



OPEN ACCESS

Reinvigorating stagnant science: implementation laboratories and a meta-laboratory to efficiently advance the science of audit and feedback

JM Grimshaw,^{1,2} Noah Ivers,^{3,4} Stefanie Linklater,¹ Robbie Foy,⁵ Jill J Francis,⁶ Wouter T Gude,⁷ Sylvia J Hysong,^{8,9} on behalf of the Audit and Feedback MetaLab

For numbered affiliations see end of article.

Correspondence to

Dr JM Grimshaw, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa ON K1Y 4E9, Canada; jgrimshaw@ohri.ca

Received 18 May 2018

Revised 8 January 2019

Accepted 12 February 2019

Published Online First

9 March 2019

ABSTRACT

Audit and feedback (A&F) is a commonly used quality improvement (QI) approach. A Cochrane review indicates that A&F is generally effective and leads to modest improvements in professional practice but with considerable variation in the observed effects. While we have some understanding of factors that enhance the effects of A&F, further research needs to explore when A&F is most likely to be effective and how to optimise it. To do this, we need to move away from two-arm trials of A&F compared with control in favour of head-to-head trials of different ways of providing A&F. This paper describes implementation laboratories involving collaborations between healthcare organisations providing A&F at scale, and researchers, to embed head-to-head trials into routine QI programmes. This can improve effectiveness while producing generalisable knowledge about how to optimise A&F. We also describe an international meta-laboratory that aims to maximise cross-laboratory learning and facilitate coordination of A&F research.

GROWING LITERATURE, STAGNANT AUDIT AND FEEDBACK SCIENCE

Audit and feedback (A&F) is a commonly used intervention in quality improvement (QI) programmes. The latest Cochrane systematic review of A&F included 140 randomised trials and found a median absolute improvement of +4% in provider performance with an IQR of +0.5% to +16%. While A&F is generally effective for improving professional practice, there is substantial variation in the observed effects that was only partially explained by the effect modifiers tested in a meta-regression.¹ Cumulative meta-analyses indicated that the effect size (and IQR) became stable in 2003 after 51 comparisons from 30 trials.² By 2011, 47

additional trials of A&F against control were published that did not substantially advance our knowledge. Furthermore, many of these trials did not incorporate A&F features likely to enhance the effectiveness, leading to the suggestion that we have a stagnant science despite growing literature. As Ioannidis *et al* point out ‘although replication of previous research is a core principle of science, at some point, duplicative investigations contribute little additional value’.³

In this paper, we argue that the field has reached this point with A&F research and that further two-arm trials of A&F against (no A&F) control are unlikely to provide additional valuable insights and represent a significant opportunity cost and research waste.⁴ Although we recognise that there are specific circumstances when having a no-intervention control might be justified; for example, if there are only a small number of studies targeting a specific group or setting and where there is a good rationale to think the group or setting could be an effect modifier. Under these circumstances, we would suggest that researchers proposing a no-intervention control should justify their choice.

We argue that future research needs to progress from the question ‘in general does A&F improve healthcare professional practice?’ (it does!) to understanding how to optimise A&F to maximise its effects.⁵ In short, we need to conduct comparative effectiveness of head-to-head trials of different variants of A&F (eg, to address questions such as



© Author(s) (or their employer(s)) 2019. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Grimshaw JM, Ivers N, Linklater S, *et al*. *BMJ Qual Saf* 2019;**28**:416–423.

how should A&F be formatted? how many behaviours should be targeted? what and how many performance comparators should be used? how frequently should A&F be given?). Often, these variants of A&F have little or no cost implications, which means even small improvements in the effects of A&F are likely worthwhile at a population level (eg, if we could improve effectiveness from say 4% to 6% by simply modifying the comparator). We also need head-to-head trials of A&F with and without cointerventions (to explore the incremental benefits of cointerventions) and head-to-head trials of A&F compared with other QI interventions (to determine their comparative effectiveness).

Indeed, there is a rich research agenda that needs further exploration. Brehaut, Colquhoun and colleagues interviewed 28 A&F experts from diverse disciplinary backgrounds to identify testable, theory-informed hypotheses about how to design more effective A&F interventions resulting in 15 'low hanging fruit' suggestions⁶ and 313 specific theory-informed hypotheses that could form the basis of an initial research agenda.⁷

NO MORE BUSINESS AS USUAL – LET'S CHANGE THE MODEL

In the business world, this approach is well recognised and called *radical incrementalism*, whereby a series of small changes are enacted one after the other resulting in radical cumulative changes.⁸ Similar approaches are increasingly being used by behavioural insights units to promote evidence-informed public policy.⁸ As an exemplar of what could be done, we highlight the Reducing Antibiotic Prescribing in Dentistry (RAPiD) trial that evaluated multiple variants of A&F to reduce antibiotic prescriptions in Scottish dental community practices.⁹ Under the umbrella of National Health Service Education in Scotland, the investigators used a partial factorial design to randomise all 795 dental practices to four different factors that included comparisons of A&F versus control, A&F with or without actionable messages, A&F with or without a regional performance comparator and A&F provided two or three times over a 12-month period. (There were no trials of A&F in community dental practices in the 2012 Cochrane review justifying an A&F vs no-intervention comparison in the trial; although notably, the trial used an unequal randomisation process (4:1) in recognition of the likely benefits from A&F.) The trial demonstrated that A&F led to a 6% reduction in antibiotic prescriptions and that further significant reductions were observed if A&F was provided with actionable messages and a comparator. However, providing feedback three times rather than two times was not beneficial.

The switch from two-arm trials of A&F versus control to head-to-head trials of different variants of A&F has significant implications for the design and conduct of trials. In particular, when we evaluate

different variants of A&F head-to-head trials, we are interested in the incremental benefit of the design modifications above the general beneficial effect of A&F, suggesting that the differential effects are likely to be small. Hence, trials need to be powered for small effects which require large sample sizes that may be difficult to achieve in typical 'one-off' A&F trials where researchers need to recruit participating healthcare organisations and professionals. Furthermore, we envision that it may be difficult to persuade a research funder to fund a very large trial to detect a relatively small (but potentially important incremental) effect. Hence, we are sceptical about the feasibility of conducting head-to-head trials as one-off projects.

TOWARDS IMPLEMENTATION LABORATORIES

One solution is to establish *implementation laboratories*.⁴ Increasingly, healthcare systems and/or organisations are providing A&F to healthcare professionals as part of their routine, system-wide QI activities. Those designing such A&F programmes must make a myriad of decisions about its content, format and delivery. Frequently, such programmes do not maximally use current theories and evidence about how to optimise their programmes, given service pressures to 'do something' and lack of access to experts with theoretical and empirical content expertise. As a result, A&F programmes delivered in service settings are unlikely to be optimised. Under these circumstances, there are opportunities for researchers and healthcare system partners to conduct collaborative research both to improve the effectiveness of the specific A&F programme and to produce generalisable knowledge about how to optimise A&F that will be relevant to other A&F programmes.

Consider a healthcare system partner that has an existing A&F programme (Baseline A&F), but recognises the arbitrary nature of many of its design choices and is interested in exploring how to enhance the effectiveness of its A&F programme. It could partner with researchers to develop A&F B and conduct a head-to-head trial of its Baseline A&F against A&F B (figure 1) informed by current empirical and theoretical insights^{10 11} and utilising human-centred design approaches.¹² Given that the healthcare system partner is delivering its A&F programme at scale to healthcare organisations and/or professionals, it is likely that the trial will have a much larger sample size than a trial mounted by researchers who would need to recruit healthcare organisations and/or professionals. Furthermore, the costs of conducting the trial are likely to be modest given that the healthcare system partner is already committed to delivering A&F at scale and is routinely collecting data that could be used for outcome assessment (so the marginal costs will largely relate to the design of A&F B and trial analysis). If the trial found that A&F B was more effective than the Baseline A&F,

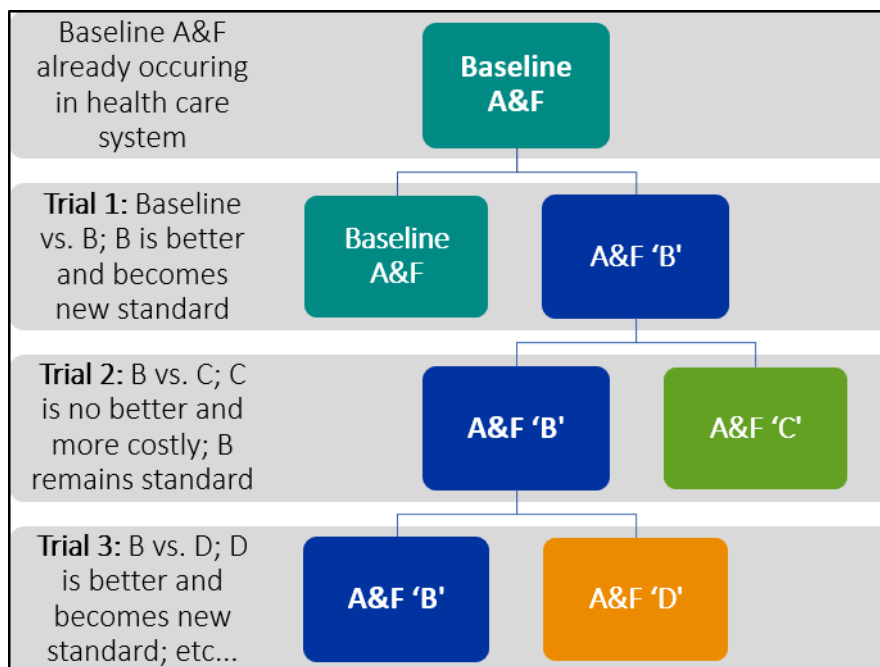


Figure 1 Schematic representation of sequential trials in an implementation science laboratory. A&F, audit and feedback.

then A&F B would become the standard version of A&F for the healthcare system partner. However, it is unlikely that all possible variations of interest in A&F would be tested in a single trial (even one as ambitious as the RAPID trial), suggesting a need for the implementation laboratory to conduct sequential trials, programmatically testing a range of variants of A&F (eg, informed by Brehaut's 15 best practice suggestions⁶). Hence, the healthcare system partner might test A&F B (now the standard practice) against A&F C, perhaps finding that A&F C is less effective, costlier to deliver or less acceptable to the target healthcare organisations and/or professionals and so it is dropped (noting that the healthcare system partner and implementation science have learnt from this no-effect study). Following this, the healthcare system partner might test A&F B against A&F D and so on. Thus, A&F implementation laboratories involve healthcare system partners in continuous improvement using rigorous methods to identify more effective A&F variants that can be routinely and sustainably embedded into their ongoing A&F programmes. Our example above uses a series of sequential two-arm trials. However, the same approach could be applied to other study designs including multi-arm trials, factorial designs or balanced incomplete block designs.

Randomised trials are considered the 'gold standard' for determining *causal description* ('did strategy *x* lead to outcome *y*?')¹³ because of the following: intervention effects are modest; there is an incomplete understanding of potential confounders and effect modifiers; and there are significant opportunity costs for healthcare systems if ineffective (or

inefficient) implementation strategies are adopted. Although causal description may be sufficient for healthcare system partner decisions ('B was better than A, so let's adopt B'), it provides an incomplete picture for advancing implementation science. Given our poor understanding of the mechanisms of action, potential confounders and effect modifiers of implementation strategies such as A&F, causal description alone is usually inadequate to interpret study findings and determine the generalisability of evaluation findings. However, the conduct of randomised trials also provides an opportunity to embed further studies to determine the *causal explanation* ('how, why and under what circumstances did strategy *x* lead to outcome *y*?')¹³ to enhance the informativeness and generalisability of evaluations including the following: *study design elements* (eg, multi-arm trials, factorial designs that allow simultaneous multiple comparison^{9 14}); *fidelity substudies* (to determine whether the content of interventions was delivered as designed, to determine whether participants received and/or opened their feedback, to determine how participants reacted to the feedback¹⁵); *mechanistic substudies* (theory-based process evaluations to determine whether interventions activated the hypothesised mediating pathways and if so, was this sufficient to lead to practice change¹⁶); *qualitative process evaluations* (to understand participants' experiences of being in a trial and whether the intervention was acceptable to them^{17 18}) and *economic evaluations* (to determine the cost-effectiveness of different A&F strategies¹⁹).

There are also opportunities for analytical approaches (eg, subgroup analyses and hierarchical

regressions) to explore potential effect modifiers in inner context variables (eg, healthcare organisation or professional characteristics)²⁰ that allow exploration of whether these factors are effect modifiers of the A&F (eg, Do teaching hospitals respond to A&F similarly to community hospitals? Is baseline performance an effect modifier for A&F?) providing valuable information for the healthcare system partner (Should we vary our feedback approach by recipient characteristics?) and for advancing implementation science. Time series analyses could also be undertaken to evaluate the effects of individual A&F variants (eg, Did Baseline A&F lead to practice change?)²¹ and to assess learning and decay effects of interventions.²²

Although we have emphasised A&F design here, this approach could be used for other aspects of the A&F process, for example evaluating the incremental benefit of cointerventions alongside A&F (such as educational meetings or computerised prompts) or evaluating approaches to improve the receipt and uptake of A&F.¹⁰

Implementation laboratories enable learning healthcare systems

Such implementation laboratories are an extension of the concept of learning healthcare systems, where the learning encompasses generalisable concepts for implementation science. As in the physical sciences, the laboratory provides an opportunity for experimentation and observation—in this case—of the processes playing a role in the uptake of best practices. However, an implementation science laboratory is developed with the express purpose of providing applied evidence for the health system partner; the engagement of the healthcare system partner throughout the process increases the likelihood that successful A&F variants are rapidly incorporated into the routine practices of the organisation in a sustainable way. The embedded nature of A&F laboratories also increases the real-world generalisability of any findings given that participants receiving A&F from health system partners are likely broadly representative of healthcare participants within that jurisdiction. The collaboration allows healthcare system partners to access researchers and their expertise and allows researchers to access experienced healthcare system personnel who have substantial tacit and practical knowledge about providing A&F within their system.

By pursuing a research question as prioritised by the healthcare system partner, researchers can be sure that the topic is one that truly needs addressing, and that the measured outcomes are the primary ones of interest. The quality of the research design and methods can be optimised by implementation researchers who are likely to have more experience with designing trials that are methodologically robust

(addressing causal description) and that embed other forms of enquiry to maximise the information yielded from any trial (addressing causal explanation). This partnership capitalises on the expertise of both the healthcare system partner and the implementation researchers to improve the quality, effectiveness and applicability of the trial.

Some practical considerations

Many practical issues need to be considered when establishing an A&F laboratory. We highlight some that are likely to be common across settings.

Need for stability in healthcare system partners

Productive relationships between the healthcare system partner and the research team can be disrupted by reorganisation or decommissioning of healthcare system partners responsible for QI activities. Our practical experience suggests a minimum of a mutual 3-year commitment to establish a viable and functioning laboratory.

Team composition

Teams need the right combination of experts, boundary spanners and facilitators. It is important to include both senior leadership, who can ensure alignment of activities and promote the value of the implementation laboratory within the organisation, and team members actively involved in the design and delivery of the current A&F programmes. Implementation science is an interdisciplinary team sport; a diverse range of research disciplines can maximise health system partner and scientific benefits.

Clarifying roles and responsibilities

Time must be invested upfront to build trust, clarify roles and expectations and develop clear governance arrangements to manage the laboratory.^{23 24} Although such arrangements will depend on local circumstances, resources and norms, it is essential that they confer a sense of joint ownership of and responsibility for the laboratory between the health system partner and the research team.

Identifying shared priorities

It is difficult to fulfill the principles of the implementation laboratory if the researchers seek to answer questions that do not match the goals, timeline or budget of the host organisation, or if the host organisation requires a 'positive' result. We suggest that researchers develop a shortlist of best practices and research opportunities for discussion^{6 7} and negotiate the best fit with healthcare system partner needs.

Aligning timelines and activities

Healthcare system partners require lag time to embed the testing of variants into their A&F cycles, typically to manage practical logistics such

as programming and preparing current and variant A&F and to ensure alignment with audit cycles. It is also worth identifying any potentially competing QI activities that may influence trial design, timing or interpretation.

Optimising design

Implementation laboratories offer opportunities to apply state-of-the-art approaches to A&F design, including robust theory-driven and user-centred design processes which require time and resources usually unavailable to healthcare organisations. In A&F laboratories, the lag time required by the healthcare system partner provides a window of opportunity for such optimisation.

Ethical considerations

Given that A&F laboratories aim to produce generalisable knowledge (alongside demonstrable service improvements), they require adequate research ethics oversight and approval before commencing.²⁵ Healthcare organisations and professionals who directly receive the A&F should usually be considered research participants but not patients who are only indirectly affected by the feedback. Altered or waived consent from research participants may be feasible if the intervention and study procedures could be considered minimum risk (which is likely given that variants of A&F are being tested), no identifiable private information about participants is released to the study team (which seems feasible to ensure), consent would not be feasible (which is likely given the large number of healthcare organisations and professionals) or if awareness of the study might invalidate its results.

Funding

Although the costs of implementation laboratories are likely to be significantly less than the costs of conducting one-off research projects, there will be additional costs for the healthcare system partner from designing and delivering two different variants of A&F. Nevertheless, we believe that research funders are more likely to fund a programme of research to optimise A&F working with and leveraging resources of healthcare system partners.

WHY STOP THERE? LET'S BUILD THE META-LABORATORY

Individual implementation laboratories cannot adequately address the remaining uncertainties around A&F for a number of reasons. First, an individual laboratory will only be able to evaluate a limited number of A&F variants. Second, implementation science is interested in the generalisability of findings from individual laboratories suggesting the need for replication across other settings. Finally, while inner context variables are likely to vary

within an individual laboratory, the outer context variables (eg, governance, accountability and reimbursement arrangements)²⁰ are likely to be common across healthcare organisations and/or professionals receiving A&F. As such, individual laboratories cannot inform us about the likely impact of these outer context variables on the effectiveness of A&F.

Fully addressing the remaining uncertainties around A&F may require a 'meta-laboratory' approach: a coordinated set of implementation laboratories, in which findings from interventions tested in one context could inform decisions for another context, to facilitate a cumulative science in the field. This approach would allow for shared expertise, opportunities for planned replication to explore replicability and outer context issues and would build an international community of healthcare system partners with shared interests. The meta-laboratory would contribute to the reduction of research waste through this shared expertise and planned replication as study results and insights from one implementation laboratory can be used to inform study design and conduct for another implementation laboratory.

A meta-laboratory could form a collaborative space encouraging linkages between researchers undertaking innovative theoretical⁷ or developmental work²⁶⁻²⁸ with researchers in implementation laboratories to ensure the timely translation of promising innovations into large-scale evaluations within implementation laboratories. Meta-laboratories could provide state-of-the-art, regularly updated resources for both healthcare systems planning A&F activities and researchers planning future research studies (including promoting full and transparent reporting of laboratories activities).

Announcing the Audit and Feedback MetaLab

There are an increasing number of A&F laboratories globally (see [table 1](#)). We have established the Audit and Feedback MetaLab involving researchers and laboratories from five countries (Australia, Canada, the Netherlands, the UK and the USA). To date, the Audit and Feedback MetaLab has held annual international symposia (Ottawa 2016, Leeds 2017, Toronto 2018 and another planned in Amsterdam 2019) each involving a scientific meeting and an open meeting with training activities for healthcare systems and organisations delivering A&F programmes. We have established a website (www.ohri.ca/auditfeedback) providing a 'one-stop shop' of resources about designing and evaluating A&F.

Implementation laboratories and meta-laboratories: a generalisable approach?

We recognise that A&F is only one of many implementations strategies that healthcare systems are interested in to improve healthcare delivery. However, there are similar challenges faced by healthcare systems interested in using other implementation strategies (eg, educational

Table 1 Participating implementation laboratories in the A&F MetaLab

Laboratory	Health system partner	A&F interventions tested	Evaluation design
AFFINITIE—enhanced A&F interventions to increase the uptake of evidence-based transfusion practice	National Comparative Audit of Blood Transfusion, UK	<ul style="list-style-type: none"> ▶ Enhanced feedback report content (eg, streamlining formatting, highlighting gaps between current and recommended practice, and suggesting practical action plans) ▶ Enhanced ‘follow on’ support for local action (ie, a web-based toolkit for use by hospital transfusion teams and telephone support to increase engagement) 	<ul style="list-style-type: none"> ▶ Two linked and sequential cluster randomised 2x2 factorial trials with parallel process and economic evaluations ▶ 147 and 137 hospitals
BORN MND—Better Outcomes Registry & Network Ontario- Maternal Newborn Dashboard	Better Outcomes Registry & Network, Ontario, Canada	<ul style="list-style-type: none"> ▶ A&F provided via the MND which consists of near real-time hospital-specific feedback about six key performance indicators, peer comparison data, a visual display of practice gaps and benchmarks to provide direction for practice change ▶ Evidence summaries are also provided for each of the key performance indicators. 	<p>Mixed methods:</p> <ul style="list-style-type: none"> ▶ Interrupted time series analysis using population-based data from all maternal newborn hospitals in Ontario covering a 3- year period before implementation of the MND and 2.5 years after ▶ Provincial survey of all hospitals and site visits with a diverse sample of hospitals to identify barriers and facilitators to use of the MND and determine factors that influence uptake to trigger and facilitate practice change ▶ 96 hospitals
Implementing Goals of Care Conversations with Veterans in Veteran Affairs (VA) Long Term Care Settings Health system partner:	VA Corporate Data Warehouse; VA National Centre for Ethics in Healthcare, USA	<ul style="list-style-type: none"> ▶ Electronically generated feedback reports providing feedback on performance in documenting goals of care conversations, coupled with learning collaboratives ▶ Different formats and delivery of feedback reports will be tested 	<ul style="list-style-type: none"> ▶ Quasi-experimental mixed-methods design using interrupted time series
NICE foundation electronic quality dashboard with action implementation toolbox to enhance quality of pain management in Dutch intensive care units	NICE quality registry, the Netherlands	<ul style="list-style-type: none"> ▶ Online quality dashboard ▶ Intervention group additionally receives access to an integrated action implementation toolbox of suggested actions and materials to further support the development and management of action plans 	<ul style="list-style-type: none"> ▶ Two-arm cluster randomised controlled trial with the same 21 ICUs allocated to feedback without or feedback with action implementation toolbox group ▶ Both arms receive telephone support to increase engagement ▶ 21 intensive care units; 72 individual professionals
OHIL—Ontario Healthcare Implementation Laboratory: Improving the impact of Practice Reports in nursing homes	Health Quality Ontario, Canada	<ul style="list-style-type: none"> ▶ Variation in comparators featured in the feedback reports (eg, overall provincial average prescribing rate versus the 25th percentile prescribing rate) ▶ Variation in the framing of the content as positive or negative (eg, recipients may be told they have prescribed potentially harmful medications to 15% of their patients versus avoiding prescription-related harms in 85% of their patients) 	<ul style="list-style-type: none"> ▶ Cluster randomised 2x2 factorial trial with embedded mixed-method process evaluation ▶ Prescribers receive three iterations of the report ▶ Prescribers receive surveys and may also participate in interviews to assess mechanism of action and additional barriers to use of reports to make changes
TRIADS—Translation Research in a Dental Setting	NHS Education Scotland	<ul style="list-style-type: none"> ▶ Enhanced feedback report content (eg, changing comparators) ▶ Enhanced persuasive messaging using behaviour change techniques ▶ Timing of feedback ▶ Impact of educational outreach visits in addition to audit ▶ Enhanced audits using behaviour taxonomy approaches 	<ul style="list-style-type: none"> ▶ Cluster randomised (sometimes factorial trials) are conducted whenever possible ▶ Time series are also used to monitor impact of national audits.

A&F, audit and feedback; ICUs, intensive care units; MND, Maternal Newborn Dashboard; NHS, National Health Service; NICE, National Intensive Care Evaluation.

outreach) and by healthcare organisations interested in using Plan-Do-Study-Act cycles to improve quality. The model of implementation laboratories and an international meta-lab could be equally applied to understand how to optimise these strategies. There are also examples of implementation laboratories addressing implementation challenges in specific sectors or jurisdictions (the Translation Research in a Dental Setting laboratory which is developing and evaluating implementation strategies alongside guideline development for community dentists in Scotland²⁹ and the Hunter Valley

Population Health laboratory which uses state-of-the-art implementation science approaches to design and evaluate population health interventions in Australia).³⁰

CONCLUSIONS

Despite widespread application, we are only beginning to understand how, when and why implementation interventions such as A&F might work best. Publications describing evaluations of these interventions may provide a snapshot of a local change initiative,

but rarely offer the sort of detail needed to ascertain potentially generalisable findings regarding mechanisms for achieving change.³¹ We urgently need to develop the scientific basis for A&F (and other QI interventions) informed by behavioural and system theory and using rigorous, large-scale evaluations. Implementation science laboratories provide an opportunity to reinvigorate the field by supporting healthcare system partners to optimise their own A&F programmes while efficiently advancing the scientific basis of A&F. This paper demonstrates that productive implementation science laboratories are feasible and that they can be mutually beneficial for both implementation scientists and health system partners. We hope it is received as a call to action for those organisations involved in the design and delivery of A&F, for scientists interested in driving forward related scientific findings, and for research funders interested in maximising impact in both these domains. The Audit and Feedback MetaLab can facilitate communication and collaboration across individual laboratories and act as a resource for healthcare systems planning A&F programmes while providing evidence on the generalisability of scientific evidence found at local A&F laboratories.

Author affiliations

¹Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, Ontario, Canada

²Department of Medicine, University of Ottawa, Ottawa, Ontario, Canada

³Women's College Research Institute, Women's College Hospital, Toronto, Ontario, Canada

⁴Department of Family Medicine and Community Medicine, University of Toronto, Toronto, Ontario, Canada

⁵Leeds Institute of Health Sciences, University of Leeds, Leeds, UK

⁶Health Services Research and Management Division, City University of London, London, UK

⁷Department of Medical Informatics, Academic Public Health Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

⁸Center for Innovations in Quality Effectiveness and Safety, Michael E. DeBakey Veterans Affairs Medical Center, Houston, Texas, United States

⁹Section of Health Services Research, Department of Medicine, Baylor College of Medicine, Houston, Texas, United States

Correction notice This article has been corrected since it published Online First. The 'on behalf of' clause has been reinstated as it got omitted in the previous version.

Acknowledgements Jeremy Grimshaw holds a Canada Research Chair in Health Knowledge Transfer and Uptake. Noah Ivers is supported by the Women's College Hospital Chair in Implementation Science and a Clinician Scientist award from the Department of Family and Community Medicine at the University of Toronto.

Collaborators On behalf of the Audit & Feedback MetaLab. Sarah Alderson, Sylvain Boet, Jamie Brehaut, Benjamin Brown, Jan Clarkson, Heather Colquhoun, Nicolette de Keizer, Laura Desveaux, Gail Dobell, Sandra Dunn, Amanda Farrin, Robbica C Foy, Jill J Francis, Anna Greenberg, Jeremy M Grimshaw, Wouter T Gude, Suzanne Hartley, Tanya Horsley, Sylvia J Hysong, Noah M Ivers, Zach Landis-Lewis, Stefanie Linklater, Jane London, Fabiana Lorencatto, Susan Michie, Denise O'Connor, Niels Peek, Justin Pesseau, Craig Ramsay, Anne E. Sales, Ann Sprague, Simon Stanworth, Michael Sykes, Monica Taljaard, Kednapa Thavorn, Mark Walker, Rebecca Walwyn, Debra Weiss, Thomas A Willis and Holly Witteman.

Contributors This paper is the product of the A&F MetaLab. The named authors are the current Steering Group of the MetaLab. The group authors have contributed consistently to the development of these ideas over the last 3 years during several national and international meetings (Ottawa Canada 2016, Leeds UK 2017, Calgary Canada 2017). All authors have been involved in drafting of the paper and have approved the final version. JMG is the guarantor for the paper.

Funding This study was funded by the Canadian Institutes of Health Research (FDN-143269).

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0>

REFERENCES

- Ivers N, Jamtvedt G, Flottorp S, *et al.* Audit and feedback: effects on professional practice and healthcare Outcomes. *Cochrane Database Syst Rev* 2012;154.
- 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.
- Ioannidis JPA, Greenland S, Hlatky MA, *et al.* Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014;383:166–75.
- Ivers NM, Grimshaw JM. Reducing research waste with implementation laboratories. *The Lancet* 2016;388:547–8.
- 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. *Implementation Sci* 2014;9.
- Brehaut JC, Colquhoun HL, Eva KW, *et al.* Practice feedback interventions: 15 suggestions for optimizing effectiveness. *Ann Intern Med* 2016;164:435–41.
- Colquhoun H, Michie S, Sales A, *et al.* Reporting and design elements of audit and feedback interventions: a secondary review. *BMJ Qual Saf* 2017;26:54–60.
- Halpern D, Mason D, Radical incrementalism MD. Radical incrementalism. *Evaluation* 2015;21:143–9.
- Elouafkaoui P, Young L, Newlands R, *et al.* An audit and feedback intervention for reducing antibiotic prescribing in general dental practice: the rapid cluster randomised controlled trial. *PLoS Med* 2016;13:e1002115.
- Gould NJ, Lorencatto F, Stanworth SJ, *et al.* Application of theory to enhance audit and feedback interventions to increase the uptake of evidence-based transfusion practice: an intervention development protocol. *Implementation Sci* 2014;9.
- Ivers NM, Desveaux L, Pesseau J, *et al.* Testing feedback message framing and comparators to address prescribing of high-risk medications in nursing homes: protocol for a pragmatic, factorial, cluster-randomized trial. *Implementation Science* 2017;12.
- Witteman HO, Pesseau J, Nicholas Angl E, *et al.* Negotiating tensions between theory and design in the development of

- Mailing for people recovering from acute coronary syndrome. *JMIR Hum Factors* 2017;4:e6.
- 13 Shadish WR, Cook TD, Experimental CDT. *Quasi-experimental designs for generalized causal inferences*. Houghton Mifflin Company, 2002.
 - 14 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.
 - 15 Lorencatto F, Gould NJ, McIntyre SA, *et al*. A multidimensional approach to assessing intervention fidelity in a process evaluation of audit and feedback interventions to reduce unnecessary blood transfusions: a study protocol. *Implement Sci* 2016;11.
 - 16 Ramsay CR, Thomas RE, Croal BL, *et al*. Using the theory of planned behaviour as a process evaluation tool in randomised trials of knowledge translation strategies: a case study from UK primary care. *Implementation Sci* 2010;5.
 - 17 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. *Implementation Science* 2015;11.
 - 18 Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. *BMC Health Serv Res* 2017;17.
 - 19 Sculpher M. Evaluating the cost-effectiveness of interventions designed to increase the utilization of evidence-based guidelines. *Family Practice* 2000;17:26S–31.
 - 20 Damschroder LJ, Aron DC, Keith RE, *et al*. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implementation Sci* 2009;4.
 - 21 Guthrie B, Kavanagh K, Robertson C, *et al*. Data feedback and behavioural change intervention to improve primary care prescribing safety (EFIPPS): multicentre, three arm, cluster randomised controlled trial. In: *BMJ (clinical research)*. , 2016: 354, i4079.
 - 22 Ramsay CR, Eccles M, Grimshaw JM, *et al*. Assessing the long-term effect of educational reminder messages on primary care radiology referrals. *Clin Radiol* 2003;58:319–21.
 - 23 Donsbach JS, Tannenbaum SI, Mathieu JE, *et al*. A review and integration of team composition models. moving toward a dynamic and temporal framework. *Journal of Management* 2014;40:130–60.
 - 24 Salas E, Shuffler ML, Thayer AL, *et al*. Understanding and improving teamwork in organizations: a Scientifically based practical guide. *Hum Resour Manage* 2015;54:599–622.
 - 25 Weijer C, Grimshaw JM, Eccles MP, *et al*. The Ottawa statement on the ethical design and conduct of cluster randomized trials. *PLoS Med* 2012;9:e1001346.
 - 26 Gude WT, van Engen-Verheul MM, van der Veer SN, *et al*. How does audit and feedback influence intentions of health professionals to improve practice? a laboratory experiment and field study in cardiac rehabilitation. *BMJ Qual Saf* 2017;26:279–87.
 - 27 Gude WT, Roos-Blom M-J, van der Veer SN, *et al*. Electronic audit and feedback intervention with action implementation toolbox to improve pain management in intensive care: protocol for a laboratory experiment and cluster randomised trial. *Implement Sci* 2017;12.
 - 28 Gude WT, van der Veer SN, van Engen-Verheul MM, *et al*. Inside the black box of audit and feedback: a laboratory study to explore determinants of improvement target selection by healthcare professionals in cardiac rehabilitation. *Stud Health Technol Inform* 2015;216:424–8.
 - 29 Clarkson JE, Ramsay CR, Eccles MP, *et al*. The translation research in a dental setting (triads) programme protocol. *Implementation Science* 2010;5.
 - 30 Wolfenden L, Yoong SL, Williams CM, *et al*. Embedding researchers in health service organizations improves research translation and health service performance: the Australian Hunter New England population health example. *J Clin Epidemiol* 2017;85:3–11.
 - 31 Glasziou P, Altman DG, Bossuyt P, *et al*. Reducing waste from incomplete or unusable reports of biomedical research. *The Lancet* 2014;383:267–76.