Using report cards and dashboards to drive quality improvement: lessons learnt and lessons still to learn

Noah M Ivers, 1,2,3 Jon Barrett 3,4,5

More than 50 years of health services research has driven home a core lesson: unintended and inappropriate variations in care are common. 1, 2 Identification of such variation in obstetrics was the impetus for Archie Cochrane to start his work. 3 In this issue of BMJ Quality & Safety, Weiss and colleagues report an intervention developed to address inappropriate variation in aspects of maternal newborn care across Ontario, Canada’s most populous province. 4 The intervention involved systematic collection and analysis of administrative data to assess key quality indicators for all hospital births in the province and provision of this data in a ‘dashboard’ back to hospitals.

Measuring quality of care and comparing this against agreed-upon standards of practice or peer performance (ie, audit) and delivery of the results to healthcare professionals and/or administrators (ie, feedback) is a common quality improvement strategy. 5 Whether referred to as ‘audit and feedback’, ‘report cards’, ‘benchmarking’, ‘practice profiles’ or other synonyms, the underlying rationale for audit and feedback is sound. The large literature evaluating this approach indicates that (1) clinicians are relatively poor at self-assessment, 6 meaning that they tend to pursue continuing professional development or quality improvement in areas of interest (where performance is often already high) rather than areas of greatest need; (2) comparing current performance to a target can drive increased performance in motivated individuals, 7-9 meaning that when desired behaviours can be measured and presented in a formative fashion, 10 health professionals may respond positively to them; and (3) high-performing health systems tend to feature audit and feedback as an evidence-based, scalable and relatively inexpensive strategy to encourage uptake of best practices. 11

The use of dashboards to encourage reflection on quality of care is expanding. In 2009, the National Health Service adopted a maternity dashboard; several countries and institutions have shown varying results when such dashboards are evaluated. 12-14 It is tempting to compare quality indicators described in these dashboards across settings and jurisdictions, but care should be taken to ensure that both numerators and denominators, as well as method of data acquisition, are standardised. Even comparing within a jurisdiction can be fraught: the variation across providers seen in the Ontario dashboard—notable for its strict definitions applied in a standardised, rigorous audit—may partially reflect differences in underlying patient populations. However, the improvement achieved is noteworthy: significant absolute reductions in episiotomies (decrease of 1.5 per 100 women), induction for postdates in women before 41 weeks (decrease of 11.7 per 100 women) and repeat caesarean delivery in low-risk women performed before 39 weeks (decrease of 10.4 per 100 women). Even small absolute improvements in the rates of important healthcare processes (like caesarian sections, as achieved in the project by Weiss and colleagues) can be meaningful and cost-effective when the intervention can be implemented across entire jurisdictions. 15

In the paper by Weiss and colleagues, effect sizes fell within the range expected...
based on the Cochrane review of audit and feedback, which found a median absolute improvement in guideline-concordant care of 4%. The usual concerns over non-randomised studies arise when the intervention produces a large effect size—too good to be true, in some cases. In this case, the rigorous quasi-experimental study by Weiss et al produced a believable effect size consistent with the existing literature. The authors compared their observed effects to contemporaneously measured outcomes observed in other jurisdictions and to quality indicators not included in the dashboard initiative. The collection and analysis of these ‘control indicators’ both from within and without the jurisdiction is a strength vis-à-vis causal attribution. We also complement the authors for allowing enough time to observe the effects. Too frequently, investigators publish evaluations of audit and feedback after only a single iteration of feedback. Indeed, despite an increasing number of trials over time testing interventions that feature audit and feedback after only a single iteration of feedback.

Tentative best practices for the design and delivery of audit and feedback interventions have been published. Prospective research is now needed to test these recommendations.

With this in mind, we encourage those interested in the conduct or evaluation of quality improvement to pursue projects that ask more than just whether audit and feedback might work. It is time now to ask how to make it work best. In this regard, health services organisations delivering audit and feedback can partner with researchers interested in generalisable knowledge regarding how to optimise the effects of this intervention. Tentative best practices for the design and delivery of audit and feedback interventions have been published. Prospective research is now needed to test these recommendations.

Over time, more local obstetricians have come to appreciate the provincial dashboard not as a critique of their professional practice but as a way of encouraging alignment in improvement efforts throughout the system. Now that
the dashboard—and its evidence for impact—is established, there is an opportunity to consider how to apply the approach to address areas of highest potential system and patient impact. For example, preterm birth accounts for 80% of all our perinatal morbidity and mortality. An Ontario alliance for the prevention of preterm birth and stillbirth will soon pursue use of a dashboard to help inform and monitor the effects of a suite of preterm birth and stillbirth prevention interventions. While it is clear from this paper that a dashboard can indeed result in improved quality of obstetrical care over time, a key question is how to complement the dashboard to encourage rapid improvement in targeted indicators. For example, 7 of 17 hospitals in the Southern Ontario obstetrical network have collaborated on an intervention focused on postpartum haemorrhage. Incidence of postpartum haemorrhage within the network hospitals has reduced, but the rate of blood transfusion increased per haemorrhage (unpublished data, shared with permission of Dr Jon Barrett), indicating both the promise of this approach and the need for rigorous evaluation. Thus, next steps for the dashboard could include development and evaluation of co-interventions to support providers in the implementation of key processes for quality improvement.

We envision a day when data from administrative sources can be analyzed rapidly and accurately and then pushed as actionable information in near real time to providers and patients themselves to prompt evidence-based actions. In such a scenario, we could prompt patients and their obstetrical care providers to discuss appropriate treatments options to, for instance, reduce the risk of preterm birth. Meanwhile, evaluations testing such feedback and reminder systems would provide generalizable lessons and support incremental improvement. By showing how data can drive system improvement, Weiss and colleagues provide an encouraging step in the direction of this type of learning health system. Having established baseline effectiveness (and hopefully thereby assuring ongoing funding) the dashboard initiative described by Weiss and colleagues could offer an opportunity to pursue additional research questions that examine the benefits and potential harms of this vision for data-driven quality improvement.

Contributors Both authors contributed equally.

Competing interests None declared.

Provenance and peer review Commissioned; internally peer reviewed.

REFERENCES