# Patient safety and rocket science

Peter McCulloch

**Correspondence to**
Peter McCulloch, Nuffield Department of Surgery, Oxford University, Level 6, John Radcliffe Hospital, Oxford OX3 9DU, UK; peter.mcculloch@nds.ox.ac.uk

'This safety stuff, it's not rocket science'. Many readers of this journal will undoubtedly have heard this sentiment expressed by their clinical colleagues. The article by Kemper et al[1] shows just how widely this impression of patient safety misses the mark. This high-quality study confirms the trend of the recent literature by finding that teamwork training using the civil aviation Crew Resource Management (CRM) approach has no evident clinical benefit, although it does seem to change attitudes and enhance some aspects of the 'non-technical' skills involved with interacting with colleagues. In doing so, the study highlights three areas of complexity and challenge in the development and evaluation of safety interventions. First, the interventions themselves are deceptively complex; as recommended by experts, they are grounded in theory,[2] but may be entirely wrong. Second, the success of even 'simple' interventions like the WHO checklist hugely depends on the context and the implementation strategy. And third, the act of evaluation is far more difficult than it might first appear.

Let us start from the end. By the methodological standards of safety and quality intervention studies generally, this is an exceptionally well done study. It is sizeable, involving six hospitals and over 8000 patients. There is a clear 'PICO' question as recommended by evidence-based medicine (EBM) pundits; the study protocol was published in advance; the study uses mixed methods intelligently to study outcomes in a structured way, using Kirkpatrick's educational model; and there is even a control group.

As a practitioner in the same field, the writer salutes the study group for their thorough and thoughtful approach. Yet, by the exacting standards of EBM, even this study would be regarded as being at moderate-to-high risk of bias. The allocation to groups is not random—intervention hospitals needed to sign up to certain financial and organisational standards and agreements, and their ability to

do so may mean that they were in some way superior to the control hospitals. There is no attempt to blind the observers who evaluated communication, and whether the assessment of questionnaires was done in a blinded fashion is unclear. The large number of questionnaire-based instruments raises questions about their validity, reliability and independence of each other, not all of which the authors confront. Finally, there is no attempt to estimate the power of the study to detect the differences it sought, nor any explanation of what size of difference would be considered worthwhile.

In contrast with many other such studies, Kemper and colleagues provide a very clear and detailed description of the CRM intervention, which has varied widely in previous studies in terms of training quantity, content and the percentage of staff receiving it. CRM was the 'poster child' of the patient safety movement before the WHO Surgical Safety Checklist usurped this role, and CRM remains a popular and (for some) profitable intervention because of the persuasive arguments for its effectiveness. These arguments boil down to: (a) it seemed to work for aviation; (b) it must make sense to make team members aware of how things can go wrong, and what good teamwork looks like; (c) it appeals (sometimes at a rather superficial level) to the findings of important psychological work on memory, perception and decision making. However, just because things ought to work does not prove that they do, and sometimes the theoretical basis for proposals for change is dangerously naïve.

The implicit expectation in CRM studies is that the training will change the work culture in a beneficial way. Like the Francis report on egregious failings in care at Mid Staffordshire Hospital, which recommended 'culture change' to improve safety in the NHS,[3] this ignores the stark reality that changing organisational culture is a massive task. Numerous examples in health systems,

police departments and private companies show that strenuous direct efforts to change corporate culture often fail. Even at an individual level, the force we can apply to change culture looks relatively puny. Like joining the army or becoming a student at an elite school or university, becoming a member of a large healthcare system reliably causes major and long-lasting changes in attitudes and behaviour—culture change at the individual level. In all of these examples, change is associated with a mix of intensive, strenuous, stressful and sometimes coercive training and the immersive experience of more subtle but equally strong social pressures over a considerable period. It seems questionable to postulate that these effects can be strongly influenced by occasional training courses.

Once expectations for culture change are reframed by this type of reflection, the results of CRM reported in this and other studies are actually quite impressive. The relatively short course does seem to have made quite a long-lasting impact on attitudes and understanding, and to a certain extent interactions with other team members. However, CRM does not teach people how to make change in their working environment, and in complex workplaces this is not a simple matter. An important, possibly dominant strand of current thinking on patient safety improvement emphasises the use of modified industrial quality improvement techniques and pays much less regard to staff relationships. From this point of view, however well-motivated staff may become following CRM training, it seems unrealistic to expect them to make important structural changes to their work systems with no training in the relevant techniques and with just one session of expert post-CRM 'mentoring'.

The third problem this study raises is not explicitly discussed in the paper, but calls attention to itself in subtle ways, which only others working in the same field might notice. I refer to the importance of context and implementation strategy in safety interventions. Most workers who have tried to initiate these interventions in a live clinical setting have been deeply impressed by how unexpectedly difficult it is. However important safety may be in theory, clinical activity, target achievement and the financial bottom line are always likely to trump it in practice. I suspect it was a realistic understanding of this which led the authors to set a standard for inclusion in the intervention arm, which presumably gave them some assurance about what the hospital management were willing to do—and spend—to support the study.

There are at present no good tools for assessing the structural properties and culture of a clinical organisation in a way which can reliably predict its reaction to a safety improvement programme, but it is clear that some hospitals are much more ready for such interventions than others. One of the most frequently quoted studies in support of the CRM approach is the large study by Neily *et al*[4] of the Medical Team Training Programme introduction in the Veterans Health Administration System. One explanation for the compelling 50% reduction in surgical mortality beyond the secular trend seen in control hospitals lies in the commitment to the intervention evinced by the 2 months of preparation and planning with each facility's implementation team as well as the day-long onsite learning session, which involved closing operating theatres for the day. Moreover, easily missed deep in the Methods section, is the admission that implementation of the training programme occurred not in a random order, but in order of readiness to participate. This practical approach to rolling out the intervention, starting with hospitals ready to implement, creates a bias in favour of the intervention group. The study compares hospitals that have undergone the intervention with those who had not yet done so—a group, it appears, selected precisely because they were deemed to have a context inimical to the intervention. In the current study, any bias in favour of the intervention from the selection of the hospitals seems to have had little effect. But, the difficulty of evaluating and neutralising context-related barriers remains.

So, is it all terribly difficult? Yes, but there are important positive lessons to be drawn from this 'negative' study. First, the authors are quite right to try to do good science in this field, and we should continue to strive for the highest standards of rigorous research. Retreating to a nihilistic position, maintaining that the complexity of routine practice makes proper evaluation impossible, serves little purpose. We may, however, need to adapt our tactics. The limitations in the evaluations of many teamwork interventions are very difficult to avoid, given the challenges of conducting quasi-experimental studies in a complex social environment where many influences remain completely beyond the control of investigators.

To avoid these sometimes overwhelming challenges, we may need to begin with small proof of principle studies based on a psychology paradigm in well-controlled settings outside the clinical environment. Once an intervention shows clear effects in such studies, pilot studies with 'tinkering' iterative adaptations could follow, using an approach similar to a Quality Improvement paradigm, to illustrate clinical feasibility.[5] If these are essentially attempts to successfully implement what earlier studies demonstrated in principle, there is a perfectly respectable argument for saying they do not need to be controlled. The final step of a large formal controlled trial will still be necessary to confirm that apparent major advances are effective in a range of contexts, and these should be conducted using a mixed methods approach according to a recognised theoretical template. As far as CRM is concerned, the evidence seems to be increasingly clear that it is relatively ineffective in changing clinical outcomes when used alone.[6] However, there is evidence

that its effectiveness, demonstrated again in the current study, in changing staff attitudes and non-technical interactions, may considerably enhance the effects of other quality improvement interventions.[7] Finally, we need to do more work to understand how context influences the outcome of safety interventions, and to engage with those in the business community and elsewhere who have developed coherent theories of organisational change: we need one which works for clinical organisations.

Bill Shankly, one of England's most revered football managers, was once asked whether he regarded football as a matter of life and death. He replied without hesitation "Oh no. It's much more serious than that." We should bear his words in mind next time rocket science is mentioned to us in the context of patient safety.

**Competing interests** None declared.

**Provenance and peer review** Commissioned; internally peer reviewed.

## REFERENCES

1 Kemper, et al. Crew resource management training in the Intensive Care Unit. A multi-site controlled before-after study. BMJ Qual Saf 2015. In press. doi:10.1136/bmjqs-2015-003994
2 Davidoff F, Dixon-Woods M, Leviton L, et al. Demystifying theory and its use in improvement. BMJ Qual Saf 2015;24:228–38.
3 Dekker SW, Hugh TB. A just culture after Mid Staffordshire. BMJ Qual Saf 2014;23:356–8.
4 Neily J, Mills PD, Young-Xu Y, et al. Association between implementation of a medical team training program and surgical mortality. JAMA 2010;304:1693–700.
5 McCulloch P Cook JA, Altman DG, et al. IDEAL Group. IDEAL framework for surgical innovation 1: the idea and development stages. BMJ 2013;346:f3012.
6 Boet S, Bould MD, Fung L, et al. Transfer of learning and patient outcome in simulated crisis resource management: a systematic review. Can J Anaesth 2014;61:571–82.
7 Robertson E, Morgan L, New S, et al. Quality improvement in surgery combining lean improvement methods with teamwork training: a controlled before-after study. PLoS One 2015;10: e0138490.