

# Implementating evidence-based practices to improve global health

Dr Jeremy Grimshaw

Senior Scientist, Ottawa Hospital Research Institute

Professor, University of Ottawa

jgrimshaw@ohri.ca @GrimshawJeremy

16<sup>th</sup> May 2024



The Ottawa  
Hospital  
Research Institute

L'Hôpital  
d'Ottawa  
Institut de recherche

**Inspired** by research. **Inspiré** par la recherche.  
**Driven** by compassion. **Guidé** par la compassion.

# What gets me out of bed each morning?

- Consistent evidence of failure to translate research findings into clinical practice
- 30-40% patients do not get treatments of proven effectiveness
- 20–25% patients get care that is not needed or potentially harmful
- Suggests that implementation of research findings is fundamental challenge for healthcare systems to optimise care, outcomes and costs

Schuster, McGlynn, Brook (1998). *Milbank Memorial Quarterly*  
Grol R (2001). *Med Care*

# Implementation science

- Key element of implementation science is evaluating the effectiveness and efficiency of implementation programs
- Randomised controlled trials of implementation programs desirable because:
  - Effects of implementation programs small to modest
  - Incomplete understanding of confounders (and non-specific effects)
  - Significant opportunity costs if incorrect conclusions drawn

# Implementation science

## Behavioural perspective

- Implementation depends on behaviour
  - Citizens, patients, health professionals, managers, policy makers
- To improve care, we need to change behaviour
- To change behaviour, it helps to understand determinants of current behaviour and how behaviour changes

# 1<sup>st</sup> generation trials (1976-2005)

Intervention development



# 1<sup>st</sup> generation trials

- Technical – clustering ignored, small numbers of units, unrealistic effect sizes, unit of analysis remain common
- Design – majority are two arm trials (intervention vs control)
- Intervention – little rationale provided for the choice of intervention, few explicitly theory based, insufficient feasibility testing
- Limited efforts to explore causal mechanisms of any observed changes
- Economic evaluation – largely ignored
- Reporting – insufficient details of context, intervention, and methods

# 1<sup>st</sup> generation trial – NEXUS (2001)

## ARTICLES

### Effect of audit and feedback, and reminder messages on primary-care radiology referrals: a randomised trial

Martin Eccles, Nick Steen, Jeremy Grimshaw, Lois Thomas, Paul McNamee, Jennifer Soutter, John Wilsdon, Lloyd Matowe, Gillian Needham, Fiona Gilbert, Senga Bond

#### Summary

**Background** Radiological tests are often used by general practitioners (GPs). These tests can be overused and contribute little to clinical management. We aimed to assess two methods of reducing GP requests for radiological tests in accordance with the UK Royal College of Radiologists' guidelines on lumbar spine and knee radiographs.

**Methods** We assessed audit and feedback, and educational reminder messages in six radiology departments and 244 general practices that they served. The study was a before-and-after, pragmatic, cluster randomised controlled trial with a 2x2 factorial design. A random subset of GP patients' records were examined for concordance with the guidelines. The main outcome measure was number of radiograph requests per 1000 patients per year. Analysis was by intention to treat.

**Findings** The effect of educational reminder messages (ie, the change in request rate after intervention) was an absolute change of -1.53 (95% CI -2.5 to -0.57) for lumbar spine and of -1.61 (-2.6 to -0.62) for knee radiographs, both relative reductions of about 20%. The effect of audit and feedback was an absolute change of -0.07 (-1.3 to 0.9) for lumbar spine of 0.04 (-0.95 to 1.03) for knee radiograph requests, both relative reductions of about 1%. Concordance between groups did not differ significantly.

**Interpretation** 6-monthly feedback of audit data is ineffective but the routine attachment of educational reminder messages to radiographs is effective and does not affect quality of referrals. Any department of radiology that handles referrals from primary care could deliver this intervention to good effect.

#### Introduction

General practitioners (GPs) can overuse radiological tests, particularly lumbar spine<sup>1,2</sup> and knee radiographs.<sup>3</sup> Such tests are frequently of little clinical use. Guidelines for use of these investigations are in the UK Royal College of Radiologists' publication *Making the best use of a radiology department*.<sup>4</sup> However, few studies have been done of interventions designed to change GPs' behaviour. Although these studies showed that GPs altered their use of radiological tests, they were badly designed,<sup>5,6</sup> used inappropriate analysis,<sup>7</sup> had short duration of follow-up,<sup>8</sup> or omitted cost considerations.<sup>9</sup> Grol<sup>10</sup> and Lomas<sup>11</sup> have summarised the theory of how to change doctors' behaviour, and Oxman and colleagues<sup>12</sup> have reviewed the effectiveness of interventions. Specific prompts at the time of consultation are a powerful strategy<sup>13</sup> and have been shown to alter GPs' behaviour—eg, when referring patients for infertility investigations<sup>14</sup>—but the effect of the widely-used strategy of audit and feedback is not so certain.<sup>15,16</sup>

We assessed two methods (audit and feedback, and educational messages) of reducing GPs' requests for radiological tests in accordance with the UK Royal College of Radiologists' guidelines. Our hypothesis was that either intervention alone would be more effective than a control and that both interventions together would be more effective than either alone.

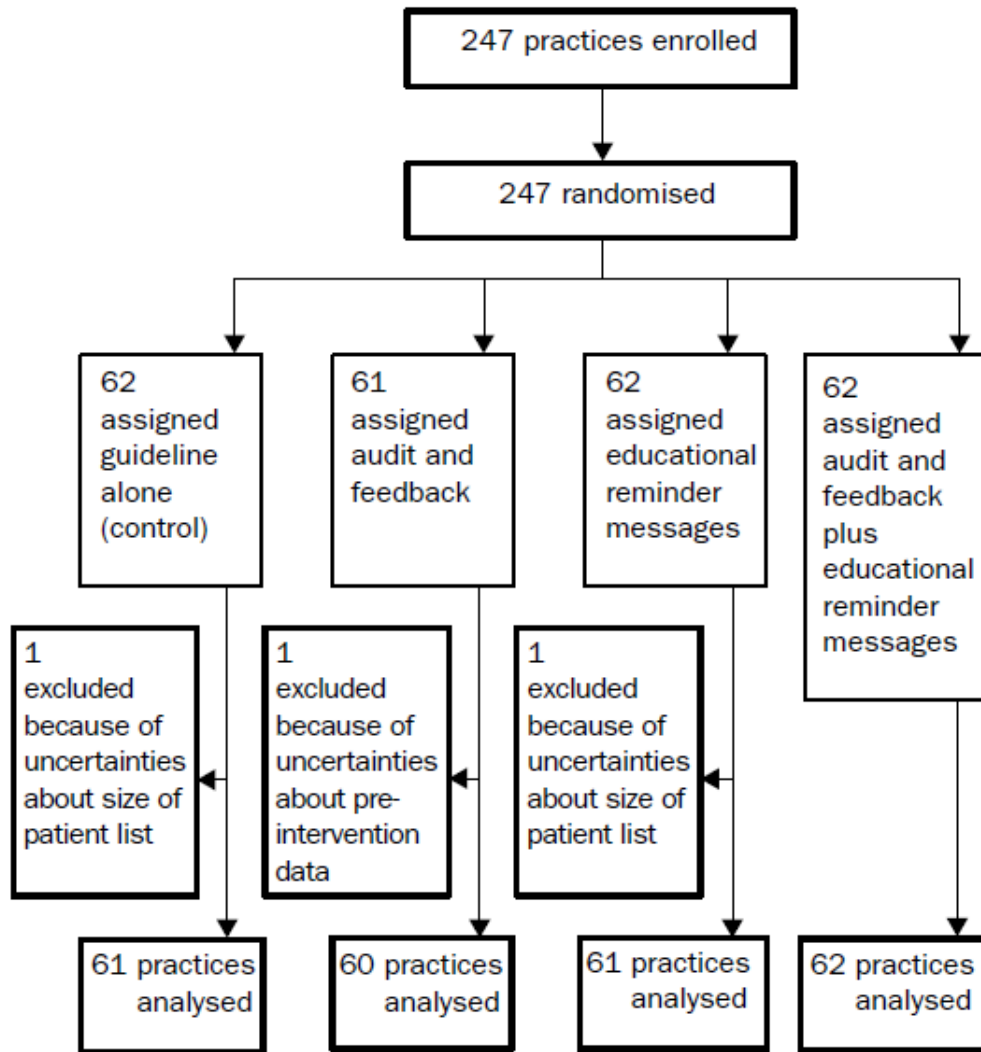
#### Methods

##### Study design

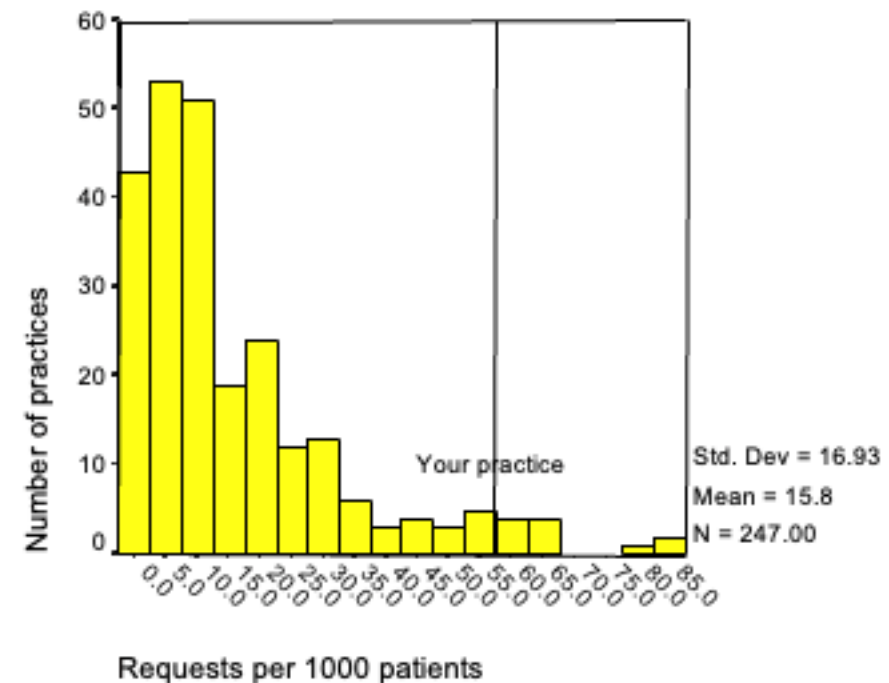
The study was based in six radiology departments in the north-east of England and Scotland and in GPs' surgeries (practices) that referred patients exclusively to them. The study was a before-and-after, pragmatic, cluster randomised controlled trial, with a 2x2 factorial design—practices were the units of randomisation and analysis.<sup>17</sup> Randomisation, stratified by radiology department and practice size, was done by the study

- Pragmatic 2 x 2 factorial RCT of audit and feedback and educational messages to 240 general practices in the North-East of England and Scotland to reduce unnecessary lumbar spine and knee x-rays.

# 1<sup>st</sup> generation trial – NEXUS (2001)



Educational message - “in adults with knee pain, without serious locking or restriction in movement, radiograph is not routinely indicated”





# 1<sup>st</sup> generation trial – NEXUS (2001)

The effect of educational reminder messages was an absolute change of  $-1.53$  (95% CI  $-2.5$  to  $-0.57$ ) for lumbar spine radiographs and of  $-1.61$  ( $-2.6$  to  $-0.62$ ) for knee radiograph requests; these estimates are both relative reductions of about 20%. The effect of audit and feedback was an absolute change of  $-0.07$  ( $-1.3$  to  $0.9$ ) for lumbar spine and  $0.04$  ( $-0.95$  to  $1.03$ ) knee radiograph requests. Relative reductions were about 1% (knee) and almost no change (lumbar spine). For both types of radiograph, interaction between interventions was not significant—ie, there was no increased effect of receiving both interventions.

# 1<sup>st</sup> generation trials

- 1000s of trials -> substantial evidence base
- Many interventions successful but significant unexplained variations in effects
- Not efficiently advancing implementation science



**Audit and feedback: effects on professional practice and healthcare outcomes (Review)**

Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, O'Brien MA, Johansen M, Grimshaw J, Oxman AD

# 2nd generation trials (2005-2015)

- Increasing sophistication of intervention design (varied use of theory)
- Technical issues (clustering) better addressed, modest sample sizes
- Design – majority are two arm trials (intervention vs control)
- Some effort to explore causal mechanisms
- Economic evaluation – largely ignored
- Reporting – insufficient details of context, intervention, and methods

# 2<sup>nd</sup> generation trial – Quality In Acute Stroke Care (QASC) (2011)

Articles ■

**Findings** 19 ASUs were randomly assigned to intervention (n=10) or control (n=9). Of 6564 assessed for eligibility, 1696 patients' data were obtained (687 pre-intervention; 1009 post-intervention). Results showed that, irrespective of stroke severity, intervention ASU patients were significantly less likely to be dead or dependent (mRS  $\geq 2$ ) at 90 days than control ASU patients (236 [42%] of 558 patients in the intervention group vs 259 [58%] of 449 in the control group,  $p=0.002$ ; number needed to treat 6.4; adjusted absolute difference 15.7% [95% CI 5.8–25.4]). They also had a better SF-36 mean physical component summary score (45.6 [SD 10.2] in the intervention group vs 42.5 [10.5] in the control group,  $p=0.002$ ; adjusted absolute difference 3.4 [95% CI 1.2–5.5]) but no improvement was recorded in mortality (21 [4%] of 558 in intervention group and 24 [5%] of 451 in the control group,  $p=0.36$ ), SF-36 mean mental component summary score (49.5 [10.9] in the intervention group vs 49.4 [10.6] in the control group,  $p=0.69$ ) or functional dependency (Barthel Index  $\geq 60$ : 487 [92%] of 532 patients vs 380 [90%] of 423 patients;  $p=0.44$ ).

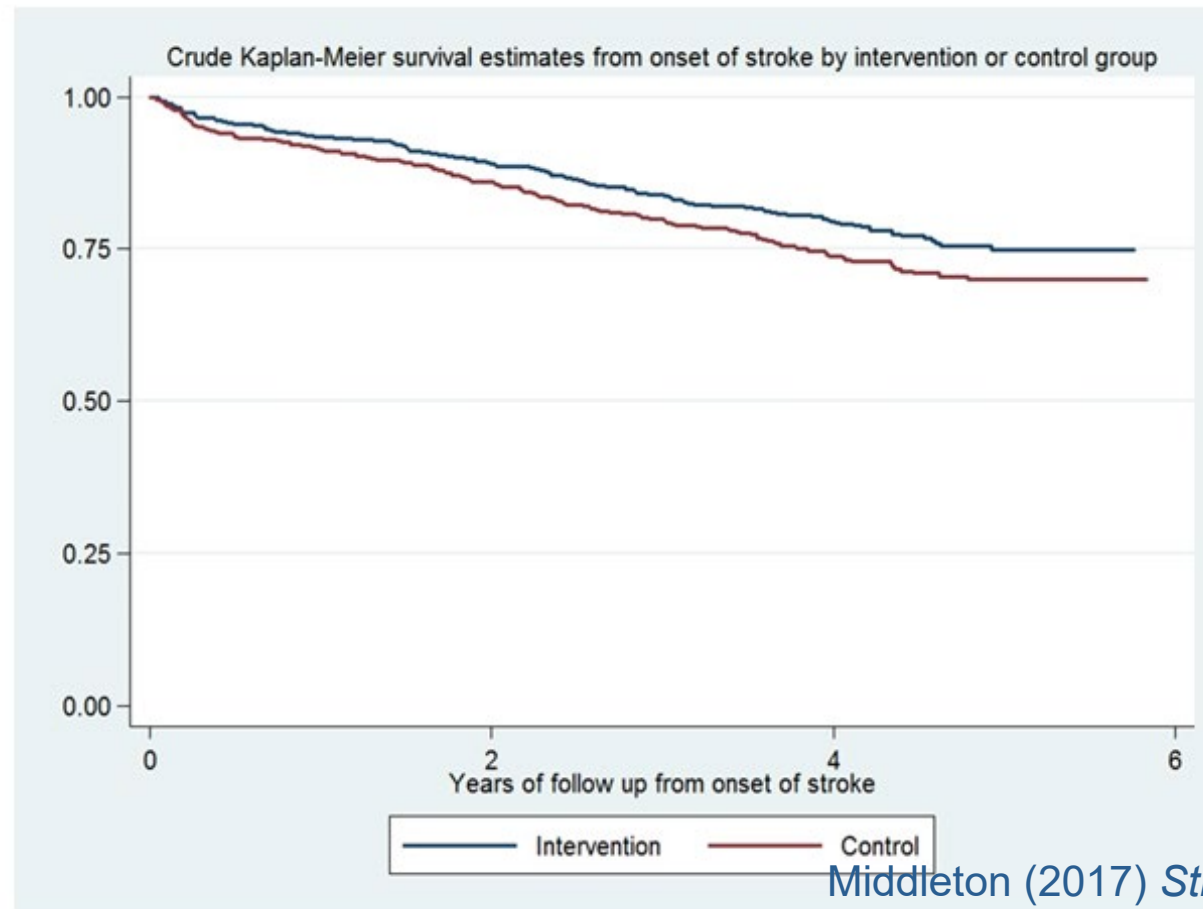
**Interpretation** Implementation of multidisciplinary supported evidence-based protocols initiated by nurses for the management of fever, hyperglycaemia, and swallowing dysfunction delivers better patient outcomes after discharge from stroke units. Our findings show the possibility to augment stroke unit care.

ASUs received treatment protocols to manage fever, hyperglycaemia, and swallowing dysfunction with multidisciplinary team building workshops to address implementation barriers. Control ASUs received only an abridged version of existing guidelines. We recruited pre-intervention and post-intervention patient cohorts to compare 90-day death or dependency (modified Rankin scale [mRS]  $\geq 2$ ), functional dependency (Barthel index), and SF-36 physical and mental component summary scores. Research assistants, the statistician, and patients were masked to trial groups. All

Sydney and School of Nursing  
(NSW & ACT), Australian  
Catholic University, NSW,  
Australia (Prof S Middleton PhD,  
S Dale BA(Hons), P Drury MN);  
Centre for Clinical Outcomes  
Research (Ch COOR), Australia



# 2<sup>nd</sup> generation trial – QASC (2011)



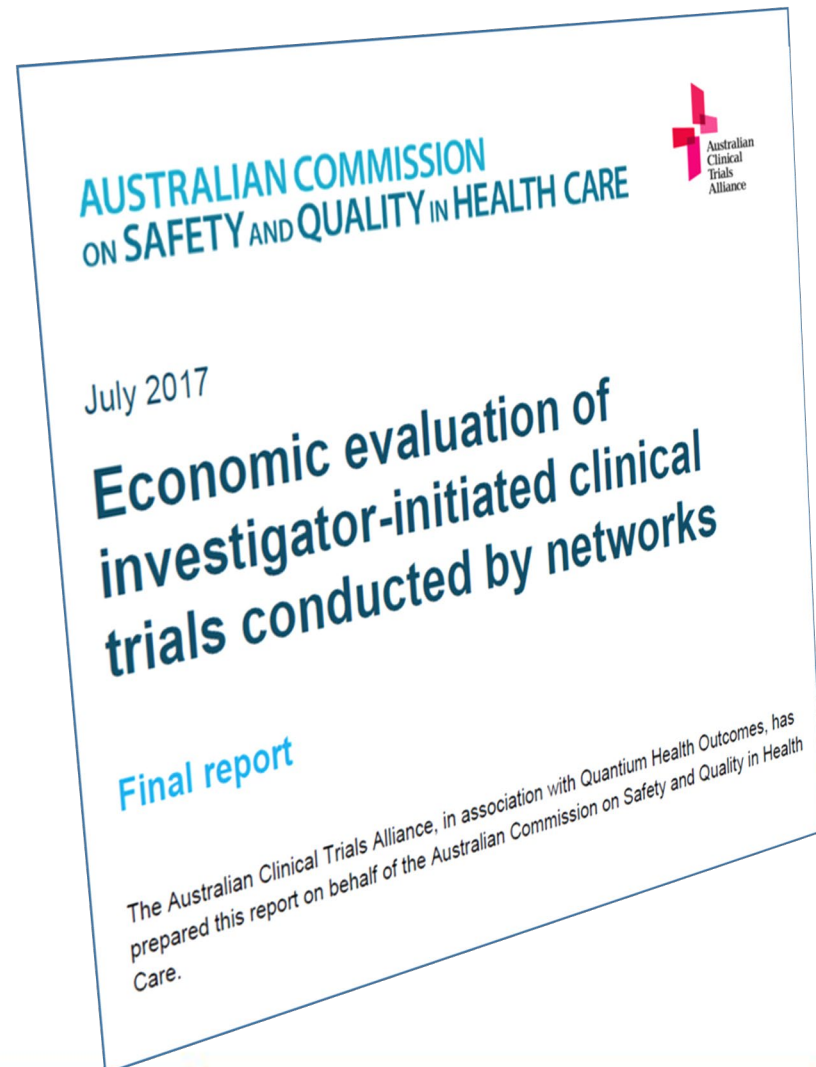
Middleton (2017) *Stroke*

## 2<sup>nd</sup> generation trial – QASC (2011)

- Improvements in proportions of patients that received protocol based care:
  - Fever (intervention: 31%; control: 15%  $P<0.001$ ),
  - Hyperglycaemia (intervention: 66%; control: 45%  $P<0.001$ )
  - Swallowing (intervention: 48%; control: 26%  $P<0.001$ )

Drury et al (2013) *International Journal of Stroke*

# 2<sup>nd</sup> generation trial – QASC (2011)



## Independent QASC Economic Evaluation

If FeSS protocols were implemented in 65% of the eligible Australian patient populations for one year **the total economic benefit (saving) would be \$281 M**

## 2nd generation trials

- Address some (but not all) of issues with 1st generation of trials
- Continue to ask does the intervention work (within limited setting of two arm trial) and not how, why and when does the intervention work and how can we optimize it?



# 3<sup>rd</sup> generation trials (2015 - )

Ivers et al. *Implementation Science* 2014, **9**:14  
<http://www.implementation-science.com/content/9/1/14>



## DEBATE

## Open Access

### No more 'business as usual' with audit and feedback interventions: towards an agenda for a reinvigorated intervention

Noah M Ivers<sup>1\*</sup>, Anne Sales<sup>2</sup>, Heather Colquhoun<sup>3</sup>, Susan Michie<sup>4</sup>, Robbie Foy<sup>5</sup>, Jill J Francis<sup>6</sup> and Jeremy M Grimshaw<sup>7</sup>

#### Abstract

**Background:** Audit and feedback interventions in healthcare have been found to be effective, but there has been little progress with respect to understanding their mechanisms of action or identifying their key 'active ingredients.'

**Discussion:** Given the increasing use of audit and feedback to improve quality of care, it is imperative to focus further research on understanding how and when it works best. In this paper, we argue that continuing the 'business as usual' approach to evaluating two-arm trials of audit and feedback interventions against usual care for common problems and settings is unlikely to contribute new generalizable findings. Future audit and feedback trials should incorporate evidence- and theory-based best practices, and address known gaps in the literature.

**Summary:** We offer an agenda for high-priority research topics for implementation researchers that focuses on reviewing best practices for designing audit and feedback interventions to optimize effectiveness.

**Keywords:** Audit and feedback, Synthesis, Best practice, Implementation, Optimization

#### Background

Audit and feedback (A&F) involves providing a recipient with a summary of their performance over a specified period of time and is a common strategy to promote the implementation of evidence-based practices. A&F is used widely in healthcare by a range of stakeholders, including research funders and health system payers, delivery organizations, professional groups and researchers, to monitor and change health professionals' behaviour, both to increase accountability and to improve quality of care. A&F is an improvement over self-assessment [1] or self-monitoring [2] as it can provide objective data regarding discrepancies between current practice and target performance, as well as comparisons of performance to other health professionals. The recognition of sub-optimal performance can act as a cue for action, encouraging those who are both motivated and capable to take action to reduce the discrepancy.

The effectiveness of A&F has been evaluated in the third update of a Cochrane review, which included 140 randomized trials of A&F conducted across many clinical conditions and settings around the world. The review found that A&F leads to a median 4.3% absolute improvement (interquartile range 0.5% to 16%) in provider compliance with desired practice [3]. One-quarter of A&F interventions had a relatively large, positive effect on quality of care, while another quarter had a negative or null effect. The challenge of identifying factors that differentiate more and less successful A&F interventions is exacerbated by poor reporting of both intervention components and contextual factors in the literature [4]. Furthermore, most A&F interventions tested in RCTs are designed without explicitly building on previous research or extant theory [5,6]. As a result, there has been little progress with respect to identifying the key ingredients for a successful A&F intervention or understanding the mechanisms of action of effective A&F interventions.

## Head-to-head arm trials evaluating:

- alternative ways of designing and/or delivering audit and feedback
- audit and feedback vs audit and feedback plus co-interventions
- audit and feedback versus alternative interventions

# 3<sup>rd</sup> generation trials (2015 -)

## RESEARCH AND REPORTING METHODOLOGY



OPEN ACCESS

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

JM Grimshaw,<sup>1,2</sup> Noah Ivers,<sup>3,4</sup> Stefanie Linklater,<sup>1</sup> Robbie Foy,<sup>5</sup> Jill J Francis,<sup>6</sup> Wouter T Gude,<sup>7</sup> Sylvia J Hysong,<sup>8,9</sup> 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

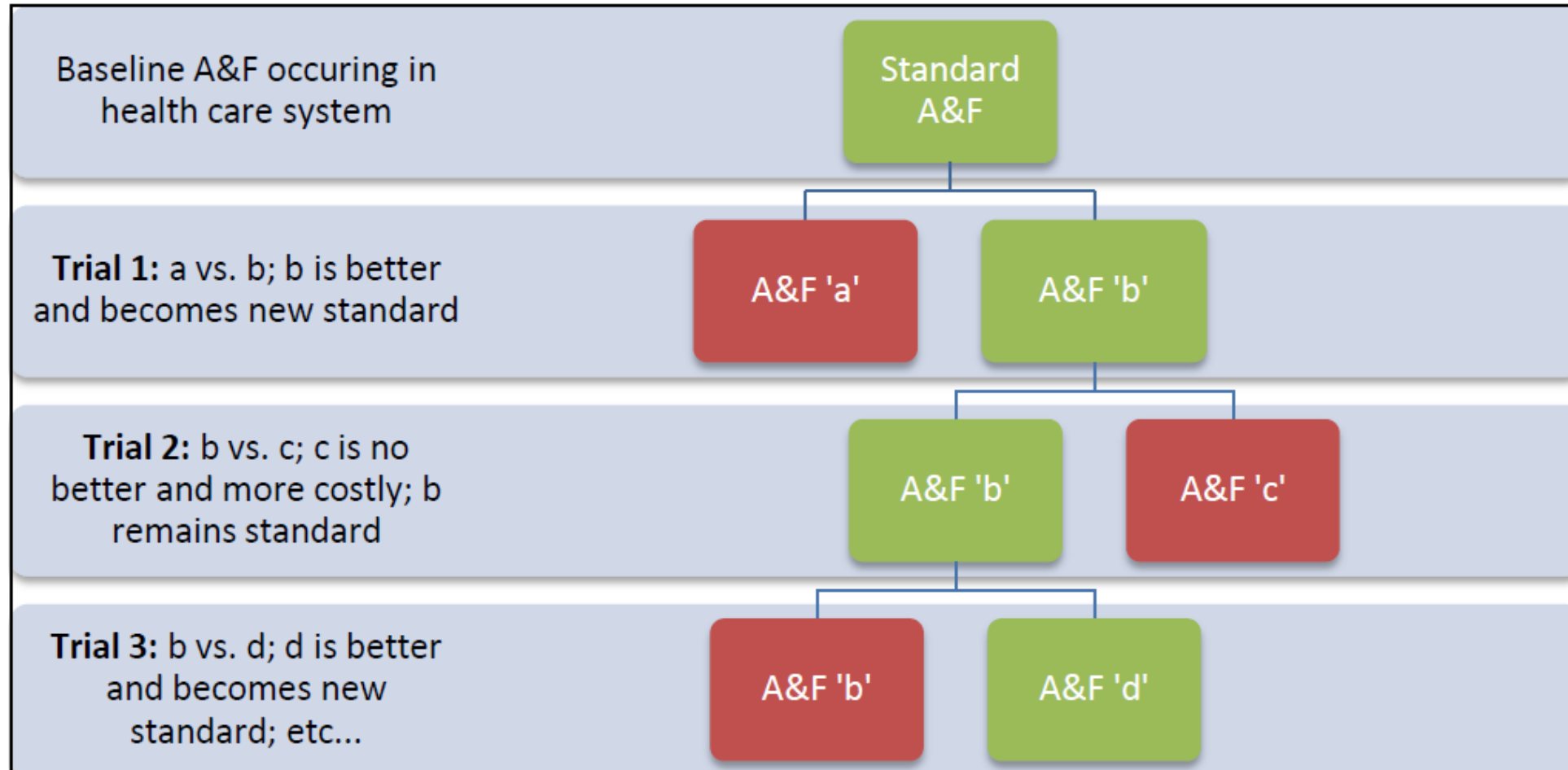
Received 18 May 2018  
Revised 8 January 2019  
Accepted 12 February 2019

### 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

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’.<sup>3</sup>

# 3<sup>rd</sup> generation trials (2015 -)



# 3<sup>rd</sup> generation trial – RAPID (2016)



## RESEARCH ARTICLE

### An Audit and Feedback Intervention for Reducing Antibiotic Prescribing in General Dental Practice: The RAPID Cluster Randomised Controlled Trial

**Paula Elouafkaoui<sup>1,2</sup>, Linda Young<sup>1\*</sup>, Rumana Newlands<sup>3</sup>, Eilidh M. Duncan<sup>3</sup>, Andrew Elders<sup>4</sup>, Jan E. Clarkson<sup>1,2</sup>, Craig R. Ramsay<sup>3</sup>, Translation Research in a Dental Setting (TRiADS) Research Methodology Group<sup>†</sup>**

**1** NHS Education for Scotland (NES), Dundee Dental Education Centre, Frankland Building, Dundee, United Kingdom, **2** Dental Health Services Research Unit (DHSRU), University of Dundee, Park Place, Dundee, United Kingdom, **3** Health Services Research Unit (HSRU), University of Aberdeen, Health Sciences Building, Foresterhill, Aberdeen, United Kingdom, **4** NMAHP Research Unit, Glasgow Caledonian University, Cowcaddens Road, Glasgow, United Kingdom

<sup>†</sup> Membership of the Translation Research in a Dental Setting (TRiADS) Research Methodology Group is provided in the Acknowledgments.

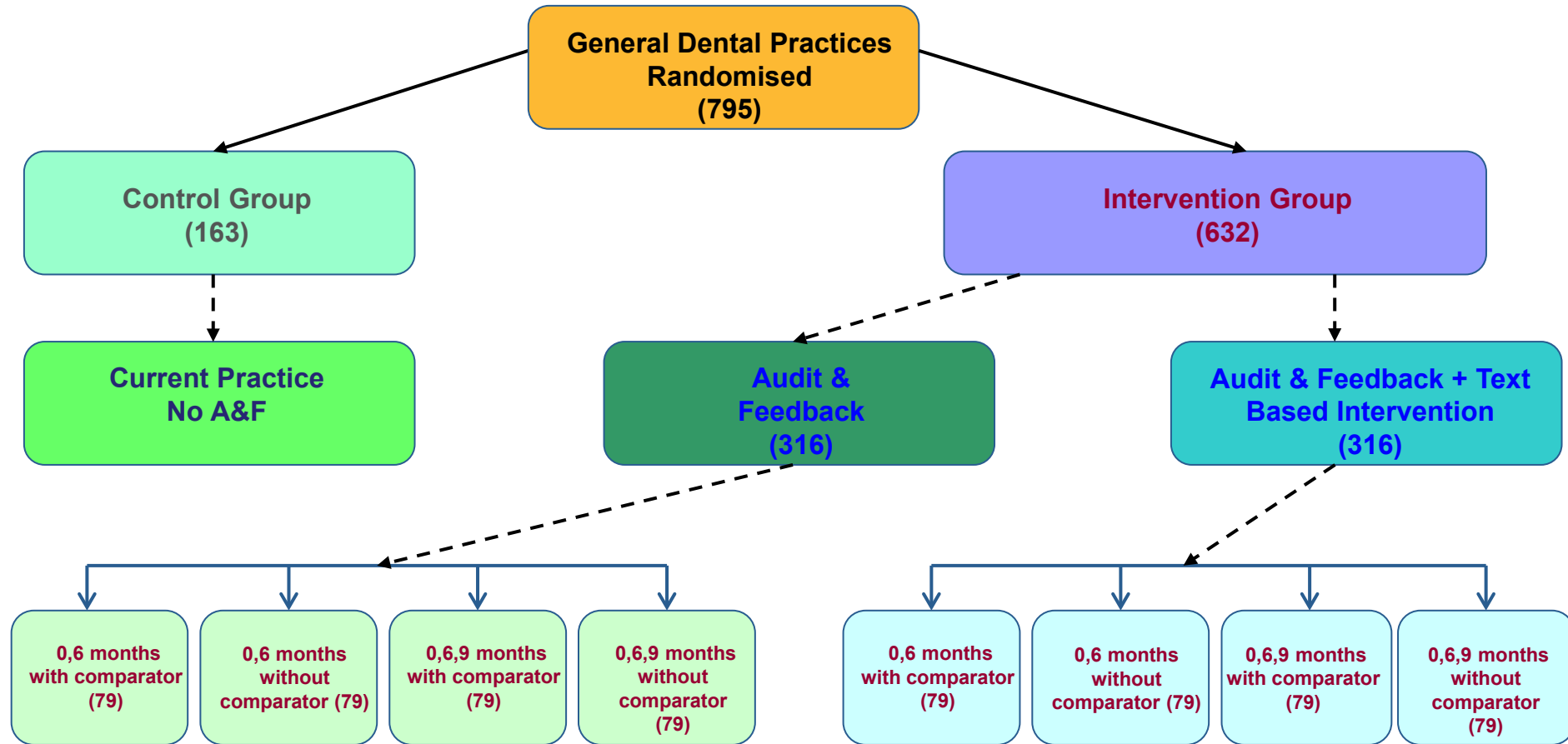
\* [linda.young@nes.scot.nhs.uk](mailto:linda.young@nes.scot.nhs.uk)



OPEN ACCESS

Citation: Elouafkaoui P, Young L, Newlands R, Duncan E, Elders A, Clarkson JE, Ramsay CR, et al. (2016) An Audit and Feedback Intervention for Reducing Antibiotic Prescribing in General Dental Practice: The RAPID Cluster Randomised Controlled Trial. *PLOS Medicine* 12(12): e1002000. doi:10.1371/journal.pmed.1002000

# 3<sup>rd</sup> generation trial – RAPID (2016)



# Summary

- Randomised trials of implementation programs are essential to advancing implementation knowledge
- 10,000s of implementation randomized trials but issues in the design and conduct of trials has limited development of implementation knowledge
- Future trials need to more directly address current knowledge gaps with robust methods at scale.

# Questions?

Dr. Jeremy Grimshaw  
jgrimshaw@ohri.ca  
@GrimshawJeremy

