining how managers approach their jobs” (p. 439).

Cultural assumptions may also account for features of the way in which the results of studies are reported. The difference between the statistical significance of an effect and its strength is often insufficiently acknowledged, and, more generally, research results are often presented in ways which undermine the ease with which their potential application within candidate situations can be appreciated (Wye et al. 2015).
We would also observe that the intrinsic quality of studies is variable. Often this is unavoidable due to the nature of the study: for example, it is sometimes not possible to blind effectively within an RCT. (Lack of blinding, according to Carter 2010, can lead to the overestimation of treatment effects of the order of 15%.)

The conclusion we draw is that, within the healthcare domain, the dominant evidential paradigm is that of evidence-based practice, and that it is strongly weighted in terms of empirical evidence, in particular as generated by RCTs. Ultimately, the paradigm argues, such evidence requires nothing additional (for example, a theoretical understanding of the cause-effect relationship under scrutiny) to make its case. This is fine, as long as the evidence is trustworthy, relevant, and persuasive in the proposed application area. But if, as we contend is frequently the case, these conditions are not met, then the power of pure empirical evidence is undermined. A broader perspective is required, accepting of more diverse sources and types of evidence.

3 THE EVIDENTIAL TRIANGLE

We have found it useful to develop the concept of the evidential triangle. Following Hjorland’s (2011) terminology, the three corners of this equilateral triangle represent evidence derived within the traditions, respectively, of empiricism, rationalism, and historicism. Empiricism privileges empirical data gathered purposefully and systematically for the express purpose of providing evidence - for example, it might be data generated by RCTs. Rationalism invokes theoretical understanding (for example, a theoretical account of cause and effect). Historicism emphasizes the social context of knowledge (in the form of anecdote, for example). Our argument, in short, is that the case for the effectiveness of any particular action or policy is composed of one or (hopefully) more pieces of evidence that lie in various positions within the triangle. In order to build a case for a policy that is simultaneously trustworthy, relevant and persuasive, the aggregate balance of these pieces should lie towards the center of the triangle: the ideal should be a case based on strong empirical, rational and historical evidence. However, we would suggest that currently the ideal position is considered by many in the healthcare field to lie close to the empiricist apex of the triangle: empirical data, which, in its pure form, is atheoretical, tends to dominate clinician thinking within the healthcare domain. This is unfortunate, both because such evidence is often unavailable, and when it is available it frequently lacks relevance. Meanwhile, managers are often drawn to seeking a case focused on the historicist data. Once again, such evidence is often scarce, and to the extent that it is available, it frequently lacks rigor.

We characterize models (and, in particular, simulation models) as sources of rationalist evidence. Broadly, we understand models to be theoretical constructs which embody rationally constructed arguments from explicit premises. The models, to be sure, may have different evidential “flavors” – a system dynamics simulation model, for example, might be based on cause-effect relationships which are derived from observation rather than theoretically understood (tending to an empiricist flavor), or might be designed to reproduce real, idiosyncratic, events in a real case (tending to the historicist). However, the overall model would constitute a rationalist construct – a theoretical representation of the situation to be studied. Modeling and simulation are rationalist evidential instruments. Since we consider that both clinical and management practice lack the rational dimension to evidential arguments (clinical practice being strongly empiricist, and management being strongly historicist), modeling and simulation have much to offer in terms of filling a “rationalist evidential gap” in healthcare management. Indeed, modeling fulfills this role in other domains, too, although in many its rationalist perspective tends already to be implicitly accepted.

4 USE OF EVIDENCE

The Festival of Evidence was intended to develop ideas around evidence as well as consider evidence itself. The discussions that this broader intention generated are hard to report, but they created a great deal of insight. These sessions tended to lead to diagrams – often generated between sessions and circulated electronically – together with a supporting narrative.
A recurrent theme was that evidence on its own was rarely sufficient to encourage people to adopt simulation and modeling. The importance of the role of the champion was addressed in more than one talk, especially Mike Carter’s presentation, and is implicit in the literature. Noting how the advertising industry can often encourage people to adopt even unhealthy lifestyles, the discussion noted the importance of stories. The concept of a modeling adoption cycle within an organization emerged. Such a cycle begins with a general awareness of modeling. As a result of this awareness, a champion might decide to use modeling to address a particular concern. Such use might give rise to an anecdote: a modeling success story. If the champion were sufficiently well-placed, corporate support might then be forthcoming, and modeling is adopted as an approach to addressing appropriate issues. Then, as examples of successful modeling accumulate, modeling can become the cultural norm. This conceptualization takes a different view of evidence from, say, EBM, where the persuasiveness of evidence is generally related to its empirical strength and reliability (though not necessarily its relevance). Here, persuasiveness is related to the narrative strength of the evidence.

Pursuing this theme led to consideration of the nature of evidence and the scenarios in which different kinds of evidence might prove satisfactory. The tension between stories and statistics (or, narrative and numbers) as evidential forms was identified and related to the contrast between clinical and operational scenarios. Many journeys would start with a compelling narrative with which a well-placed clinician could identify. Corporate support would require evidence as narrative and as numbers, to persuade the decision-making communities at that level. Finally, ubiquitous modeling embedded in the culture would require a balance of clinically and operationally oriented evidence, both as numbers and narratives.

Further tensions between varying forms of evidence were teased out. Evidence might be range from real-time, available as feedback to decisions almost immediately, to long-term, where it might only emerge from analysis and review a significant period of time after the events to which it related. Some evidence is paper-based while other evidence is electronic; the latter is generally in a form which makes it more easily subject to analysis. While some evidence might be explicitly held in databases or similar (in silico), other evidence might exist only in people’s heads (in vivo). Some evidence relates to small problems well-solved, while other evidence relates to wicked problems only partially overcome – which is more valid, or valuable, or persuasive? Evidence may relate to efficiency (doing things with less resource usage) or efficaciousness (doing things right) or effectiveness (doing the right things right). Evidence may provide concrete yes-no answers, or just change confidence levels.

One of the main outcomes of these discussions was that the focus of the debate moved away from a concept of evidence as persuasive by virtue of its very existence, to one that recognizes that evidence needs to be persuasive – its social dimension needs to be acknowledged. Those who seek evidence need to be persuaded, those attempting to evangelize need to have material that enables them to persuade, and evidence needs to be able to relate problem to solution in a manner that is appropriately persuasive to the audience who may be exposed to it. The most useful piece of this argument is probably the way in which it allows plans for using evidence to support change to be formulated. Given that more than half a century of healthcare modeling has led to so little adoption, this is an important step.

Figure 1, below, illustrates one of the arguments explored during the discussions. The horizontal axis contrasts story-based evidence with statistics-based evidence. The vertical axis relates to an organization’s appetite for change – hunger at one extreme, hostility at the other. Within this diagram we might identify trajectories of persuasion amongst various people. Two such trajectories are shown. On the left of the diagram, the trajectory from bottom-left to top-left quadrant depicts people who move from a position of skepticism to change to one of openness purely through persuasion by appropriate stories. However, it is possible to identify those for whom the journey might be more complicated (the trajectory from bottom-right to top-right quadrant) – in which the narrative plays a catalytic effect, perhaps, in moving them from being sceptics to evangelists. In this case, their natural affinity to numbers-based evidence marks the start and the end points of their journey.
The main impact of these discussions was to drive the Cumberland Initiative towards a more multifaceted approach to evidence. The main conclusion was that the quest for more evidence is always going to be difficult. It will, for instance, be almost impossible to generate evidence at scale without evidence at scale to start with. The inability of the healthcare community to scale up pilots is a clear demonstration of this.

However, these discussions focused attention on the persuasiveness of the evidence, the appropriateness of the evidence, and the nature of the evidence against the task faced by a putative problem owner. In a sense, none of this is new, but trying to bring a number of theories together in order to change the cultural outlook of the largest and most complex sector in the world – that has some novelty. We might conclude that there is a need for a development of an understanding of the cultural context in the healthcare sector.

5 MODELING IN HEALTHCARE: SURVEYING THE LITERATURE

Under the aegis of the EPSRC-funded RIGHT project (Research Into Global Healthcare Tools) led by Terry Young, Brailsford et al. (2009) explored the research literature on modeling and simulation in healthcare using a sampling methodology. They found that only 5.3% of the papers they surveyed could be classified as “implemented”, with 50% not reporting a practical implementation, and 44.7% being academic studies with no healthcare practitioner input. Other reviewers (for example: Fone et al 2003; Brailsford and Vissers 2010; Katsaliaki and Mustafee 2011) have reached similar conclusions. Reflecting on the lack of implementation, Brailsford and De Silva (2015) suggest that access to both policy-makers and clinicians is an important factor, while Harper and Pitt (2004) propose a project lifecycle for modeling in healthcare to understand successful implementation. Brailsford et al. (2009) argue that there is much healthcare modeling work described in the “grey” literature, but point to the difficulty in accessing it systematically and comprehensively, and that such literature may have more useful
material on implementation issues than the conventional academic literature. Innovative tools that can search grey literature effectively would be welcome.

6 OTHER WORK IN THE PUBLIC DOMAIN

MASHnet, the UK Network for Modeling and Simulation in Health, was established in 2005 to improve the application of modeling and simulation in health and social care. A booklet published by MASHnet (2014) describing five successful modeling case studies provides much material to confirm the value of modeling and simulation approaches as providing valuable support for the introduction of new practices within healthcare organizations.

Scenario Generator, produced by Simul8 Corporation in 2009 as part of a joint venture with the then NHS Institute for Innovation and Improvement (NHS III), is a simulation tool in which users can design care pathways for particular conditions or types of patient within a geographical region, which they can then test against demographic and disease incidence and prevalence data. This enables the implications of particular healthcare commissioning decisions to be explored prior to committing to them. Brailsford et al. (2013) carried out a study of the use of Scenario Generator, and concluded that factors leading to its successful use in organizations included:

- The presence of a champion who would enthuse to colleagues about the value of modeling.
- A business-critical model to work on.
- External factors encouraging the adoption of modeling.
- Sustainability of expertise within the organization.

Barriers to adoption were also identified: in particular, lack of time, capacity, analytic capability and data availability.

7 SUBMITTED EVIDENCE

Participants had been invited to submit evidence to the FoE: 23 such submissions were gathered. These clearly represent a minute proportion of all applications of simulation and modeling in health. It was noteworthy that, of the 23, only seven of the submissions involved academics. Two of these seven were solely by academics, the data sets they utilized, though provided by hospitals, were for illustrative purposes only. Another two of the seven were academic-led but constituted academic-hospital collaborative partnerships, and the studies were used by the hospitals. The other three academic submissions arose from an initiative led by Professor Paul Harper at the University of Cardiff, in which five modelers hold joint appointments at the University and at Aneurin Bevan University Health Board (ABUHB).

Of the other sixteen submissions, ten were commercial, while six were from NHS/Local Authority groups. Some of the commercial examples had a marketing orientation: they described software tools or systems, describing the capabilities of the systems, but not including any evidence of benefit. Two of the software providers, though, did report on users and benefits. The remaining commercial examples were submitted by modeling consultancies. Of the six submissions from NHS organizations or Local Authorities, two reported joint work with commercial partners, while the others described work they had undertaken themselves.

In terms of the distinction between empiricist, rationalist and historicist evidence that we formulated earlier in the paper, we note that most of the evidence that was presented fell into the empiricist (RCTs and similar) or rationalist (simulation and modeling, arguably) categories, or mixes thereof, rather than the historicist (stories and anecdotes) category. Perhaps, given the nature of the FoE, this was inevitable. However, the point we would emphasize is that we suggest that, ideally evidence should be balanced at the center of the evidential triangle – and, indeed, one or two of the submissions presented a mix of evidence that conformed to this ideal.
8 THE CHALLENGE

While it is clear that there is a substantial amount of modeling and simulation activity in the area of healthcare, it is fragmented and there is little sign of diffusion of ideas and approaches. Rather, on each occasion, modelers and their clients have to start from scratch, although there may be many instances of successful modeling interventions in similar situations elsewhere. Our experience suggests several reasons for this. First, the “not invented here” syndrome – a skepticism in accepting that something that worked elsewhere might be applicable in the current situation. Second, the data that may be needed to inform a model may be hard to access. Third, modeling is sometimes viewed as a “dehumanizing” approach that fails to acknowledge patients as unique human individuals. The challenge for those who wish the benefits of modeling to be fully exploited within the healthcare sector is to find ways of convincing senior healthcare managers that these benefits are real and substantial. In terms of evidence, what evidence is required to meet this challenge, and what form should this evidence take?

9 A NEW HOPE

On 23rd October 2014, the fourth day of the Festival of Evidence, Simon Stevens, Chief Executive of NHS England, published his NHS Five Year Forward View (NHS 2014), which contains the following: “working with NIHR and the Department of Health we will expand NHS operational research, RCT capability and other methods to promote more rigorous ways of answering high impact questions in health services redesign” (p. 34). The explicit mention operational research in this document did much for the morale of the FoE participants, and may be interpreted, perhaps, as the dawn of a more balanced approach to healthcare planning within the UK, with greater emphasis on the rationalist paradigm in particular.

Other encouraging signs of an increasing recognition of the value of modeling and simulation in UK healthcare may be detected. Many of the NIHR Collaborations for Learning in Applied Health Research and Care (CLAHRCs) include groups of modeling personnel. Academic Health Science Networks (AHSNs) have been set up and are already beginning to realize their potential. As alluded to earlier in this paper, a joint initiative between a Health Board in Wales (ABUHB) and the University of Cardiff now employs five modelers. Commercial organizations are setting up membership groups (such as the Whole Systems Partnership Workforce Collaborative and Simul8’s Scenario Generator User Groups) to actively support NHS personnel in their use of modeling tools.

10 OVERALL CONCLUSIONS AND RECOMMENDATIONS

Overall, the Festival of Evidence both confirmed and strengthened many of our expectations concerning the role of modeling, and simulation in particular, within healthcare, as well as leading to new insights. Generally, we can conclude that modeling and simulation are extremely useful techniques to improve decision-making, but have yet to achieved the potential in healthcare that they have realized in other domains. There is an abundance of evidence, both from within the UK and beyond, for the benefits of simulation and modeling, but it takes diverse forms, is fragmented, and is frequently anecdotal. Successful “rollout” of modeling interventions beyond the original client context seems comparatively rare. Within the UK, there are encouraging signs that modeling and simulation are becoming more embedded within NHS organizations.

There are different forms of evidence; evidence means different things to different people, and different people are convinced by different types of evidence. For this reason, and others, the best evidence upon which to base healthcare interventions is a mix of congruent empiricist, rationalist and historicist types. However, to assemble such a mix of evidence takes time. We would generally expect the journey of evidential support to start from an empiricist, rationalist or historicist extreme, and move towards a more balanced profile. We see a particular role for modeling and simulation in supplying rationalist evidence: simulation modelers may therefore find it useful to consider the evidence their work generates as complementing other forms of evidence, and direct their efforts accordingly.
Specifically addressing the issue of scaling up the use of modeling and simulation within healthcare, we suggest that for them to become part of the NHS management “toolkit” it is essential that they are embedded within NHS organizations, and are not seen either as requiring exceptional skills, or as activities that can only be carried out by specialist consultants. Promotion of modeling and simulation should not be left to a few enthusiasts: it is vital that senior NHS management has a real and non-superficial appreciation of their value. There is some indication that a statutory requirement to demonstrate the use of modern management approaches can be helpful in bringing this about.

The cultural divide between clinical and management perspectives in healthcare organizations inhibits the appropriate use of evidence in healthcare. Efforts should be made, not to eliminate the divide, but rather to create a pluralist “meta-culture” in which each side appreciates the values and strengths of the other, and acknowledges its own limitations. Meanwhile, researchers (in the areas of both medicine and management) should take care to present their findings in ways which facilitate their adoption where appropriate by potential users. In particular, evidence from modeling organizational innovations and interventions needs to be presented to clinicians and managers in a different way to evidence for new treatments: this requires a culture change within the NHS. The modeling community would benefit from stronger links with the implementation science community.

Regarding the future, we conclude that evidence-based practice should be based on a view of evidence generation that allows for more diverse evidence generation, and recognizes the strengths of such generation methods. There is a need for new ways of categorizing and evaluating different types of evidence, especially for organizational interventions. More academic research should be directed towards the “grey” literature, and new and more rigorous tools are required for analyzing and interpreting its contents. We need to develop evidence-based practice to a degree where it matches evidence-based medicine in terms of rigor and acceptability, while recognizing that it is likely that EBM can benefit from the insights that a broader-based EBP can provide.

REFERENCES


Brailsford, Klein, and Young


**AUTHOR BIOGRAPHIES**

SALLY BRAILSFORD is Professor of Management Science at the University of Southampton, UK. She received a BSc in Mathematics from the University of London, and MSc and PhD in Operational Research from the University of Southampton. Her research interests include simulation modeling methodologies, system dynamics, health service research and disease modeling, and the modeling of human behavior in healthcare systems. She is chair of the European Working Group on OR Applied to Health Services (ORAHS) and is an Editor-in-Chief of the journal Health Systems. She is on the editorial boards of Health Care Management Science, the Journal of Simulation, and Operations Research for Health Care. Her email address is s.c.brailsford@soton.ac.uk.

JONATHAN H. KLEIN is an Associate Professor of Management Science at the Southampton Business School, University of Southampton, UK. He has BSc and MSc degrees in Physics from the University of Sussex, an a PhD in Operational Research from Royal Holloway (University of London). His research interests include problem structuring and knowledge management, with particular application to healthcare management. His email address is J.H.Klein@soton.ac.uk.

TERRY YOUNG joined Brunel University as Professor of Healthcare Systems in 2001 after 16½ in industrial research and technology. Since then he has specialised in leading large, cross-disciplinary, multi-university projects addressing value in care delivery. His research interests are the organisation and delivery of care, at all levels, from the role and value of technology to modeling, simulations and business models of new services and structures. Terry is also a leader of the Cumberland Initiative; a unique collaboration between health, industry and academe, committed to building a national centre in services and systems excellence. His email address is Terry.Young@brunel.ac.uk.