

Simulation: a key tool for refining guidelines and demonstrating they produce the desired behavioural change

Mark Fan,¹ Patricia Trbovich²

¹Research and Innovation, North York General Hospital, Toronto, Ontario, Canada

²Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada

Correspondence to

Dr Patricia Trbovich;
patricia.trbovich@utoronto.ca

Accepted 13 July 2020

Guidelines aim to align clinical care with best practice. However, simply publishing a guideline rarely triggers behavioural changes to match guideline recommendations.^{1–3} We thus transform guideline recommendations into actionable tasks by introducing interventions that promote behavioural changes meant to produce guideline-concordant care. Unfortunately, not much has changed in the 25 years since Oxman and colleagues concluded that we have no ‘magic bullets’ when it comes to changing clinician behaviour.⁴ In fact, far from magic bullets, interventions aimed at increasing the degree to which patients receive care recommended in guidelines (eg, educational interventions, reminders, audit and feedback, financial incentives, computerised decision support) typically produce disappointingly small improvements in care.^{5–10}

Much improvement work aims to ‘make the right thing to do the easy thing to do.’ Yet, design solutions which hard-wire the desired actions remain few and far between. Further, improvement interventions which ‘software’ such actions—not guaranteeing that they occur, but at least increasing the likelihood that clinicians will deliver the care recommended in guidelines—mostly produce small improvements.^{5–9} Until this situation changes, we need to acknowledge the persistent reality that guidelines themselves represent a main strategy for promoting care consistent with current evidence, which means their design should promote the desired actions.^{11 12}

In this respect, guidelines constitute a type of clinical decision support. And, like all decision support interventions, guidelines require: (1) user testing to assess if

the content is understood as intended and (2) empirical testing to assess if the decision support provided by the guideline does in fact promote the desired behaviours. While the processes for developing guidelines have received substantial attention over the years,^{13–18} surprisingly little attention has been paid to empirically answering basic questions about the finished product: do users understand guidelines as intended? And, what version of a given guideline engenders the desired behaviours by clinicians?

In this issue of *BMJ Quality and Safety*, Jones *et al*¹⁹ address this gap by using simulation to compare the frequency of medication errors when clinicians administer an intravenous medication using an existing guideline in the UK’s National Health Service (NHS) versus a revised and user-tested version of the guideline that more clearly promotes the desired actions. Their findings demonstrate that changes to guideline design (through addition of actionable decision supports) based on user feedback does in fact trigger changes in behaviour that can improve safety. This is an exciting use of simulation, which we believe should encourage further studies in this vein.

ENSURING END USERS UNDERSTAND AND USE GUIDELINES AS INTENDED

Jones and colleagues’ approach affords an opportunity to reflect on the benefits of user testing and simulation of guidelines. The design and evaluation of their revised guidelines provides an excellent example of a careful stepwise progression in the development and evaluation of a guideline as a type of decision support for clinicians. First, in a prior study,²⁰ they user



► <http://dx.doi.org/10.1136/bmjqs-2020-010884>



© Author(s) (or their employer(s)) 2020. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Fan M, Trbovich P. *BMJ Qual Saf* Epub ahead of print: [please include Day Month Year]. doi:10.1136/bmjqs-2020-011999

tested the original NHS guidelines to improve retrieval and comprehension of information. The authors produced a revised guideline, which included reformatted sections as well as increased support for key calculations, such as for infusion rates. The authors again user tested the revised guideline, successfully showing higher rates of comprehension. Note that user testing refers to a specific approach focused on comprehension rather than behaviour²¹ and is distinct from *usability* testing. Second, in the current study, Jones *et al* evaluated whether nurse and midwife end users exhibited the desired behavioural changes when given the revised guidelines (with addition of actionable decision supports), compared with a control group working with the current version of the guidelines used in practice. As a result, Jones and colleagues verify that end users (1) understand the content in the guideline and (2) actually change their behaviour in response to using it.

Simulation can play a particularly useful role in this context, as it can help identify problems with users' comprehension of the guideline and also empirically assess what behavioural changes occur in response to design changes in the guidelines. The level of methodological control and qualitative detail that simulation provides is difficult to feasibly replicate with real-world pilot studies, and therefore simulation fills a critical gap.

Jones *et al* report successful changes in behaviour due to the revised guidelines in which they added actionable decision supports. For example, their earlier user testing found that participants using the initial guidelines did not account for displacement volume when reconstituting the powdered drug, leading to dosing errors. A second error with the initial guidelines involved participants using the shortest infusion rate provided (eg, guidelines state '1 to 3 hours'), without realising that the shortest rate is not appropriate for certain doses (eg, 1 hour is appropriate for smaller doses, but larger doses should not be infused over 1 hour because the drug would then be administered faster than the maximum allowable infusion rate of 3 mg/kg/hour). These two issues were addressed in the revised guidelines by providing key determinants for 'action' such as calculation formulas that account for displacement volume and infusion duration, thereby more carefully guiding end users to avoid these dose and rate errors. These changes to the guideline triggered specific behaviours (eg, calculations that account for all variables) that did not occur with the initial guidelines. Therefore, the simulation testing demonstrated the value of providing determinants for action, such as specific calculation formulas to support end users, by showing a clear reduction in dose and rate errors when using the revised guidelines compared with the initial guidelines.

The authors also report that other types of medication-specific errors remained unaffected by

the revised guidelines (eg, incorrect technique and flush errors)—the changes made did not facilitate the desired actions. The initial guidelines indicate 'DO NOT SHAKE' in capital letters, and there is a section specific to 'Flushing'. In contrast, the revised guidelines do not capitalise the warning about shaking the vial, but embed the warning with a numbered sequence in the medication preparation section, aiming to increase the likelihood of reading it at the appropriate time. The revised guidelines do not have a section specific to flushing, but embed the flushing instructions as an unnumbered step in the administration section. Thus, the value of embedding technique and flushing information within the context of use was not validated in the simulation testing (ie, no significant differences in the rates of these errors), highlighting precisely the pivotal role that simulation can play in assessing whether attempts to improve usability result in actual behavioural changes.

Finally, simulation can identify potential unintended consequences of a guideline. For instance, Jones and colleagues observed an increase in errors (although not statistically significant) that were not medication specific (eg, non-aseptic technique such as hand washing, swabbing vials with an alcohol wipe). Given that the revised guidelines were specific to the medication tested, it is unusual that we see a tendency toward a worsening effect on generic medication preparation skills. Again, this finding was not significant, but we highlight this to remind ourselves of the very real possibility that some interventions might introduce new and unexpected errors in response to changing workflow and practice⁶; simulations offer an opportunity to spot these risks in advance.

Now that Jones *et al* have seen how the revised guidelines change behaviour, they are optimally positioned to move forward. On one hand, they have the option of revising the guidelines further in attempts to address these resistant errors, and on the other, they can consider designing other interventions to be implemented in parallel with their user-tested guidance. At first glance, the errors that were resistant to change appear to be mechanical tasks that end users might think of as applying uniformly to multiple medications (eg, flush errors, non-aseptic technique). Therefore, a second intervention that has a more general scope (rather than drug specific) might be pursued. Regardless of what they decide to pursue, we applaud their measured approach and highlight that the key takeaway is that their next steps are supported with clearer evidence of what to expect when the guidelines are released—certainly a helpful piece of information to guide decisions as to whether broad implementation of guidelines is justified.

CAVEATS AND CONCLUSION

Simulation is not a panacea—it is not able to assess longitudinal adherence, and there are limitations to

how realistically clinicians behave when observed for a few sample procedures when under the scrutiny of observers. Further, studies where interventions are implemented to assess whether they move the needle on the outcomes we care about (eg, adverse events, length of stay, patient mortality) are needed and should continue. However, having end users physically perform clinical tasks with the intervention in representative environments represents an important strategy to assess the degree to which guidelines and other decision support interventions in fact promote the desired behaviours and to spot problems in advance of implementation. Such simulation testing is not currently a routine step in intervention design. We hope it becomes a more common phenomenon, with more improvement work following the example of the approach so effectively demonstrated by Jones and colleagues.

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 required.

Provenance and peer review Not commissioned; internally peer reviewed.

REFERENCES

- Grimshaw JM, Thomas RE, MacLennan G, *et al*. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Int J Technol Assess Health Care* 2005;21:149
- Hawe CS, Ellis KS, Cairns CJS, *et al*. Reduction of ventilator-associated pneumonia: active versus passive guideline implementation. *Intensive Care Med* 2009;35:1180–6.
- Mortimer D, French SD, McKenzie JE, *et al*. Economic evaluation of active implementation versus guideline dissemination for evidence-based care of acute low-back pain in a general practice setting. *PLoS One* 2013;8:e75647.
- Oxman AD, Thomson MA, Davis DA, *et al*. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995;153:1423–31.
- Soong C, Shojania KG. Education as a low-value improvement intervention: often necessary but rarely sufficient. *BMJ Qual Saf* 2020;29:353–7.
- Soong C, Cho HJ, Shojania KG. Choosing quality problems wisely: identifying improvements worth developing and sustaining. *BMJ Qual Saf* 2020. doi:10.1136/bmjqs-2020-011054. [Epub ahead of print: 29 Apr 2020].
- Shojania KG, Jennings A, Mayhew A, *et al*. Effect of point-of-care computer reminders on physician behaviour: a systematic review. *CMAJ* 2010;182:E216–25.
- Ivers N, Jamtvedt G, Flottorp S, *et al*. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012:CD000259.
- Gondi S, Soled D, Jha A. The problem with pay-for-performance schemes. *BMJ Qual Saf* 2019;28:511–3.
- Tricco AC, Ivers NM, Grimshaw JM, *et al*. Effectiveness of quality improvement strategies on the management of diabetes: a systematic review and meta-analysis. *Lancet* 2012;379:2252–61.
- Lomotan EA, Michel G, Lin Z, *et al*. How "should" we write guideline recommendations? Interpretation of deontic terminology in clinical practice guidelines: survey of the health services community. *BMJ Qual Saf* 2010;19:509–13.
- Shiffman RN, Michel G, Rosenfeld RM, *et al*. Building better guidelines with BRIDGE-Wiz: development and evaluation of a software assistant to promote clarity, transparency, and implementability. *J Am Med Inform Assoc* 2012;19:94–101.
- Guyatt GH, Oxman AD, Sultan S, *et al*. Grade guidelines: 9. rating up the quality of evidence. *J Clin Epidemiol* 2011;64:1311–6.
- Guyatt G, Oxman AD, Akl EA, *et al*. Grade guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011;64:383–94.
- Santesso N, Carrasco-Labra A, Langendam M, *et al*. Improving grade evidence tables Part 3: detailed guidance for explanatory footnotes supports creating and understanding grade certainty in the evidence judgments. *J Clin Epidemiol* 2016;74:28–39.
- Grol R, Dalhuijsen J, Thomas S, *et al*. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. *BMJ* 1998;317:858–61.
- Gagliardi AR, Brouwers MC, Palda VA, *et al*. How can we improve guideline use? A conceptual framework of implementability. *Implement Sci* 2011;6:26.
- Kastner M, Bhattacharyya O, Hayden L, *et al*. Guideline uptake is influenced by six implementability domains for creating and communicating guidelines: a realist review. *J Clin Epidemiol* 2015;68:498–509.
- Jones MD, McGrogan A, Raynor DK, *et al*. User-testing guidelines to improve the safety of intravenous medicines administration: a randomised in situ simulation study. *BMJ Qual Saf* 2020. doi:10.1136/bmjqs-2020-010884. [Epub ahead of print: 30 Jun 2020].
- Jones MD, Franklin BD, Watson MC, *et al*. User testing to improve retrieval and comprehension of information in guidelines to improve medicines safety. *J Patient Saf* 2020. doi:10.1097/PTS.0000000000000723. [Epub ahead of print: 19 Jun 2020].
- Raynor DK, Knapp P, Silcock J, *et al*. "User-testing" as a method for testing the fitness-for-purpose of written medicine information. *Patient Educ Couns* 2011;83:404–10.