Using data science to improve patient care

Rethinking clinician responsibility

Grossmann, I.; Marang-van de Mheen, P.J.

DOI
10.1136/bmjqs-2025-018514

Publication date
2025

Document Version
Final published version

Published in
BMJ Quality and Safety

Using data science to improve patient care: rethinking clinician responsibility

Irene Grossmann [1,2] Perla J Marang-van de Mheen [3]

[1]Institute for Health Systems Science, Center for Safety in healthcare and Safety & Security Science section, Delft University of Technology, Delft, Netherlands
[2]Emergency Care, Deventer Hospital, Deventer, Netherlands
[3]Safety & Security Science and Centre for Safety in Healthcare, TU Delft TPM, Delft, Netherlands

Correspondence to
Dr Irene Grossmann;
I.grossmann@tudelft.nl

Received 17 January 2025
Accepted 10 February 2025

Linked

► https://doi.org/10.1136/bmjqs-2024-017935

Check for updates

To cite: Grossmann I, Marang-van de Mheen PJ. *BMJ Qual Saf* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjqs-2025-018514

'Knowing what you are doing' is a simple, but elemental value for any (care) professional. Acknowledging that treatments in healthcare can be inherently harmful, and the practice of medicine often involves weighing one harm (the disease) against the other (the treatment), it is obviously vital to know and understand the effects of medical interventions on humans. However, healthcare is becoming increasingly complex, not in the least due to the abundant body of in-depth knowledge that professionals need to weigh into their decisions for patients. Data science is rapidly changing healthcare as we speak, creating tools such as scores,[1 2] benchmarks provided by clinical audits[3] and guidelines that alter our clinical strategies. Artificial intelligence and machine learning solutions may be less comprehensible than the information provided by, for example, guidelines, but are revolutionising the world and healthcare at an unstoppable speed.[4] The added value of these developments is not in question, and neither is their position within healthcare.[5] The real issue is, how can we as clinicians come to trust these tools and use them to improve care?

Feedback on performance, including visual graphs such as statistical process control charts to monitor patient outcomes,[6] is important as a way to evaluate treatment and policy decisions, to enable implementation of changes if patient outcomes deteriorate. To be able to act on performance feedback, clinicians need to understand it, so feedback should be as simple and intuitive as possible. This was the rationale for the new chart developed by Cordier and colleagues in this issue of *BMJ Quality & Safety* that can be used for quality improvement in surgery.[7] Their study revolves around the dilemma of using simplified tools that may be easier to use against charts conveying more, or more statistically accurate information, but which may be too complex for surgeons to use in practice. They state that ideally '*The selected tools must balance statistical rigour with surgeon usability, enabling both statistical interpretation of trends over time and comprehensibility for surgeons, their primary users.*' Cordier and colleagues therefore combine the simpler and more intuitive Observed minus Expected (O−E) chart with the statistically robust risk-adjusted CUSUM chart to provide a balanced solution and provide an example of its use using surgical data. With this chart, surgeons are able to identify sequences of abnormal changes in outcomes and quantify the amount of associated adverse events during these sequences. This adds to other literature of dual charts, in which the informative (statistical) properties of different chart types are combined.[8]

TAPPING INTO CLINICIAN INTUITION

A key prerequisite for using a statistical chart in a clinical setting is conveying information intuitively, and Cordier and colleagues[7] considered the O−E chart as highly intuitive to visualise the safety of care over time in the combined chart. However, they did not provide evidence supporting that notion. It does operate on a procedure-by-procedure basis, which may make it better aligned with the clinician mental model as clinicians are known to have difficulty in translating aggregated data into patient-centred actions.[9] But what makes something intuitive? One of us is a surgeon, who can walk down the corridor in the ER and immediately sense that something is up, intuitively picking up various soft signals. The other is an epidemiologist who will not catch these signals when walking that same corridor

at the same time. Instead, she will immediately pick up any abnormality in a dataset or figure, which might be missed by the other. This simple example shows how the different training that we both had will determine what signals we will intuitively pick up. Whether something is intuitive or not is not a one-size-fits-all but is influenced by various factors, most notably education and training. Even where surgeons may rely on those with statistical expertise to help them select which chart to use, the question is still whether these graphs convey what they need to know to improve their work and are able to correctly interpret and use them. For instance, is it enough to know one's performance has significantly become worse for a surgeon to take action—information one can obtain from the CUSUM chart when an alert is generated—or is the absolute number of patients harmed also needed? Could providing the absolute numbers inadvertently have adverse effects, for example, not taking action if numbers are too low? Do I trust the signal given by this chart, even if I do not exactly understand what the threshold is for such a signal? These are important questions that require further testing and research. For now, the implication is that we should not assume charts are intuitive but offer education—needed in most QI initiatives[10]—or add narratives to the charts to ensure proper interpretation. But more is needed.

THE NEED FOR EXPLAINABLE DATA AND TOOLS

Professionals working in healthcare, including doctors and other staff overseeing quality and safety, will only take action when they trust the information to be correct and feel they are the designated actors to take responsibility. A known first response to data signalling deteriorating performance is often 'the data are wrong'. But even if one considers the data to be correct, the response is often 'it is not my responsibility'. How do we create an environment where the data are trusted, and it is clear who has the responsibility to act on the information? The Dutch Institute for Clinical Auditing (DICA) started to systematically gather clinical data to create a benchmark on outcomes in 2009 with the Dutch Surgical Colorectal Audit (DSCA), and was a pioneering clinician-led initiative.[3 11] The fact that data were collected by and for clinicians, together with a chart that could be understood with little explanation, likely contributed to the eventual adoption and trust by the clinical community. Not because the medical community fully understood or even agreed with all the methods used, but it had come to trust the data to be the 'best possible information'. But it still is a shaky trust, and innovations 'from the outside', in particular, encounter many obstacles to adoption in healthcare.[12]

Meanwhile, we increasingly use various tools and technologies in healthcare to support our treatment and policy decisions, such as the above-mentioned guidelines, decision support tools and other clinical scores. We need these to be able to process multiple sources of knowledge and information, as there is just too much to know, too much information to comprehend and weigh for an individual (care) professional. For these advanced tools as well as information systems on performance and patient outcomes, the issue is whether we fully understand and trust them and the information on which they are based.

Let's take as an example the decision of which anticoagulant is best for an 88-year-old female patient with paroxysmal atrial flutter, no history of smoking or signs of atherosclerosis and a tendency to fall due to inner ear problems. This decision is determined by scores and guidelines, and there is no way of knowing whether the provided advice is the best for my individual patient. It comes down to the trustworthiness of these information sources and their ability to serve the true needs of my patient. Questions that immediately come to mind are: Are the results explainable? Do these scores consider all aspects we deem relevant? What data sources are these calculations and recommendations based on? Who funds this research or researchers? In other words, can we trust the advice as truly being the best for our patient? With the necessity for, and availability of, transdisciplinary approaches in medicine, it is difficult to merge knowledge[13] and build trusted systems. Methods from other sciences, such as game theory from mathematics[14 15] and human factors engineering,[16] are being introduced into medicine and quality improvement, but such transdisciplinary approaches need guidance and leadership to flourish.[17] Instead of an individual profession, medicine is in a transition to become a transdisciplinary team science[18] and this evokes new questions in how we guide this transition.

TRUST AND RESPONSIBILITY ARE KEY

One of the key questions is how we allocate responsibility. Clinicians are both morally and legally responsible for diagnostic and treatment decisions, and clinical strategies to ensure safe care. But who is responsible if clinicians follow the recommendations from decision support or prediction tools created by non-clinicians, which turn out to be harmful for certain patient groups?[19] The relevance and weight of the (felt) responsibility by medical professionals should not be underestimated. Historically, the individual care professional is held fully accountable for outcomes of care, and this has not changed at the same pace with the transition towards healthcare as a collaborative effort that includes non-clinical disciplines. This responsibility is thoroughly felt by clinicians who therefore often strive to keep control, a behaviour that may be misconceived as a hindrance or arrogance. So, how to deal with care responsibility in a transdisciplinary environment? Can we design a better alignment of responsibilities to optimise safe care? A relevant question in the context of the article by Cordier

and colleagues could be who would be responsible if signals of deteriorating performance are missed? This issue relates to the wider problem of shared responsibility, such as the issue of healthcare institutional boards' responsibility in safety in healthcare.[20] The key issue here, whether we talk about embracing advanced methods for monitoring or clinical decision-making or governance in healthcare, is that we need to build trust. The first step is to do this through mutual understanding and collaboration, and by applying design thinking for technological innovations,[21] so that the user's needs are properly identified and addressed an iterative, transdisciplinary design process. But it also requires updated regulatory frameworks that cover the responsible design of technological innovations and appropriately share accountability.

Coming back to the article by Cordier and colleagues, is the solution to ***balance*** understandability with rigour? Or would it be better to ***maximise*** the statistical rigour and use separate design methods to inform clinicians in an understandable manner? The question should not primarily be how mathematicians can create an understandable chart, but whether and how care professionals can trust the information and signals resulting from these tools. We think that data science and advanced methods are needed to continue improving care. However, building trust through transdisciplinary understanding and collaboration in teams, with a balanced division of responsibilities and accountability, is key to responsibly use innovations in healthcare.

Contributors IG and PJMvdM both contributed to conception of the paper, IG wrote the first draft and both authors critically read and modified subsequent drafts and approved the final version. GI is the guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Commissioned; internally peer reviewed.

ORCID iDs
Irene Grossmann http://orcid.org/0000-0003-1863-6827
Perla J Marang-van de Mheen http://orcid.org/0000-0003-1439-0989

REFERENCES

1 Shewale AR, Johnson JT, Li C, *et al*. Net Clinical Benefits of Guidelines and Decision Tool Recommendations for Oral Anticoagulant Use among Patients with Atrial Fibrillation. *J Stroke Cerebrovasc Dis* 2015;24:2845–53.
2 Chua WL, Rusli KDB, Aitken LM. Early warning scores for sepsis identification and prediction of in-hospital mortality in adults with sepsis: A systematic review and meta-analysis. *J Clin Nurs* 2024;33:2005–18.
3 Beck N, van Bommel AC, Eddes EH, *et al*. The Dutch Institute for Clinical Auditing: Achieving Codman's Dream on a Nationwide Basis. *Ann Surg* 2020;271:627–31.
4 Alowais SA, Alghamdi SS, Alsuhebany N, *et al*. Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Med Educ* 2023;23:689.
5 Verma AA, Trbovich P, Mamdani M, *et al*. Grand rounds in methodology: key considerations for implementing machine learning solutions in quality improvement initiatives. *BMJ Qual Saf* 2024;33:121–31.
6 Marang-van de Mheen PJ, Woodcock T. Grand rounds in methodology: four critical decision points in statistical process control evaluations of quality improvement initiatives. *BMJ Qual Saf* 2023;32:47–54.
7 Cordier Q, Prieur H, Duclos A. Risk-adjusted observed minus expected cumulative sum (RA O-E CUSUM) chart for visualisation and monitoring of surgical outcomes. *BMJ Qual Saf* 2024.
8 Ilieş I, Anderson DJ, Salem J, *et al*. Large-scale empirical optimisation of statistical control charts to detect clinically relevant increases in surgical site infection rates. *BMJ Qual Saf* 2020;29:472–81.
9 Desveaux L, Rosenberg-Yunger ZRS, Ivers N. You can lead clinicians to water, but you can't make them drink: the role of tailoring in clinical performance feedback to improve care quality. *BMJ Qual Saf* 2023;32:76–80.
10 Soong C, Shojania KG. Education as a low-value improvement intervention: often necessary but rarely sufficient. *BMJ Qual Saf* 2020;29:353–7.
11 Van Leersum NJ, Snijders HS, Henneman D, *et al*. The Dutch surgical colorectal audit. *Eur J Surg Oncol* 2013;39:1063–70.
12 Berwick DM. Disseminating innovations in health care. *JAMA* 2003;289:1969–75.
13 Colwell R, Eisenstein R. From microscope to kaleidoscope: merging fields of vision. In: Klein JT, Häberli R, Scholz RW, et al, eds. *Transdisciplinarity: Joint problem solving among science, technology, and society: An effective way for managing complexity*. Basel: Birkhäuser Basel, 2001: 59–66.
14 Stanková K, Brown JS, Dalton WS, *et al*. Optimizing Cancer Treatment Using Game Theory: A Review. *JAMA Oncol* 2019;5:96–103.
15 Tarrant C, Stokes T, Colman AM. Models of the medical consultation: opportunities and limitations of a game theory perspective. *Qual Saf Health Care* 2004;13:461–6.
16 Gosbee J. Human factors engineering and patient safety. *Qual Saf Health Care* 2002;11:352–4.
17 Gray B. Enhancing transdisciplinary research through collaborative leadership. *Am J Prev Med* 2008;35:S124–32.
18 Stokols D, Hall KL, Taylor BK, *et al*. The science of team science: overview of the field and introduction to the supplement. *Am J Prev Med* 2008;35:S77–89.
19 Teeple S, Chivers C, Linn KA, *et al*. Evaluating equity in performance of an electronic health record-based 6-month mortality risk model to trigger palliative care consultation: a retrospective model validation analysis. *BMJ Qual Saf* 2023;32:503–16.
20 Millar R, Mannion R, Freeman T, *et al*. Hospital board oversight of quality and patient safety: a narrative review and synthesis of recent empirical research. *Milbank Q* 2013;91:738–70.
21 Kleinsmann M, Valkenburg R, Sluijs J. Capturing the value of design thinking in different innovation practices. *Int J Design* 2017;11:25–40.