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]

[1]Academic Family Health Team, Institute for Health Systems Solutions and Virtual Care, Women's College Research Institute, Women's College Hospital, Toronto, Ontario, Canada
[2]Department of Family and Community Medicine, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada
[3]Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
[4]Division of Maternal-Fetal Medicine, Department of Obstetrics and Gynecology, Sunnybrook Research Institute, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada
[5]Department of Obstetrics and Gynecology, Faculty of Medicine, University of Toronto, Toronto, Ontario, Canada

Correspondence to
Dr Noah M Ivers, Family Practice Health Centre, Women's College Hospital, Toronto, ON M5S1B3, Canada;
noah.ivers@utoronto.ca

Accepted 28 December 2017
Published Online First
9 January 2018

Linked

► http://dx.doi.org/10.1136/bmjqs-2017-007361

Check for updates

To cite: Ivers NM, Barrett J. *BMJ Qual Saf* 2018;**27**:417–420.

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[5] 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.[5] Indeed, despite an increasing number of trials over time testing interventions that feature audit and feedback as a core essential element, few studies generate new insights about how to optimise the effects of this intervention.[16] Tentative best practices for the design and delivery of audit and feedback interventions have been published[17 18]; 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.[19] In this regard, health services organisations delivering audit and feedback can partner with researchers interested in generalisable knowledge regarding how to optimise the intervention—to the benefit of both parties. In such implementation of science laboratories,[20] the health service organisations (and the patients they serve) could benefit from sequential, rigorous evaluation to iteratively improve their intervention, while researchers could access opportunities to test key implementation science hypotheses[21] about the intervention at scale. For example, in the UK, the National Health Service Blood and Transplant National Comparative Audit programme has pursued trials to test ways to enhance the effectiveness of their audit programme, including follow-on support to help organisations act on the feedback.[22] This type of ambitious programme of research, with embedded process evaluations, is most likely to provide the types of insights needed to inform future initiatives.

Each of the (presumed) best practices for audit and feedback will not be feasible in every context. For example, the audit and feedback intervention described in the paper by Weiss and colleagues featured a number of evidence-based elements including repeated, timely data fed back in a non-punitive manner from a trustworthy source.[23] It also featured clear comparators. However, despite achieving changes in the direction desired, it is not clear whether the data were reliably received by those able to take action to improve performance. If an administrative person or even department chief received the data, it is not certain that they would have perceived the message regarding performance as intended nor whether they would have reliably known how to take action on those messages to enable the desired practice changes. It is possible that future work by the team describing embedded process evaluations will unearth details about fidelity and mechanism of action[24] that might suggest ways to enhance the impact of the intervention (which should then be evaluated once again).

Despite the widely held notion that 'you can't manage what you don't measure', measurement often imposes a burden on front-line staff and/or middle managers. Thus, evaluations of interventions such as dashboards and report cards remain necessary to ensure that they achieve concrete improvements in clinical care that outweigh their administrative burden. To be worthwhile, these interventions must measure aspects of care that really matter and facilitate process changes that are likely to improve patient outcomes. A key strength of the dashboard described in the article was the rigorous process undertaken prior to the selection of these indicators. However, our local experience suggests that the health professionals in Ontario most commonly involved in intrapartum care (obstetricians) might prioritise indicators other than those featured in the dashboard. In 2015, obstetricians from 17 hospitals in Ontario created the Southern Ontario obstetric network (http://www.obgyn.utoronto.ca/gta-obs-network). One of the first outputs of this network was a dashboard, derived from the very same data represented in this article. Why did this group of obstetricians see it necessary to form a new dashboard? It is natural for obstetricians to seek feedback on indicators that they believe are both predictive of patient outcomes and within their control to change. For instance, although the evidence against elective caesarean section before 39 weeks is strong, local obstetricians felt their performance on this indicator reflected a lack of flexible access to operating room time, an aspect of the system they perceived as beyond their control (Dr Jon Barrett, personal communication). Some obstetricians also argued that rate of attempted Vaginal Birth After Caesarian (VBAC) was a result of patient choice rather than their effectiveness of their counselling, illustrating the complex interactions between self-efficacy and improvement intentions.[25]

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 analysed 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 generalisable 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

1 Wennberg J, Gittelsohn A. Small area variations in health care delivery. *Science* 1973;182:1102–8.
2 Birkmeyer JD, Reames BN, McCulloch P, *et al*. Understanding of regional variation in the use of surgery. *Lancet* 2013;382:1121–9.
3 Chalmers I, Dickersin K, Chalmers TC. Getting to grips with archie cochrane's agenda. *BMJ* 1992;305:786–8.
4 Weiss D, Dunn SI, Sprague AE, *et al*. Effect of a population-level performance dashboard intervention on maternal-newborn outcomes: an interrupted time series study. *BMJ Qual Saf* 2018;27:425–36.
5 Ivers N, Jamtvedt G, Flottorp S, *et al*. Audit and feedback: effects on professional practice and healthcare outcomes. *Cochrane Database Syst Rev* 2012;6:CD000259.
6 Davis DA, Mazmanian PE, Fordis M, *et al*. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA* 2006;296:1094–102.
7 Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation. A 35-year odyssey. *Am Psychol* 2002;57:705–17.
8 Kluger AN, DeNisi A. The effects of feedback interventions on performance: a historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychol Bull* 1996;119:254–84.
9 Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. *Psychol Bull* 1982;92:111–35.
10 Hysong SJ, Best RG, Pugh JA. Audit and feedback and clinical practice guideline adherence: making feedback actionable. *Implement Sci* 2006;1:1–9.
11 Baker GA, Axler R. *Creating A high performing healthcare system for Ontario: evidence supporting strategic changes in Ontario. Health System Reconfiguration*, 2015:1–40.
12 Gaspar J, Chagas J, Osanan GC, *et al*. Maternal and neonatal healthcare information system: development of an obstetric electronic health record and healthcare indicators dashboard. In: Bursa M, Khuri S, Renda ME, eds. *Information technology in bio- and medical informatics. ITBAM 2013. lecture notes in computer science*. 8060. Berlin, Heidelberg: Springer, 2016.
13 Chandraharan E, Sutton J, Beattie J, *et al*. O164 the role of the 'Maternity Dashboard' in improving patient care: 2 year experience at a tertiary centre. *Int J Gynaecol Obstet* 2009;107:S139.
14 Simms RA, Ping H, Yelland A, *et al*. Development of maternity dashboards across a UK health region; current practice, continuing problems. *Eur J Obstet Gynecol Reprod Biol* 2013;170:119–24.
15 Johri M, Ng ESW, Bermudez-Tamayo C, *et al*. A cluster-randomized trial to reduce caesarean delivery rates in Quebec: cost-effectiveness analysis. *BMC Med* 2017;15:96.
16 Ivers NM, Grimshaw JM, Jamtvedt G, *et al*. Growing literature, stagnant science? Systematic review, meta-regression and cumulative analysis of audit and feedback interventions in health care. *J Gen Intern Med* 2014;29:1534–41.
17 McNamara P, Shaller D, De La Mare D, *et al*. *Confidential physician feedback reports: designing for optimal impact on performance*, 2016. 16-0017-EF.
18 Brehaut JC, Colquhoun HL, Eva KW, *et al*. Practice feedback interventions: 15 suggestions for optimizing effectiveness. *Ann Intern Med* 2016;164:435–41.

19 Ivers NM, Sales A, Colquhoun H, *et al*. No more 'business as usual' with audit and feedback interventions: towards an agenda for a reinvigorated intervention. *Implement Sci* 2014;9:14.
20 Ivers NM, Grimshaw JM. Reducing research waste with implementation laboratories. *Lancet* 2016;388:547–8.
21 Colquhoun HL, Carroll K, Eva KW, *et al*. Advancing the literature on designing audit and feedback interventions: identifying theory-informed hypotheses. *Implement Sci* 2017;12:117.
22 Hartley S, Foy R, Walwyn REA, *et al*. The evaluation of enhanced feedback interventions to reduce unnecessary blood transfusions (AFFINITIE): protocol for two linked cluster randomised factorial controlled trials. *Implement Sci* 2017;12:84.
23 Dunn S, Sprague AE, Grimshaw JM, *et al*. A mixed methods evaluation of the maternal-newborn dashboard in Ontario: dashboard attributes, contextual factors, and facilitators and barriers to use: a study protocol. *Implement Sci* 2016;11:59.
24 Moore GF, Audrey S, Barker M, *et al*. Process evaluation of complex interventions: medical research council guidance. *BMJ* 2015;350:h1258.
25 Williams BW, Kessler HA, Williams MV. Relationship among practice change, motivation, and self-efficacy. *J Contin Educ Health Prof* 2014;34(Suppl 1):S5–10.