



What works, why doesn't it work all of the time, and how could we do better?

Liz Glidewell

- What is implementation?
- Why do we need implementation research, isn't it all commonsense?
- Implementation components & strategies how can we tell what works (effectiveness) and where can we find out more?
- The effectiveness of financial incentives and the importance of rigorous and generalisable implementation
- Can frameworks help us to think what is most likely to work for whom, in what circumstances?

- **The BIG IDEA: ‘patient care will improve if clinicians get timely and credible information linking local clinical practices and patient outcomes.’**
- **Growth since +++ in range of outcome measures**
- **( including clinical factors, death, disease, functional status, well-being, satisfaction and cost. )**
- **Continuous debate though about whether the approach will really lead to better quality of care (Epstein AM, 1990)**



- The study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice and, hence, to improve the quality and safety of health care
- Includes the study of influences on healthcare professional and organisational behaviour



# Plan for the session

- What is implementation?
- Why do we need implementation research, isn't it all commonsense?
- Implementation components & strategies how can we tell what works (effectiveness) and where can we find out more?
- The effectiveness of financial incentives and the importance of rigorous and generalisable implementation
- Can frameworks help us to think what is most likely to work for whom, in what circumstances?

# What's the evidence that NICE guidance has been implemented?



UNIVERSITY OF LEEDS

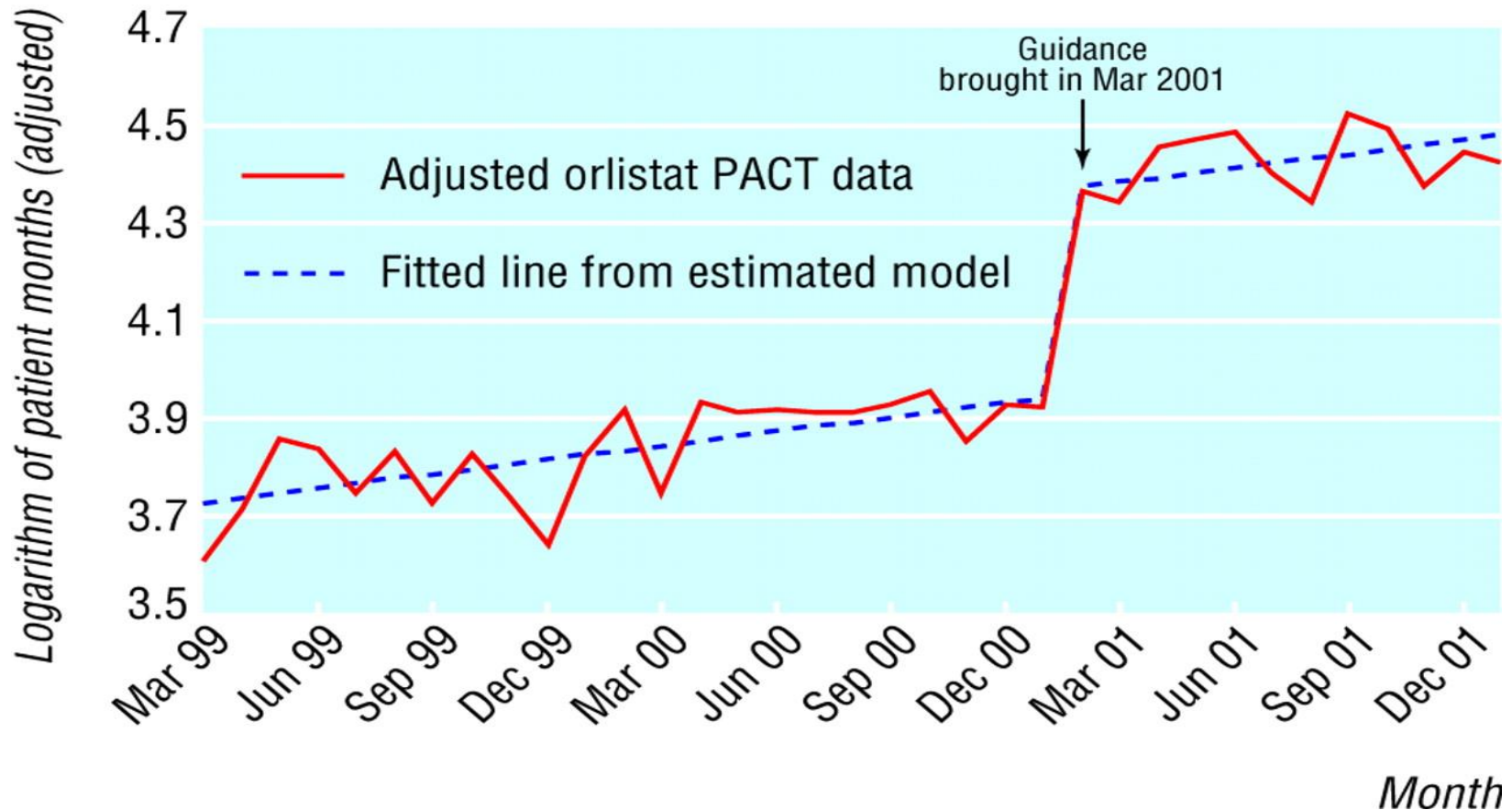
National evaluation using time series analysis, audit of patients' notes, and interviews

Sheldon TA, et al. BMJ 2004;329:999

# Guidelines change practice: Use of orlistat for obesity



UNIVERSITY OF LEEDS



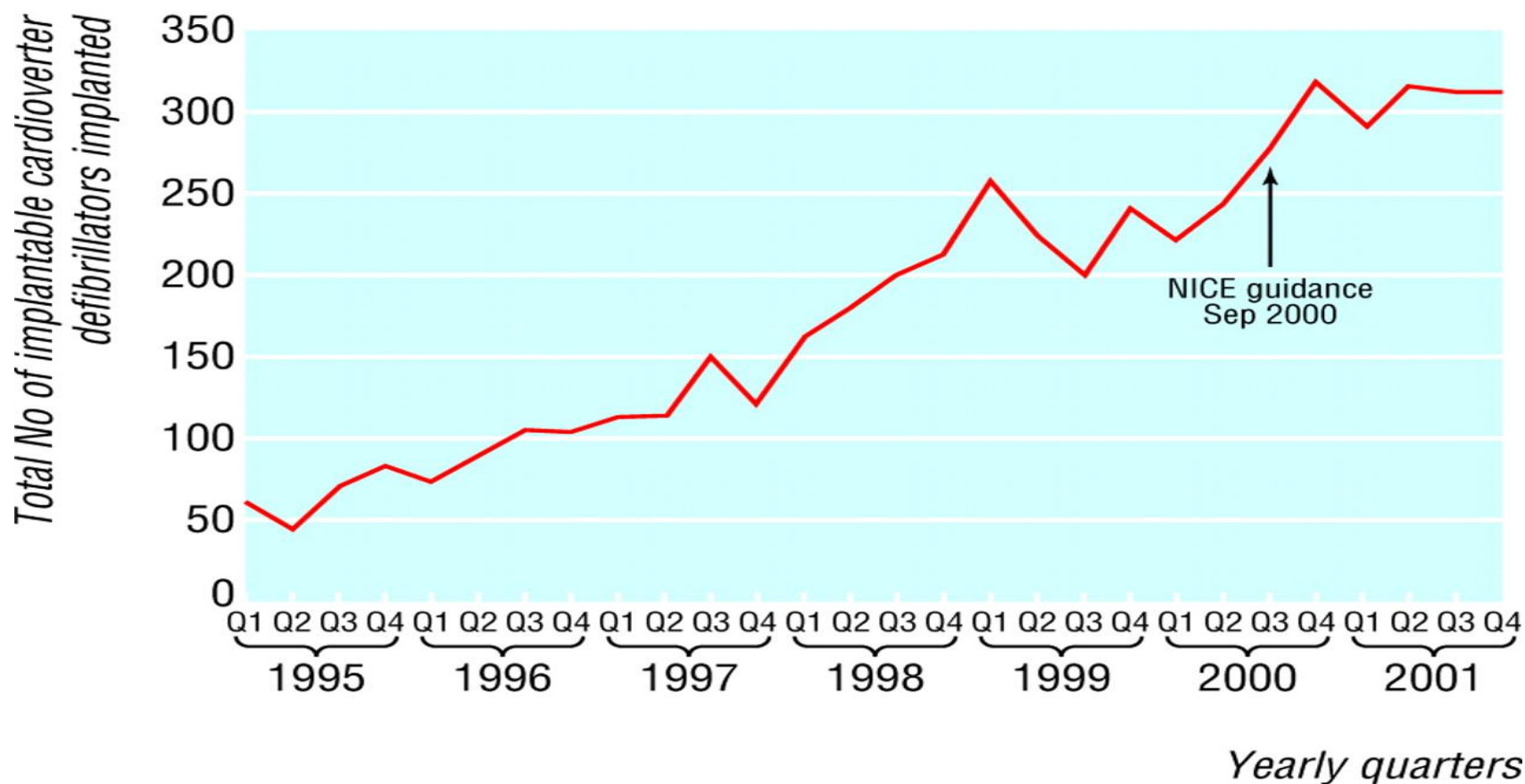
Sheldon, T. A et al. BMJ 2004;329:999

BMJ

# Or do they? Total number of implantable cardioverter defibrillators implanted by quarter, 1995-2001



UNIVERSITY OF LEEDS



Sheldon, T. A et al. BMJ 2004;329:999

BMJ

# Why do we need implementation research?



UNIVERSITY OF LEEDS

“In almost all studies the process of care did not reach the standards set out in national guidelines or set by the researchers themselves.”

Seddon ME, Marshall MN, Campbell SM, Roland MO.  
*Qual Health Care* 2001, 10:152-158.

The US spends 16% of gross domestic product on health care yet...

*Patients receive only 55% of recommended health services and disadvantaged patients and minorities fare worse*

McGlynn EA, Asch SM, Adams J, et al.  
*N Engl J Med* 2003;348:2635-2645

# What's the evidence that NICE guidance has been implemented?



UNIVERSITY OF LEEDS

- Some clinical practice has changed in line with NICE guidance, in particular around prescribing
- Other technologies have been adopted in line with NICE guidance but continued pre-existing practice patterns
- Evidence that NICE guidance has been less influential in surgical procedures and use of medical devices
- Routine data not sufficient to assess compliance with guidance
- Impact of NICE guidance likely to be greater if
  - more effort is devoted to clarity of guidance and its relevance to practice
  - adequate funding provision
  - getting professional support
  - encouraging healthcare organisations to set up formal mechanisms for handling guidance



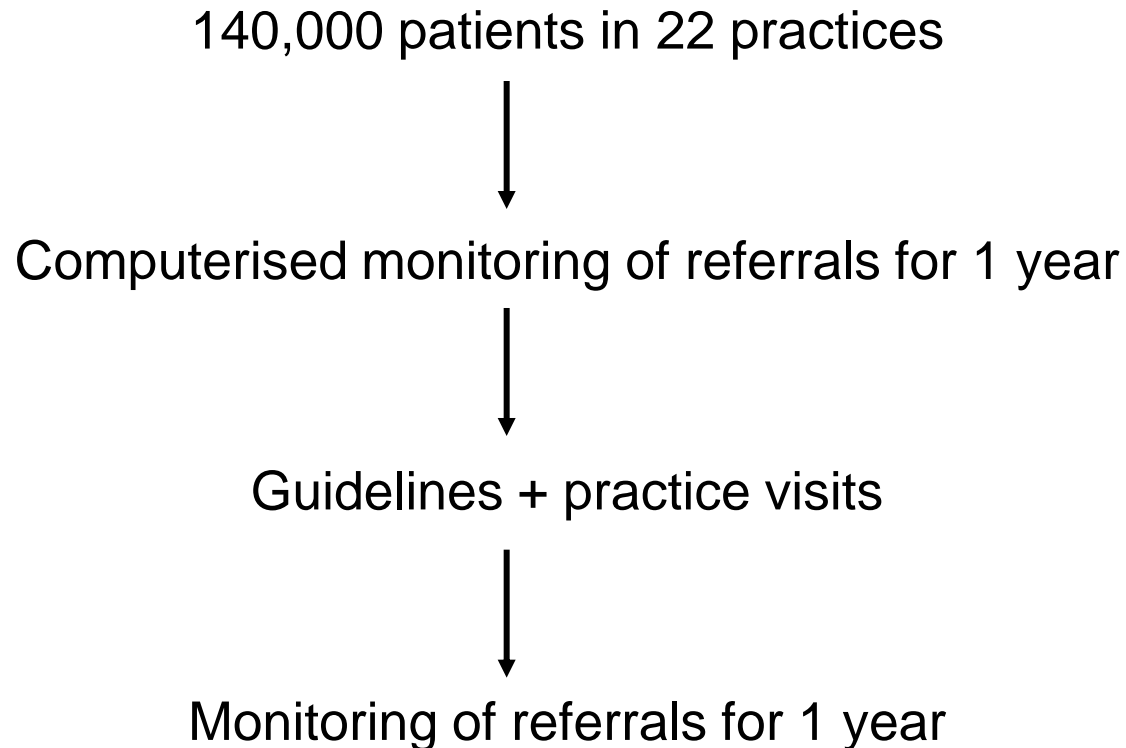
- Resources available for health care are limited
- Variations in clinical practice are unexplained by characteristics of patients, their illnesses or setting of care
- There is evidence of unacceptable standards of practice
- There is a lag between the emergence of research evidence and its incorporation into routine practice

# Why do we need implementation research?

## Influence of radiology guidelines on x-ray referrals



UNIVERSITY OF LEEDS



Royal College of Radiologists Working Party BMJ 1993;306:110-11



## Referral rates

12 months before intervention	88/1000 registered patients
	↓
12 months after intervention	77/1000 registered patients

Royal College of Radiologists Working Party BMJ 1993;306:110-11



“Application of these results throughout the whole of the UK would reduce referrals by at least 0.6 million and produce potential savings of £5.4 million.”

Royal College of Radiologists Working Party BMJ 1993;306:110-11

Debate

Open Access

## Designing theoretically-informed implementation interventions

The Improved Clinical Effectiveness through Behavioural Research Group (ICEBeRG) \*2006



Imagine an initial trial of a drug to reduce the likelihood of acute stroke in high-risk patients, where the drug is described as "**the red pill**" rather than in terms of its pharmacological properties. Over two to three years the "**red pill**" produces positive outcomes across a range of randomised controlled trials of patients at high risk of stroke. It is trialled in patients with moderate risk and low risk, again producing positive outcomes.

Debate

Open Access

## Designing theoretically-informed implementation interventions

The Improved Clinical Effectiveness through Behavioural Research Group (ICEBeRG) \* 2006



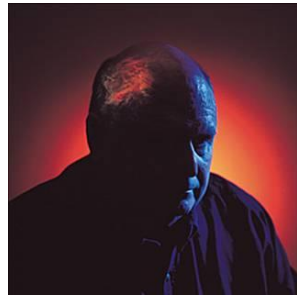
Clinicians are impressed by the "red pill's" (unknown) properties and so begin to investigate its role in the treatment of a range of other conditions, though these are chosen on an ad hoc basis as there is no underlying rationale for its use. Equally impressed by the effects of red pills, a number of pharmaceutical companies launch other versions of red pills – the magenta pill, the crimson pill, and the vermillion pill.

Debate

Open Access

## Designing theoretically-informed implementation interventions

The Improved Clinical Effectiveness through Behavioural Research Group (ICEBeRG) \*2006



After ten years of trials the Cochrane Collaboration Red Pill Review Group begins to conduct systematic reviews of the effectiveness of "red pills" in the treatment of patients with stroke, asthma, epilepsy, and migraine to establish the generalisable messages about the effectiveness of "red pills."



- Ideally, should be based upon an understanding of barriers and enablers
- Most research has focused on characteristics of clinicians or health care organisations, such as local attitudes or preparedness to change
- Less attention has been paid to the characteristics of guideline recommendations themselves

- Relative simplicity/complexity of tasks, e.g.
  - Step up asthma therapy
  - Introducing and monitoring of beta-blockers for heart failure
  
- Starting new or stopping established habitual practices, e.g.
  - Checking BMI
  - Prescribing antibiotics for upper respiratory tract infections



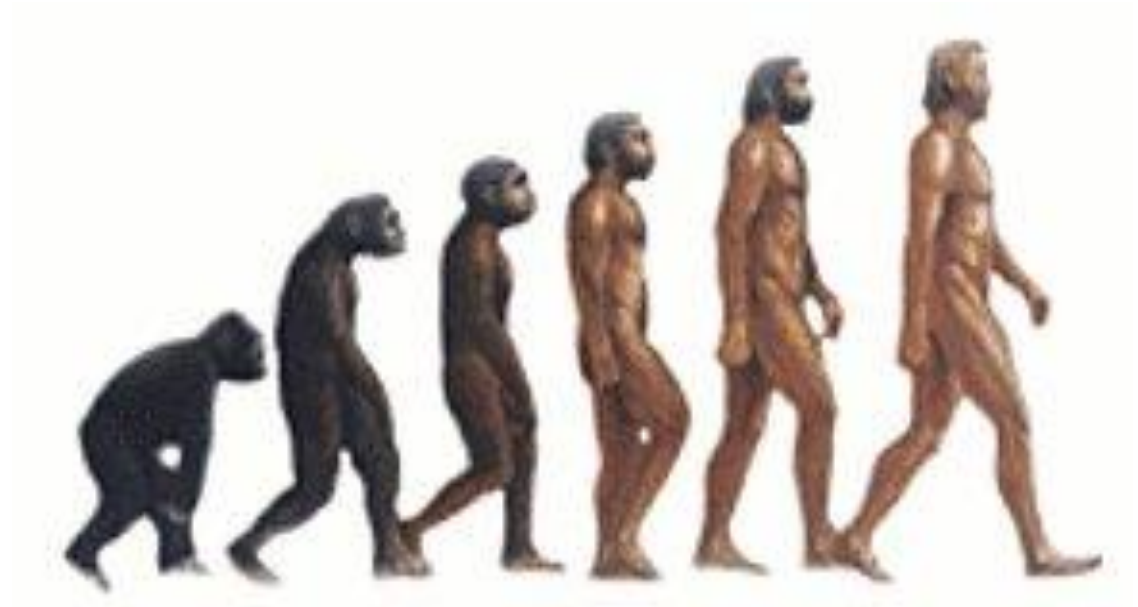
# Characterising recommendations

## Order of work



UNIVERSITY OF LEEDS

- Earlier work
  - Rogers
  - Grilli
  - Grol
  - Foy (Grimshaw)
  - Shekelle
- Recent work
  - Michie



# Characterising recommendations

## Review of 23 published studies



UNIVERSITY OF LEEDS

- Compliance rates with 143 different recommendations
- Wide range of contexts and specialities
- The researchers (not targeted clinicians) made judgements about the characteristics of each recommendation
- Recommendation compliance higher with:
  - **Low complexity** (mean 56% vs 42% for high complexity)
  - **High trialability** (56% vs 39% for low trialability)
  - No differences for observability and insufficient data available to assess relative advantage and complexity

Grilli & Lomas. Medical Care 1994;32:202-13



- 47 recommendations from 10 national clinical guidelines
- Consensus process to judge degree to which recommendations possessed 16 characteristics
- Post-dissemination compliance assessed during a clinical audit using self-recorded data from 61 GPs
- Multivariate analysis
- Compliance lower if recommendations were:
  - Vaguely worded
  - Incompatible with clinician norms and values
  - Disruptive to routine practice



- 16 hospital gynaecology units in an audit and feedback project
- 42 clinical practice recommendations from 6 guidelines
- Panel of 7 gynaecologists rated the extent to which each recommendation possessed each of 13 attributes
- Multivariate analysis
- Higher compliance if recommendations
  - Compatible with clinician values
  - Not requiring changes to fixed routines
- *Greater change* if recommendations
  - Incompatible with clinician values

# Are non-specific practice guidelines potentially harmful?



UNIVERSITY OF LEEDS

- Randomised trial of the effect of practice guidelines on self-reported physician test ordering behaviour in response to a series of 12 clinical vignettes
- Two different versions of a practice guideline for the use of electrodiagnostic tests (EDT) for low back problems
- The two guidelines were similar in content but varied in the specificity of their recommendations

Shekelle et al. *Health Serv Res* 2000;34:1429-1448

# Are non-specific practice guidelines potentially harmful?



UNIVERSITY OF LEEDS

	No guideline	Non-specific guideline	Specific guideline
Proportion of appropriate vignettes for which EDTs were ordered ( $p = 0.002$ )	77%	71%	79%
EDTs ordered for inappropriate vignettes ( $p = 0.08$ )	32%	32%	26%

# Improving guideline recommendations



UNIVERSITY OF LEEDS

- Psychological research shows that the more precisely behaviours are specified, the more they are likely to be carried out
- Rewriting guidelines to increase behavioural specificity may be the simplest, most effective method of increasing implementation
- Specifying what, who, when, where, and how will assist implementation
- Writing guidelines with high behavioural specificity in conjunction with the use of "plain English" may be a simple and effective method of increasing their implementation

Michie & Johnston. Changing clinical behaviour by making guidelines specific. *BMJ* 2004;328:343-345

Michie & Lester. Words matter: increasing the implementation of clinical guidelines. *Qual. Saf. Health Care* 2005; 4:367-370.



# Plan for the session

- What is implementation?
- Why do we need implementation research, **isn't it all commonsense?**
- Implementation components & strategies how can we tell what works (effectiveness) and where can we find out more?
- The effectiveness of financial incentives and the importance of rigorous and generalisable implementation
- Can frameworks help us to think what is most likely to work for whom, in what circumstances?

“No magic bullets”



UNIVERSITY OF LEEDS

# No magic bullets?

Oxman AD, Thomson MA, Davis DA, et al. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *Can Med Assoc J* 1995;153:1423-1431

# What do we know



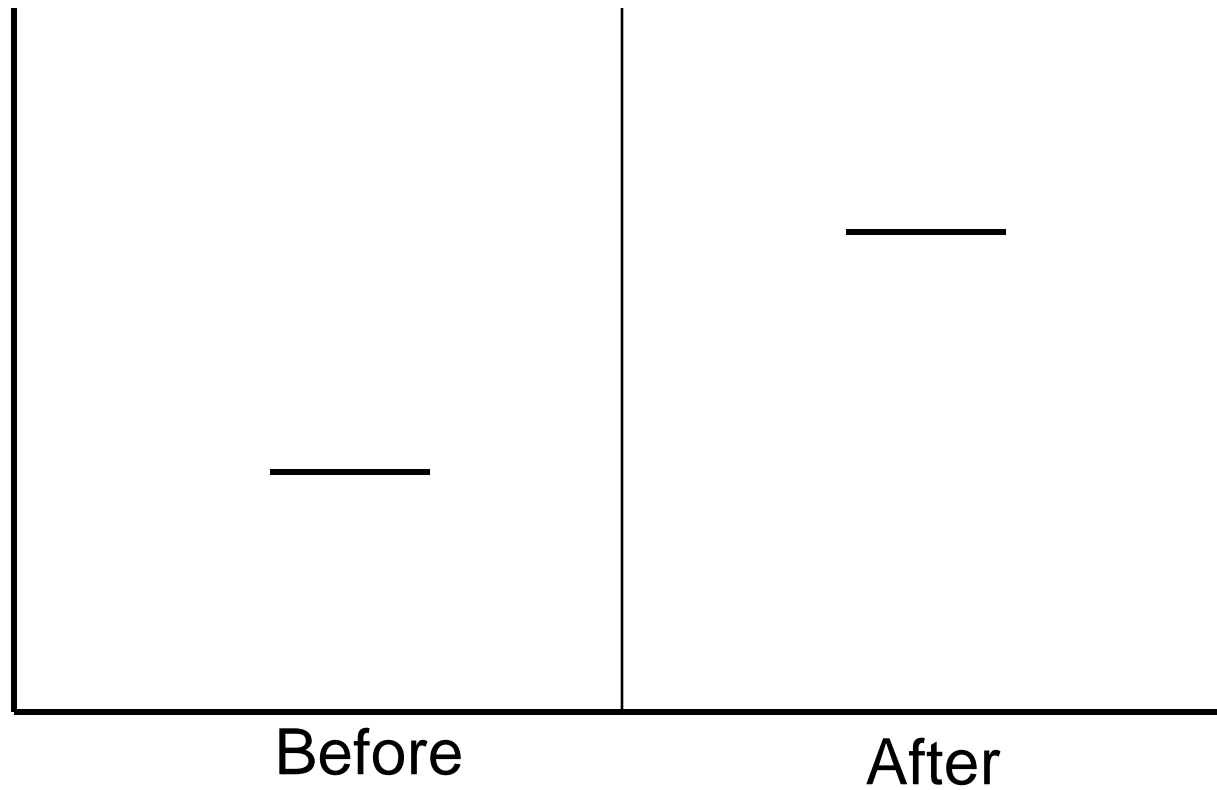
UNIVERSITY OF LEEDS



# Uncontrolled before and after



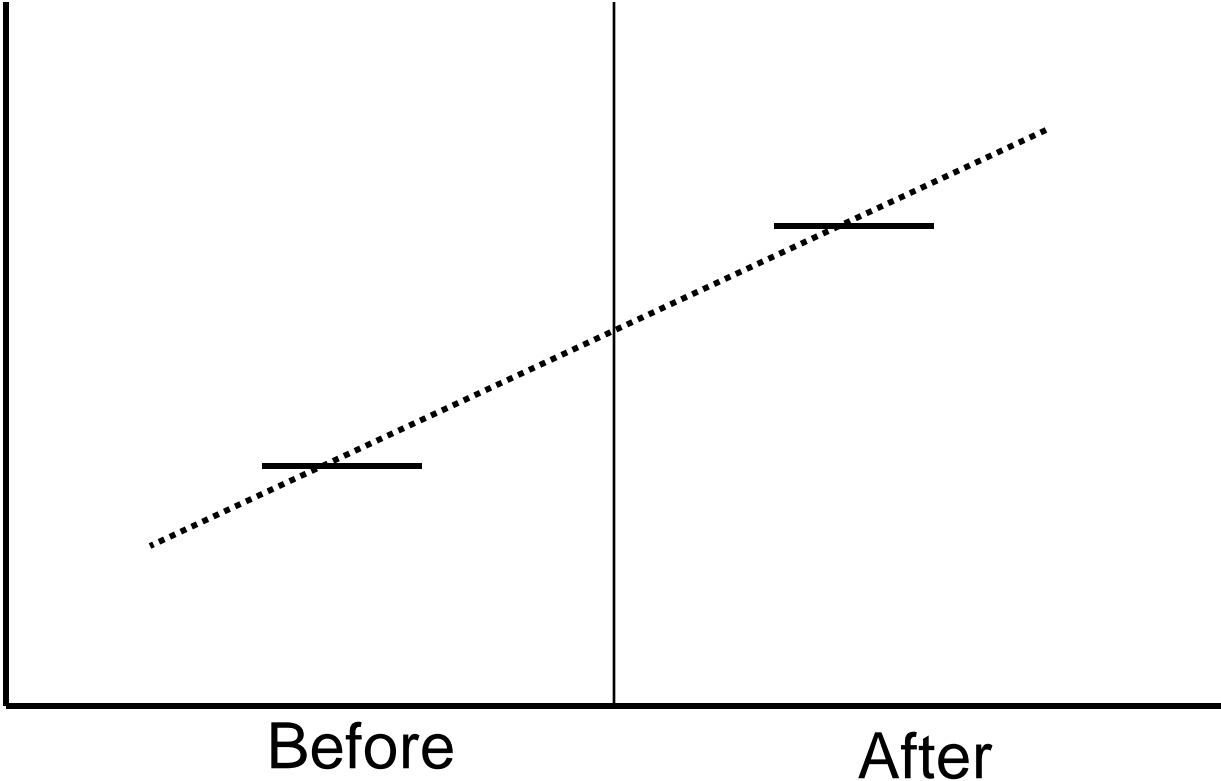
UNIVERSITY OF LEEDS



# Uncontrolled before and after



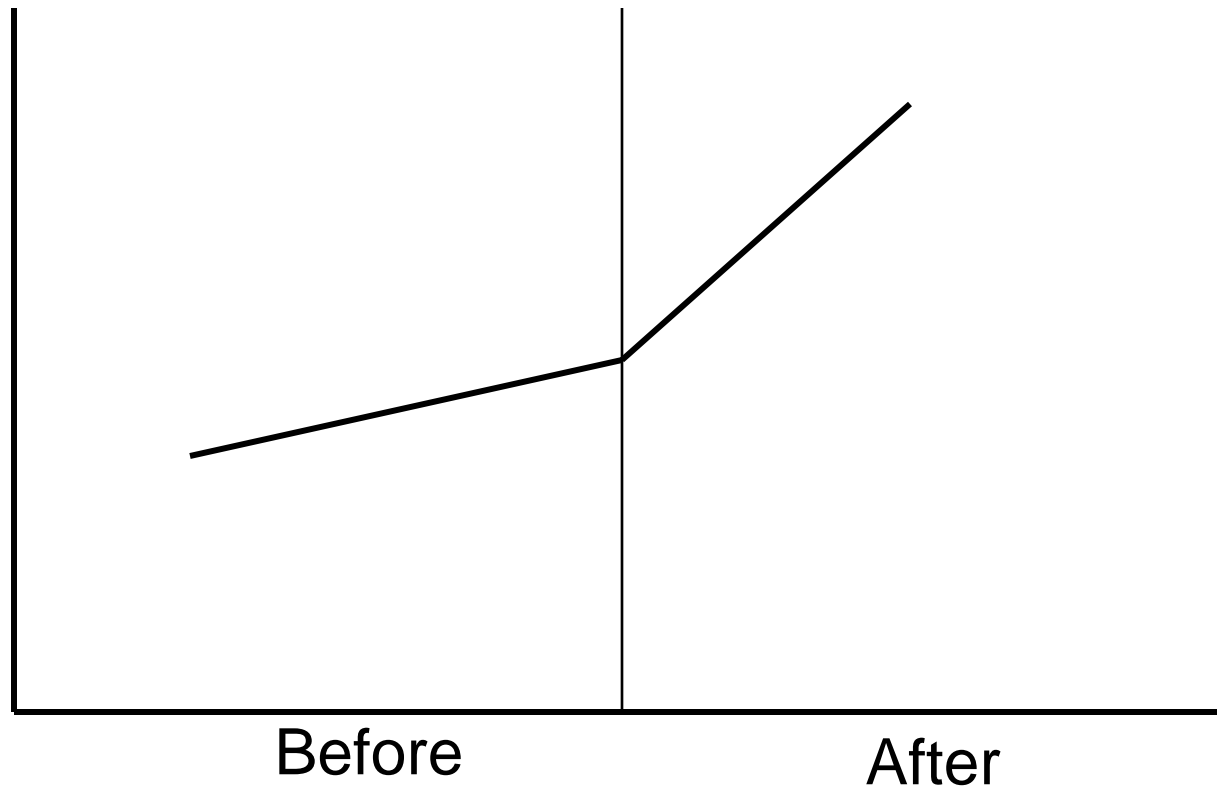
UNIVERSITY OF LEEDS



# Interrupted time series



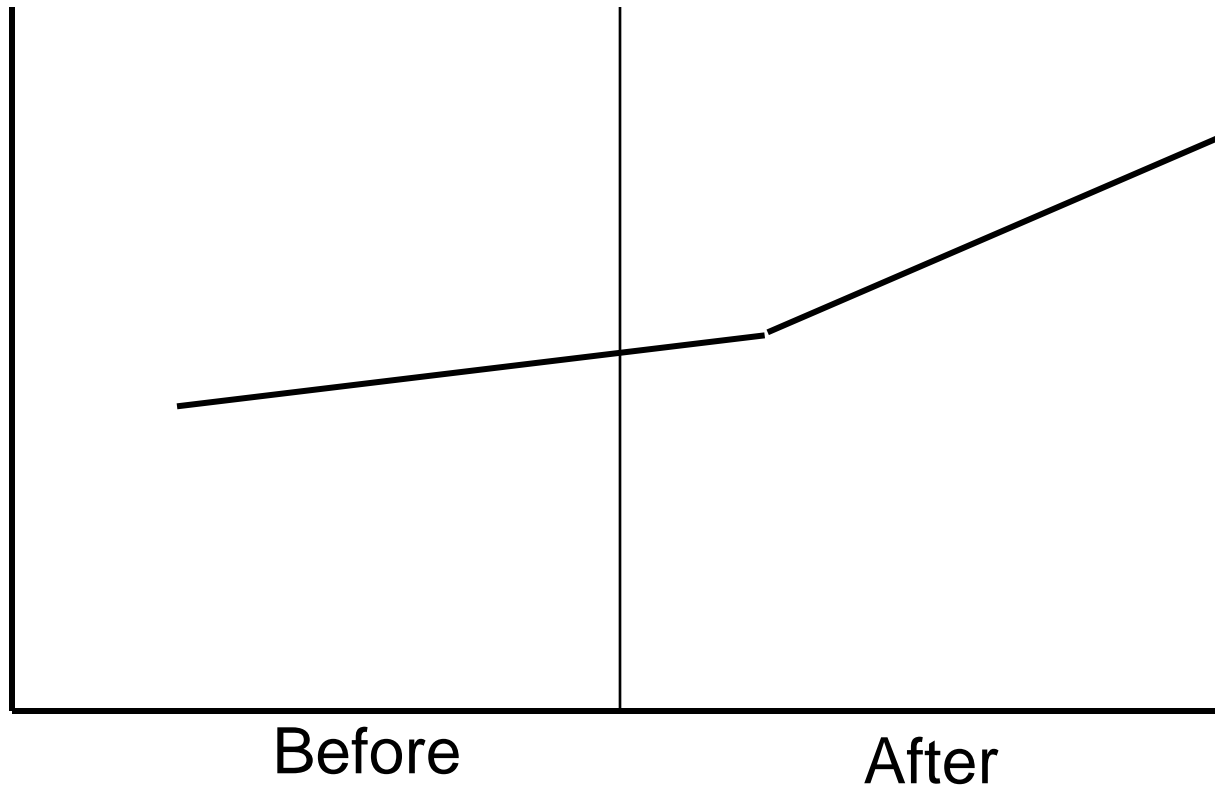
UNIVERSITY OF LEEDS



# Interrupted time series



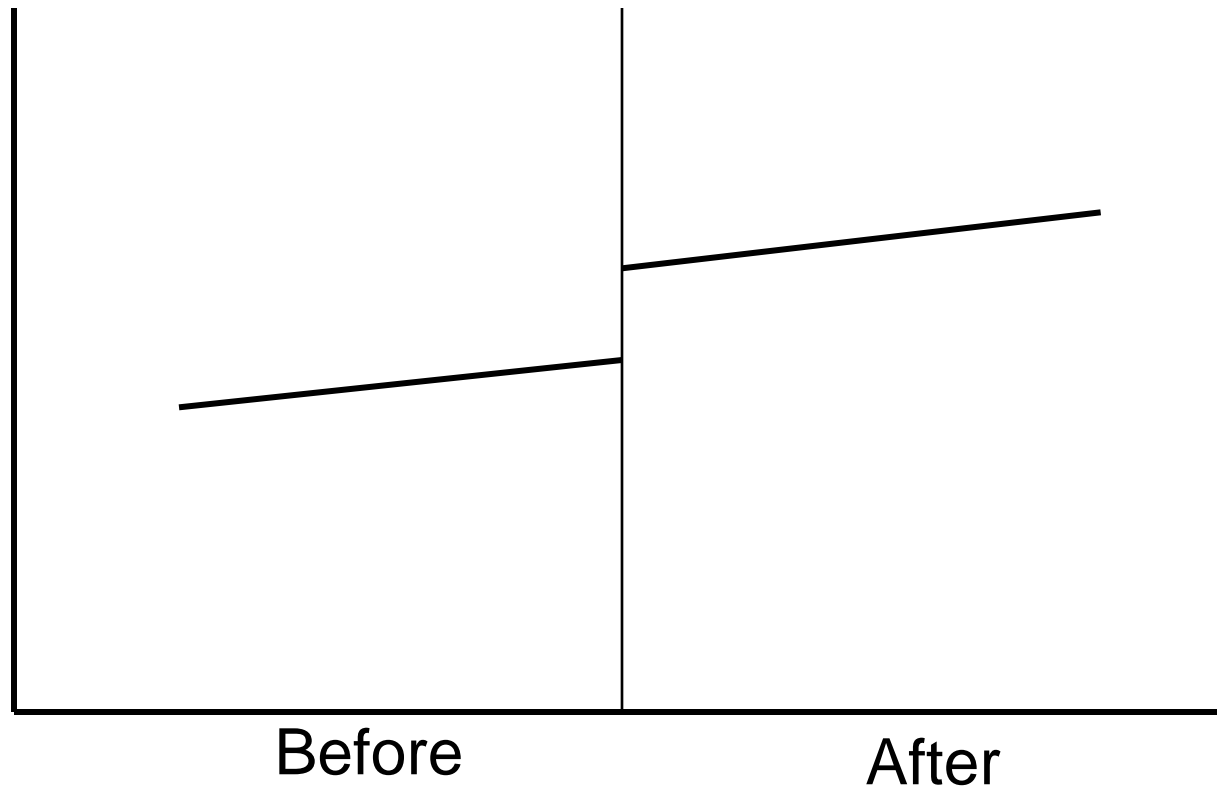
UNIVERSITY OF LEEDS



# Interrupted time series



UNIVERSITY OF LEEDS

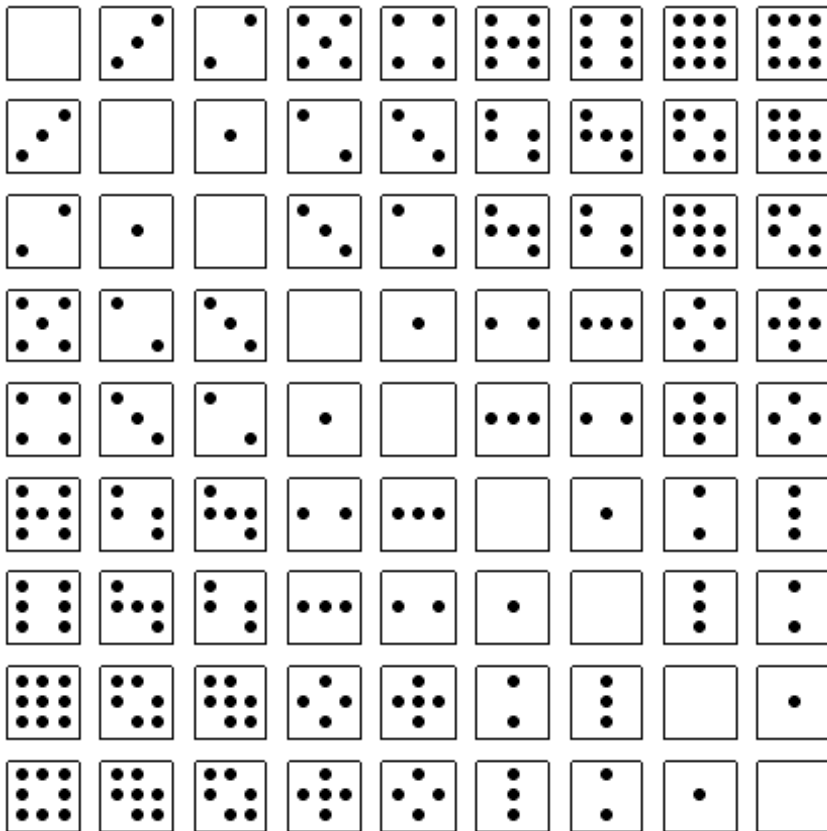




# How can we tell what works?

Review of interventions to improve prescribing

Poorly controlled studies more likely to show effect than well controlled studies



Soumerai SB, McLaughlin TJ, Avorn J. *Milbank Quarterly*, 1989;67:268-317

# How can we tell what works?

## Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials

Gordon C S Smith, Jill P Pell

### Abstract

**Objectives** To determine whether parachutes are effective in preventing major trauma related to gravitational challenge.

**Design** Systematic review of randomised controlled trials.

**Data sources:** Medline, Web of Science, Embase, and the Cochrane Library databases; appropriate internet sites and citation lists.

accepted intervention was a fabric device, secured by strings to a harness worn by the participant and released (either automatically or manually) during free fall with the purpose of limiting the rate of descent. We excluded studies that had no control group.

### Definition of outcomes

The major outcomes studied were death or major trauma, defined as an injury severity score greater than 15.<sup>o</sup>

### Meta-analysis

Our statistical approach was to assess outcomes in parachute and control groups by odds ratios and quantified the precision of estimates by 95% confidence intervals. We chose the Mantel-Haenszel test to assess heterogeneity, and sensitivity and subgroup analyses and fixed effects weighted regression techniques to explore causes of heterogeneity. We selected a funnel plot to assess publication bias visually and Egger's and Begg's tests to test it quantitatively. Stata software, version 7.0, was the tool for all statistical analyses.

### Results

Our search strategy did not find any randomised controlled trials of the parachute.

### Discussion

**Evidence based pride and observational prejudice**  
It is a truth universally acknowledged that a medical intervention justified by observational data must be in want of verification through a randomised controlled

Department of Obstetrics and Gynaecology, Cambridge University, Cambridge CB2 2QQ

Gordon C S Smith  
*professor*

Department of Public Health, Greater Glasgow NHS Board, Glasgow G3 8YU  
Jill P Pell  
*consultant*

Correspondence to:  
G C S Smith  
gcs2@cama.ac.uk

BMJ 2005;327:1459-61

**Conclusions** As with many interventions intended to prevent ill health, the effectiveness of parachutes has not been subjected to rigorous evaluation by using randomised controlled trials. Advocates of evidence based medicine have criticised the adoption of interventions evaluated by using only observational data. We think that everyone might benefit if the most radical protagonists of evidence based medicine organised and participated in a double blind, randomised, placebo controlled, crossover trial of the parachute.

led trials of parachutes.

### Methods

#### Literature search

We conducted the review in accordance with the QUOROM (quality of reporting of meta-analyses) guidelines.<sup>5</sup> We searched for randomised controlled trials of parachute use on Medline, Web of Science, Embase, the Cochrane Library, appropriate internet sites, and citation lists. Search words employed were "parachute" and "trial." We imposed no language restriction and included any studies that entailed jumping from a height greater than 100 metres. The

s of using  
uma,  
domised  
ntended to  
hutes has  
y using  
vidence  
i of  
vational  
if the most  
icine  
d,  
rial of the

itary sector,  
rthopaedic,  
ravitational  
ng from an  
e a success-  
al evidence.  
se is associ-  
oath failure  
ications? In  
all indicate  
e does not  
e therefore  
ed control-



Parachutes reduce the risk of injury after gravitational challenge, but their effectiveness has not been proved with randomised controlled trials



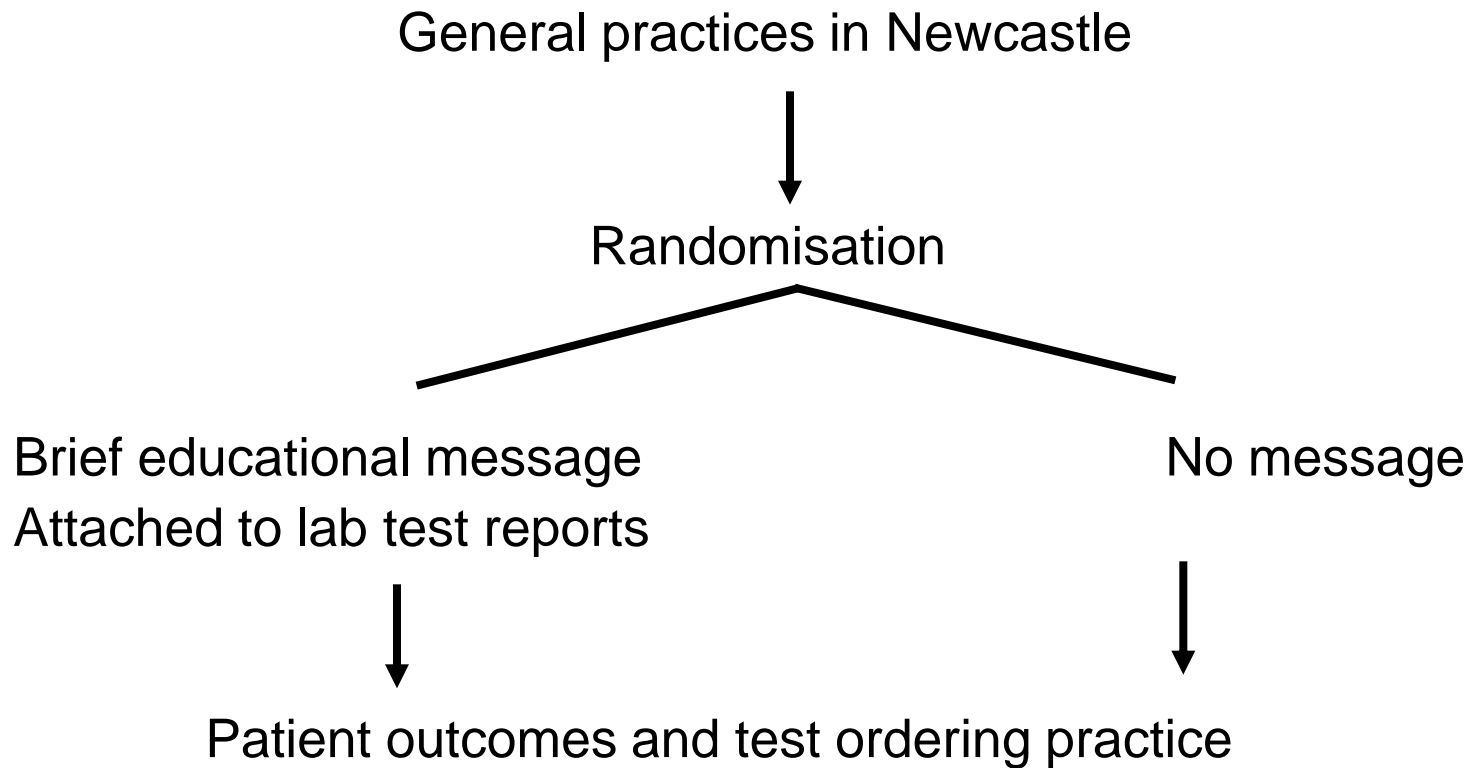
**“ ... rest in bed should be continued for from six to eight weeks to ensure firm cicatrization of the ventricular wall; during the whole of thus period the patient is to be guarded by day and night nursing and helped to avoid voluntary movement or effort. Patients have lost their lives... by neglect of these precautions.”**

Lewis T. Diseases of the heart 1934:49  
(In: Dixon AS. Medical Care 1990;28:202-220)

# How can we tell what works (effectiveness)?



UNIVERSITY OF LEEDS



Foy R, Hawthorne G, Gibb I, et al. A cluster randomised trial of educational prompts in diabetes care. *Implementation Science* 2007; 2: 22

# Beware of the hype e.g. drug company evidence



UNIVERSITY OF LEEDS

- Study designs
- Main outcomes
- Who is evaluating effectiveness?



### A randomised controlled trial of a tailored multifaceted strategy to promote implementation of a clinical guideline on induced abortion care

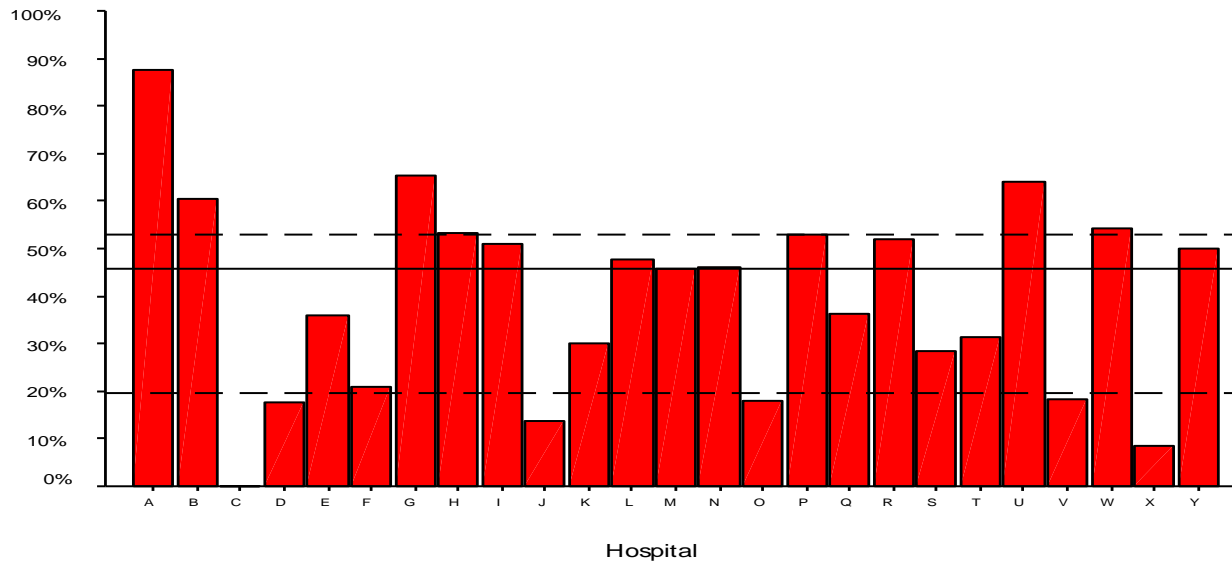
R. Foy,<sup>a</sup> G.C. Penney,<sup>a</sup> J.M. Grimshaw,<sup>b,c</sup> C.R. Ramsay,<sup>d</sup> A.E. Walker,<sup>d</sup>  
G. MacLennan,<sup>d</sup> S.C. Stearns,<sup>e</sup> L. McKenzie,<sup>f</sup> A. Glasier<sup>a,g</sup>

**Objective** To evaluate the effectiveness and efficiency of a tailored multifaceted strategy, delivered by a national clinical effectiveness programme, to implement a guideline on induced abortion.

**Design** Cluster randomised controlled trial.

**Setting and participants** All 26 hospital gynaecology units in Scotland providing induced abortion care.

**Intervention** Following the identification of barriers to guideline implementation, intervention units received a package comprising audit and feedback, unit educational meetings, dissemination of structured case records and promotion of a patient information booklet. Control units received printed guideline summaries alone.





# Plan for the session

- What is implementation?
- Why do we need implementation research, isn't it all commonsense?
- Implementation components & strategies how can we tell what works (effectiveness) and where can we find out more?
- The effectiveness of financial incentives and the importance of rigorous and generalisable implementation
- Can frameworks help us to think what is most likely to work for whom, in what circumstances?



# How can we tell what works?

## Possibility that QI can cause harm

- 12 out of 66 studies on improving diabetes care examined rates of hypoglycaemia
- Hypoglycaemia was more common in intervention groups in 7 out of 12 reports



Shojania, Ranji, McDonald et al. JAMA 2006; 296: 427-40

- **Professional**
  - E.g. audit and feedback, continuing medical education
- **Financial**
  - E.g. general practitioner quality payments
- **Organisational**
  - E.g. expanded roles of nurses, continuous quality improvement
- **Regulatory**
  - E.g. professional accreditation

# What are the limitations of what we know?



UNIVERSITY OF LEEDS

- Systematic review of guideline dissemination & implementation strategies (235 studies)
- Randomised and rigorous quasi-experimental studies
- Objective process or patient outcome measures
- 84 comparisons of single intervention against no intervention control
  - Reminders median effect 14%
  - Educational materials 8%
  - Audit and feedback 7%
- Average effect size ~ 9%



Grimshaw et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. Health Technol Assess 2004; 8(6)

## ***Largely ineffective***

E.g. Dissemination of written educational materials

## ***Variably effective***

E.g. Audit and feedback

## ***Largely effective***

E.g. interactive educational workshops

Bero et al. BMJ 1998; 317:465-8

# Comparing effect sizes



UNIVERSITY OF LEEDS

<b>Treatment</b>	<b>Outcome</b>	<b>Crude absolute effect sizes</b>
Physician advice to quit smoking	Smoking cessation	1-3%
SSRI's versus placebo for depression in primary care	Clinical response	13%
Antibiotics for acute otitis media in children	Pain at 2-7 days	7%
Oral anticoagulation in atrial fibrillation	Stroke prevention	4%
		The Cochrane Library

# What are the limitations of what we know?



UNIVERSITY OF LEEDS

Number of components in study arm	Median effect (% absolute improvement)
1	<b>10.2</b> (n=52)
2	<b>7.4</b> (n=36)
3	<b>11.0</b> (n=25)
4	<b>4.0</b> (n=11)
5	<b>21.8</b> (n=4)
6	<b>15.0</b> (n= 1)
7	-



*It is unclear whether tailored strategies are more effective than non-tailored strategies or no strategy*

Shaw B, Cheater F, Baker R, Gillies C, Hearnshaw H, Flottorp S, Robertson N.  
Tailored interventions to overcome identified barriers to change:  
effects on professional practice and health care outcomes.  
*Cochrane Database of Systematic Reviews* 2005, Issue 3

# How can we tell what works?



UNIVERSITY OF LEEDS

- The costs of some interventions may actually outweigh any potential benefits or anticipated cost savings of a change in practice



Mason et al. *JAMA* 2001:2988-2992

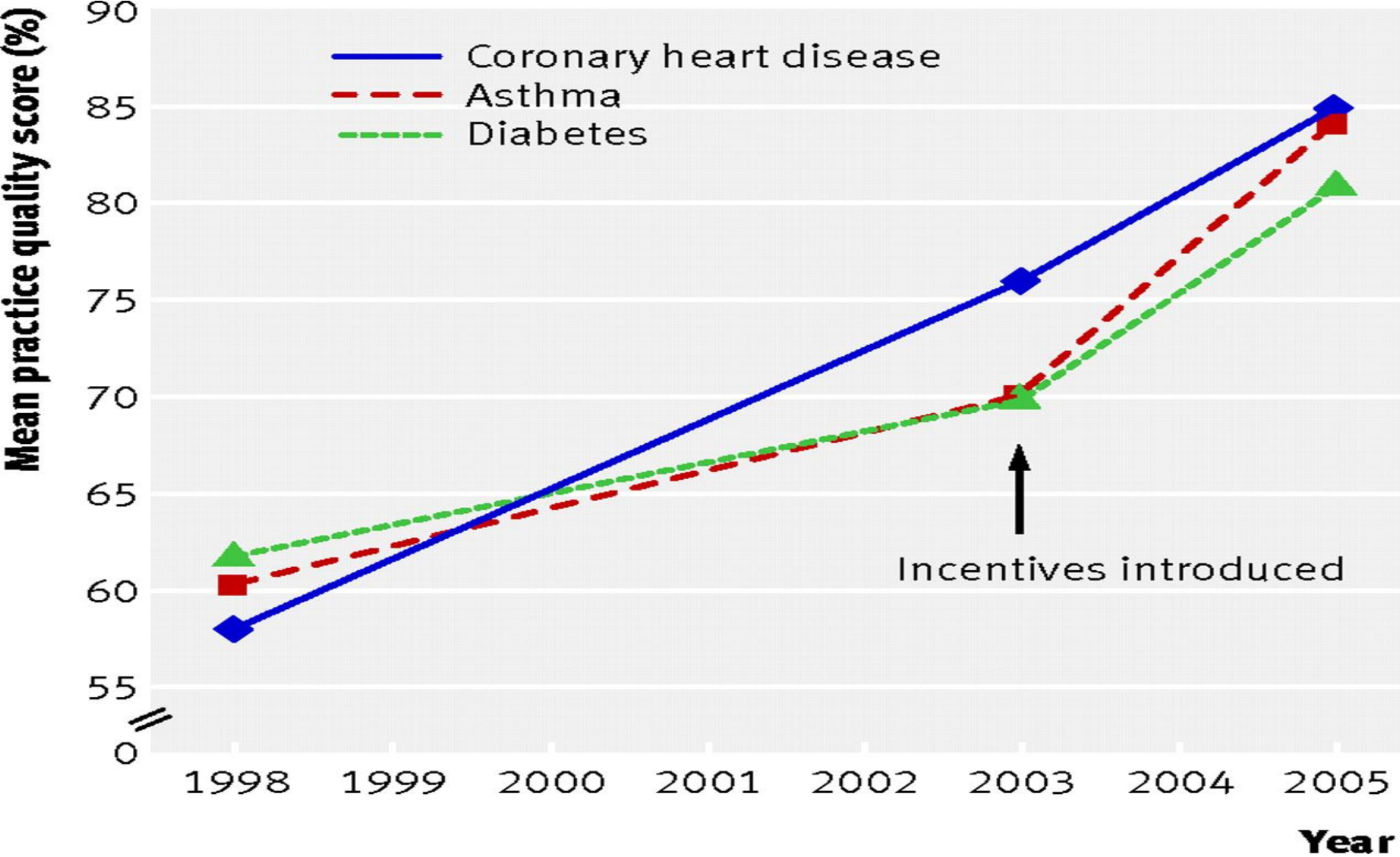


# Plan for the session

- What is implementation?
- Why do we need implementation research, isn't it all commonsense?
- Implementation components & strategies how can we tell what works (effectiveness) and where can we find out more?
- The effectiveness of financial incentives and the importance of rigorous and generalisable implementation
- Can frameworks help us to think what is most likely to work for whom, in what circumstances?



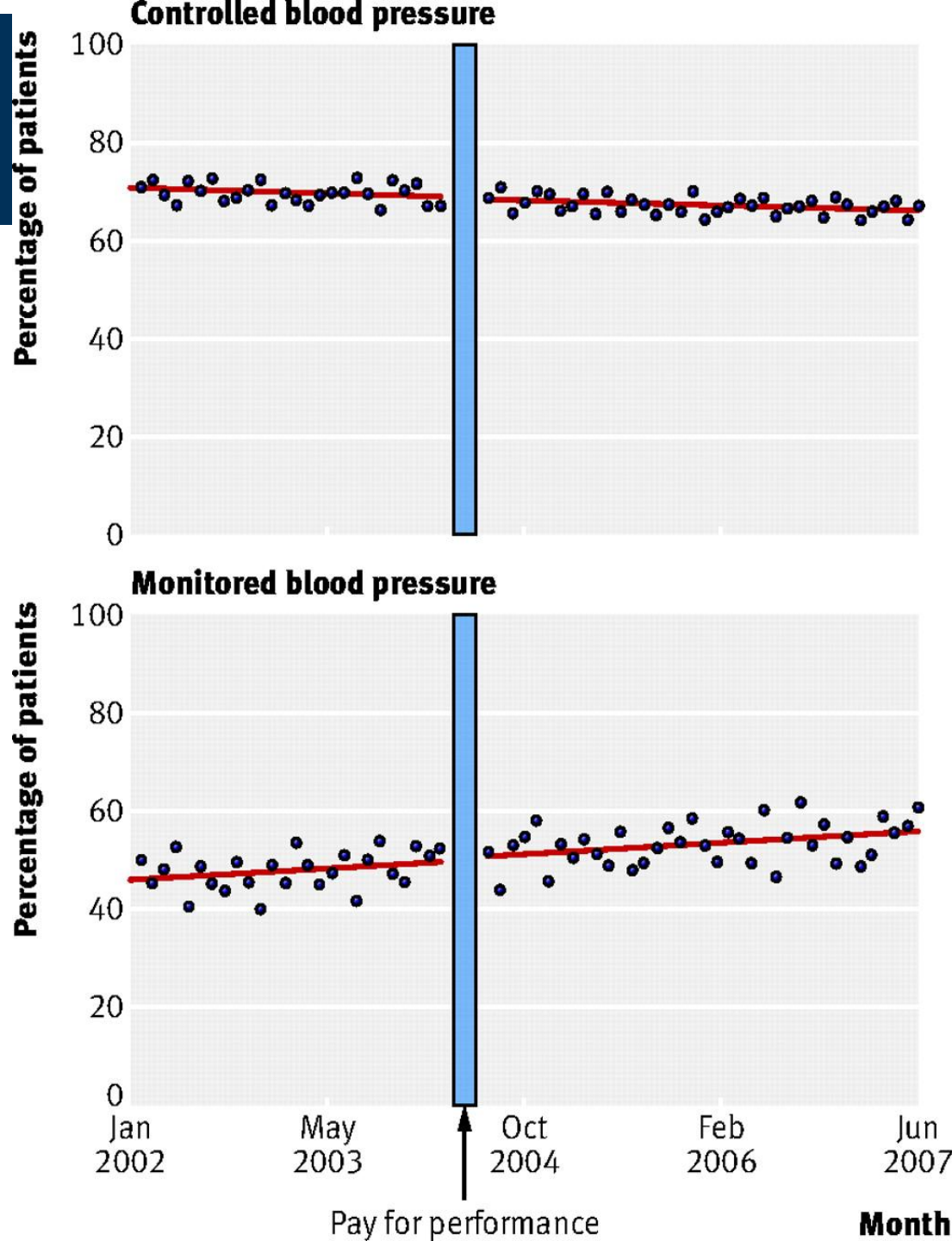
- Quality Outcomes Framework introduced in 2004
- Up to 1000 quality points available per practice
- Each point worth £125
- But did it 'work?'



Mean scores for clinical quality at practice level for coronary heart disease, asthma, and type 2 diabetes, 1998 to 2005. Reproduced with permission from Campbell et al. *N Engl J Med* 2007;357:181-190

Brown, C. et al. *BMJ* 2008;337:a2764

# Effect of pay for performance on blood pressure control and monitoring in UK



BMJ

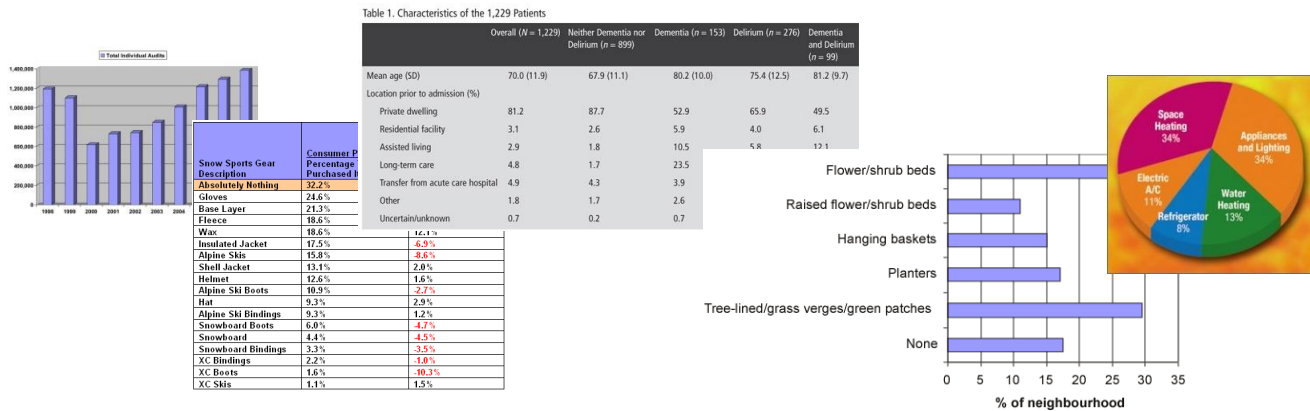
Serumaga B et al. BMJ 2011;342:bmj.d108

# What do we know (or not) about what works?



UNIVERSITY OF LEEDS

*Audit and feedback can be effective in improving professional practice. When it is effective, the effects are generally small to moderate.*



Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD.  
 Audit and feedback: effects on professional practice and health care outcomes.  
 Cochrane Database of Systematic Reviews 2006, Issue 1

# What are the limitations of what we know?



UNIVERSITY OF LEEDS

How best to do audit for diabetes in primary care...

- Does it work for diabetes in primary care?
- Does it work across all dimensions of care?
- Should feedback be comparative or anonymised?
- How intensive should feedback be?
- Who should deliver the feedback?
- What other activities should accompany feedback?
- What to do about the poorest performers?



# What are the limitations of what we know?



UNIVERSITY OF LEEDS

How best to do audit for diabetes in primary care...

**Varying only five elements of audit and feedback produces 288 combinations**

The Improved Clinical Effectiveness through Behavioural Research Group (ICEBeRG). Implementation Science 2006;1:4

# What are the limitations of what we know?



UNIVERSITY OF LEEDS

- NICE produces around 10 clinical guidelines each year relevant to general practice
- Barriers to implementation can vary as much within as between guidelines
- Each guideline probably contains around 4 particularly important recommendations
- An implementation trial for every guideline and/or recommendation is not feasible (not accounting for permutations in interventions, etc)
- Lack of basis for generalising clinical behaviours



# What is an intervention?

What makes for a perfect intervention?

- Cost-effective: benefits outweigh harms and costs
- Sustainable: can be embedded within existing structures and routines
- Predictable: we will know how and when it works best
- Transferable: we also know how it could be adapted for another problem or context



## What makes a complex intervention?

- Number of interacting components within the experimental and control interventions
- Number and difficulty of behaviours required by those delivering or receiving the intervention
- Number of groups or organisational levels targeted by the intervention
- Number and variability of outcomes
- Degree of flexibility or tailoring of the intervention permitted

- What works is not necessarily common sense
- Active interventions not necessarily better than passive interventions
- Multifaceted not necessarily better than single interventions
- We can do harm

# What are the limitations of what we know: Conclusions



UNIVERSITY OF LEEDS

- If you want to show that an intervention works
- ... use a less rigorous study design and be prepared to live with the consequences
  
- But churning out more trials in itself is not sufficient
- ... without building conceptual frameworks with which to describe key elements of targeted behaviours, interventions and contexts, we will fail to build a cumulative understanding of what works



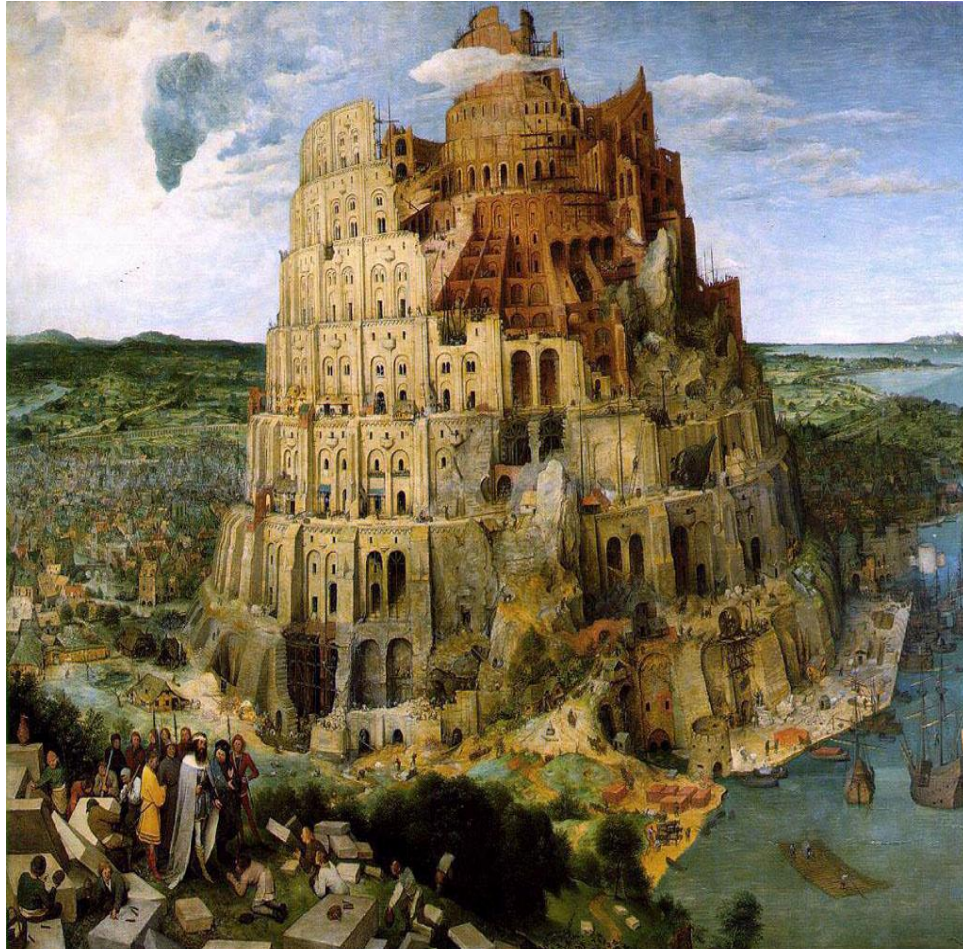
# Plan for the session

- What is implementation?
- Why do we need implementation research, isn't it all commonsense?
- Implementation components & strategies how can we tell what works (effectiveness) and where can we find out more?
- The effectiveness of financial incentives and the importance of rigorous and generalisable implementation
- Can frameworks help us to think what is most likely to work for whom, in what circumstances?

# What do we need to do?



UNIVERSITY OF LEEDS



Develop a common language with which to describe salient features of interventions, targeted behaviours and settings

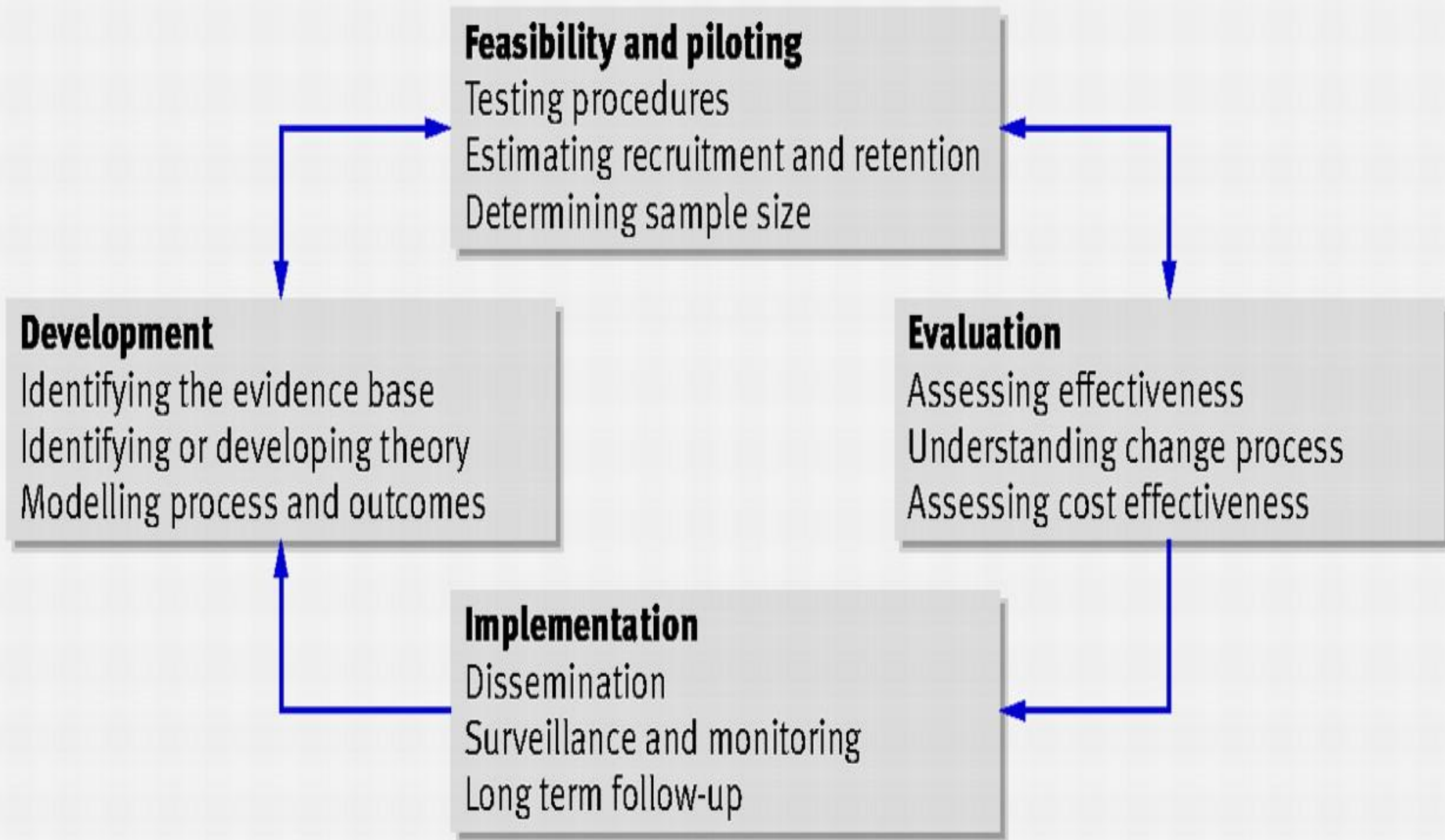
Build a cumulative rather than 'hit and miss' science



## What is a theory?

An organized, heuristic, coherent, and systematic articulation of a set of statements related to significant questions that are communicated in a meaningful whole for the purpose of providing a *generalisable form of understanding*

Meleis A. *Theoretical Nursing: Development and Progress*. New York: Lippincott, 1997



## Key elements of the development and evaluation process

Craig, P. et al. *BMJ* 2008;337:a1655

**Larger system/ environment**

**Organisation**

**Team/ clinical microsystem**

**Individual**

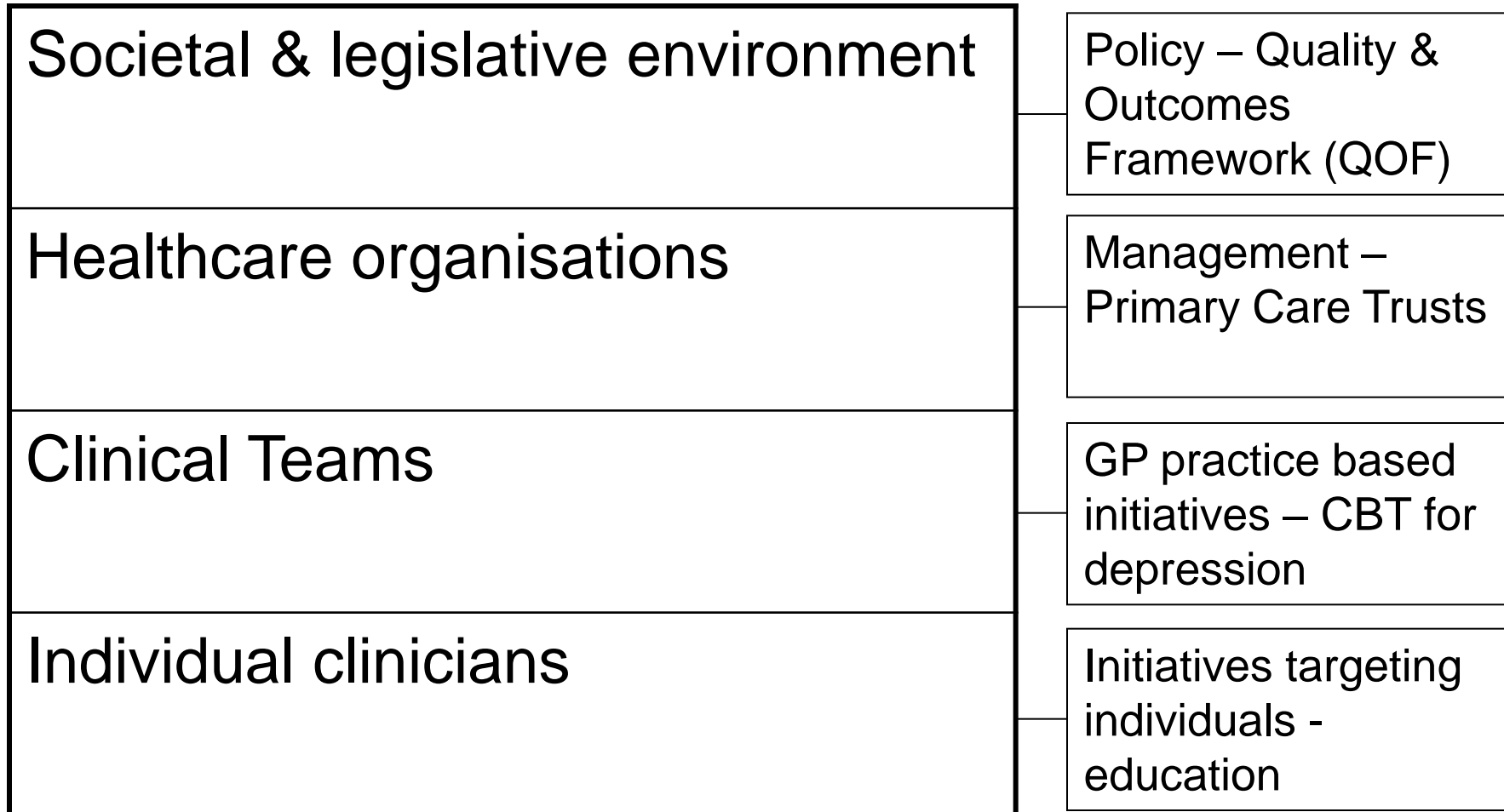


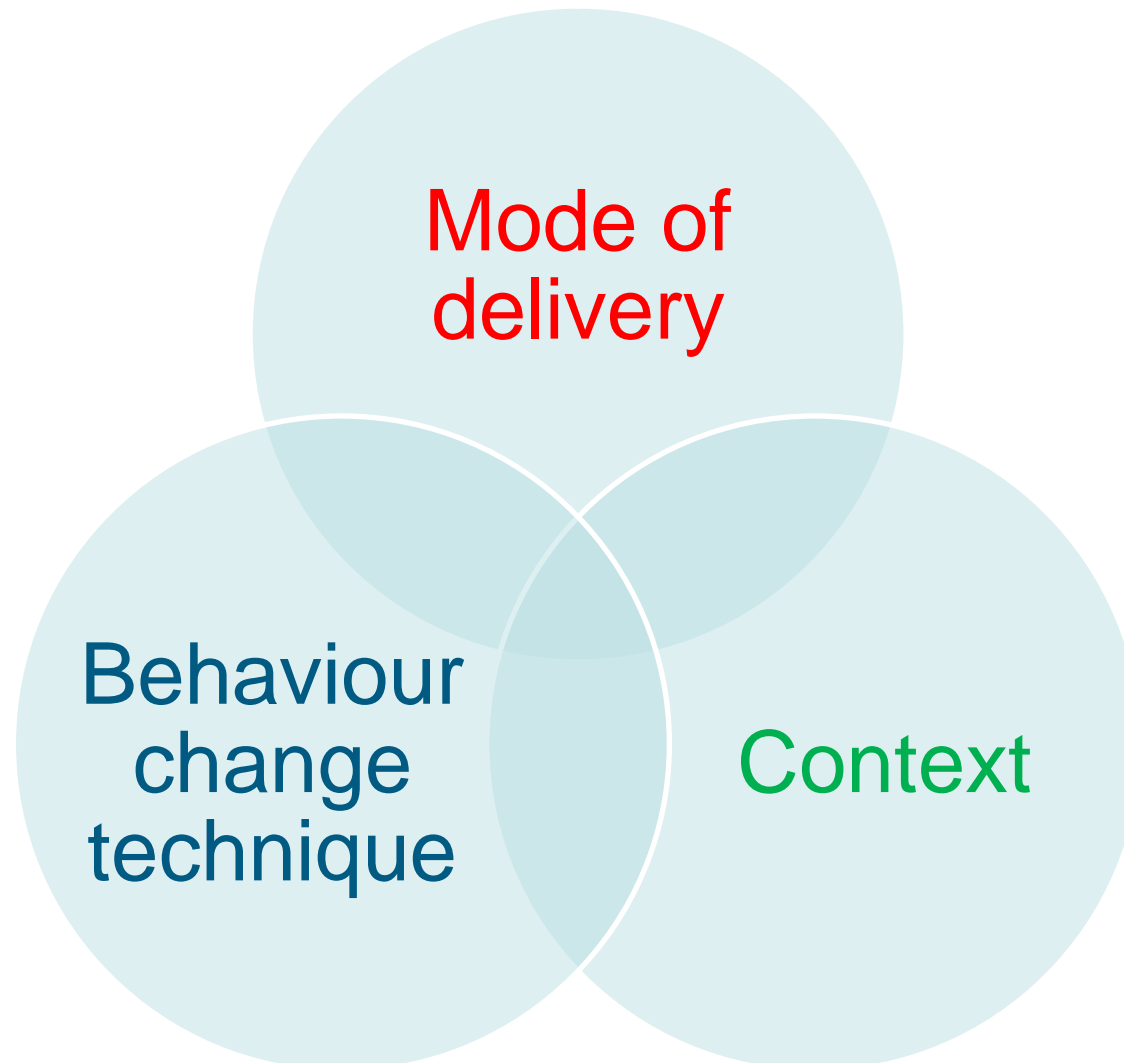
# Multi-level Framework

(Ferlie & Shortell 2001)



UNIVERSITY OF LEEDS







How can we use theories?

1. Explaining behaviour
2. Intervention development (define targetted behaviour, intervention components and contexts)
3. Evaluation of mechanisms of action

Foy et al. *BMJ Qual Saf* 2011;20:453-459

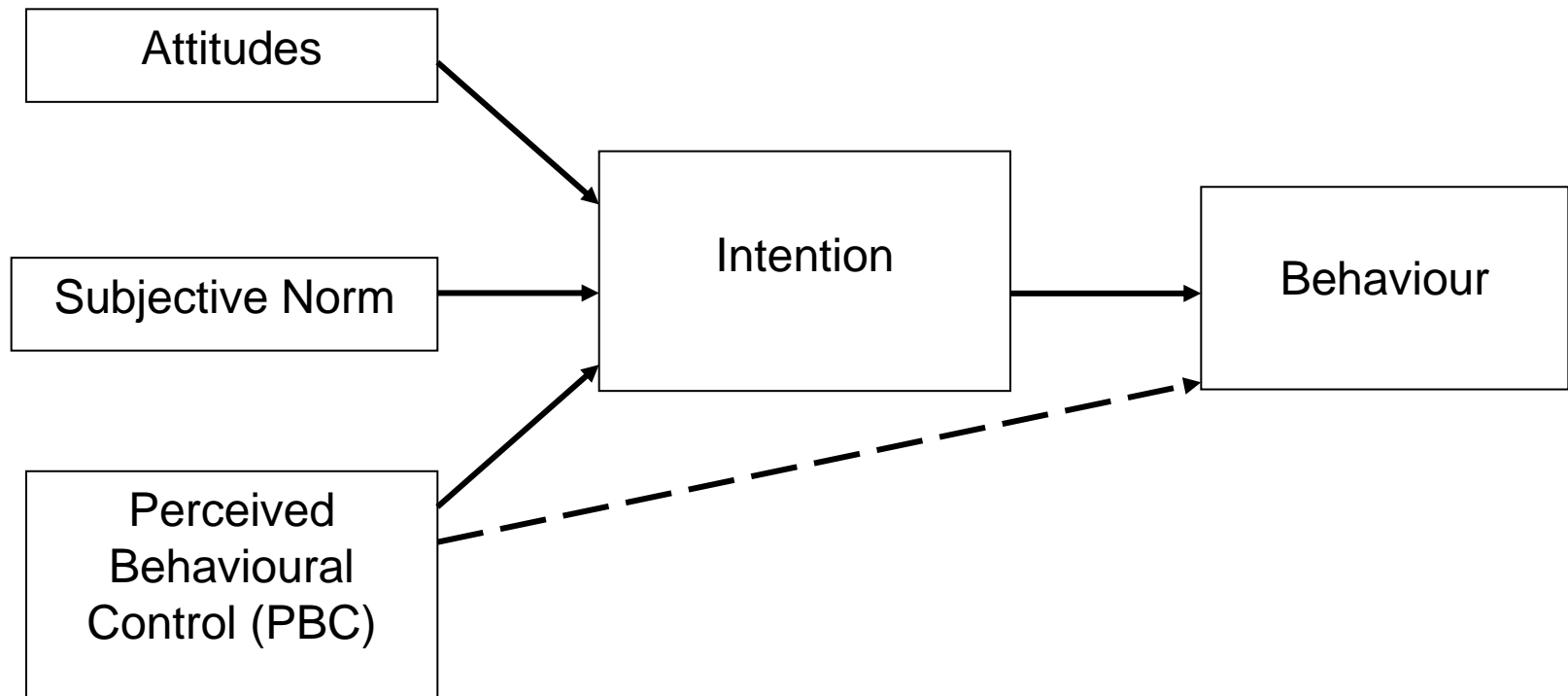


# Re-thinking interventions

- The behaviour targeted
- Likely barriers and enablers to change
- The likely active ingredients of the intervention
- Expectations of the intervention

## High priority contexts for patient safety research

- External Factors such as regulatory requirement, the presence of public reporting or pay-for-performance
- Organizational Characteristics, such as size, complexity and financial status or strength
- Teamwork, leadership and patient safety culture
- Management/Implementation tools such as training resources, internal organization incentives, audit and feedback, etc.



Theory of Planned Behaviour (Ajzen, 1988, 1991)



- Proximal predictor of behaviour
- Substitute measure for behaviour
- I intend to, I aim to, I expect to, I will

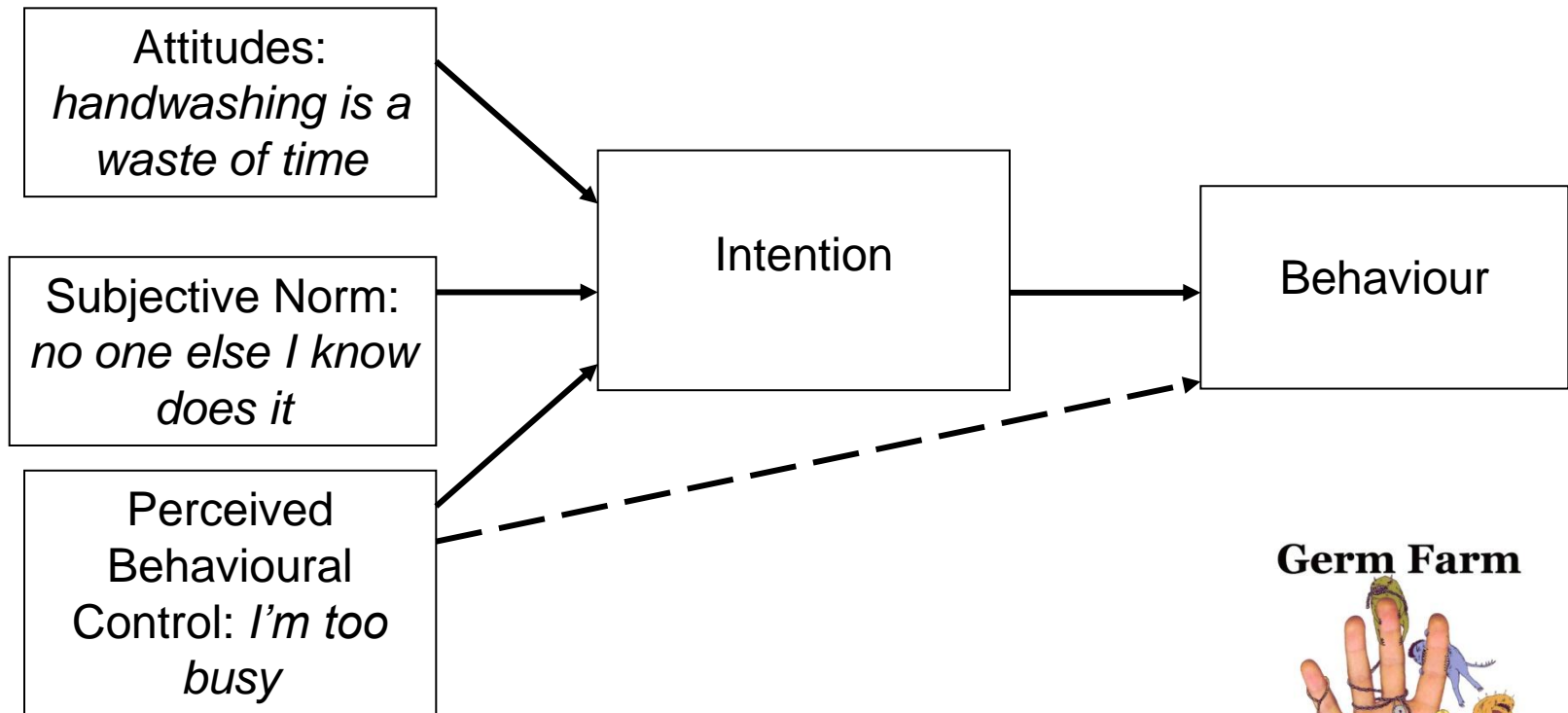


- Clear definition of the behaviour being measured.
  - e.g. “*prescribing antibiotics*”
- Specific & general behaviours.
  - (e.g. “*prescribing antibiotics for sore throat / in general*”)
- “Principle of compatibility”.
  - all measures should have the same level of generality

# Explaining behaviour



UNIVERSITY OF LEEDS



**Germ Farm**



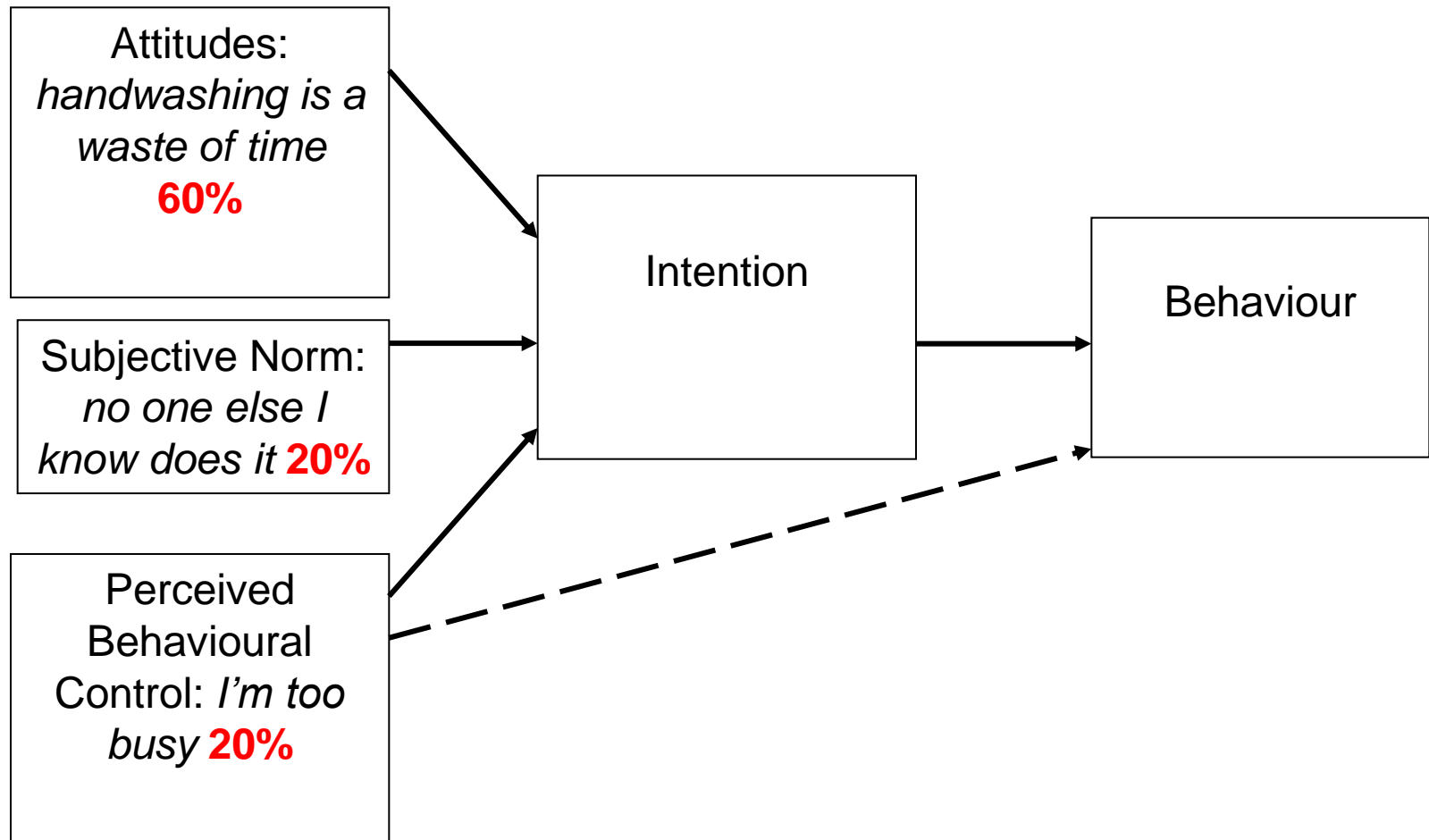
**Scrub'em!**

[www.1st-in-handwashing.com](http://www.1st-in-handwashing.com)

# Explaining behaviour



UNIVERSITY OF LEEDS



<b>Targeted behaviour</b>	<b>Complex/ simple</b>	<b>Habitual/ new</b>
Beta-blockers in heart failure	Complex	New
Avoidance of antibiotics for URTI	Simple	Habitual
Anticoagulation for atrial fibrillation	Complex	New

Foy R, MacLennan G, Grimshaw J, Penney G, Campbell M, Grol R. J Clin Epidemiol 2002; 55: 717-22

<b>Computerised decision support (‘off the shelf’)</b>	
Statement of guideline recommendation	
Guideline endorsed by professional body	
Statement of benefit to patients	
Primed at point of clinical decision	
Signed up to by all or most people in clinical team	

# What are theories good for?



UNIVERSITY OF LEEDS

<b>Computerised decision support ('off the shelf')</b>	<b>Active ingredients (based on behavioural theories)</b>
Statement of guideline recommendation	Knowledge
Guideline endorsed by professional body	Social role and identity
Statement of benefit to patients	Beliefs about consequences
Primed at point of clinical decision	Memory and attention
Signed up to by all or most people in clinical team	Social influence

# Whether an intervention works in *theory* as well as in practice



UNIVERSITY OF LEEDS

Use of theory in a review of studies evaluating effectiveness of implementation strategies

- 53 out of 235 used theory in any way
- Only 14 were explicitly theory-based

Use of (qualitative) process evaluations parallel to randomised trials remains uncommon

Davies P, Walker AE, Grimshaw JM. *Implementation Science* 2010, **5**:14

Lewin S, Glenton C, Oxman A. *BMJ* 2009, **339**:b3496

# What do we need to do?

## Some reflections



UNIVERSITY OF LEEDS

- No one study by itself can answer all salient practice and policy questions, e.g.
  - Parallel process evaluations
  - Cost-effectiveness studies
- Need for interdisciplinary evaluation
- Value of multiple studies across a range of targeted behaviours and settings to understand variation in effects between contexts
- Plausible theories can be wrong or need refining (and randomised trials are usually the ultimate test of a theory)
- Increasing challenges with increasing complexity of interventions

# Make it easier to do the right rather than the wrong thing



UNIVERSITY OF LEEDS



“That's the one point it does get a little bit annoying when that comes up and you think `well I'm seeing them for their big toe'”

# Make it easier to do the right rather than the wrong thing



UNIVERSITY OF LEEDS



“They're off [patient has left], I turn back and go back into this and then select the problem title, and then I say right well I've looked at ischaemic heart disease and then this comes up—and there [system has activated] and the patient's already gone by this stage.”

Rousseau et al. *BMJ* 2003; 326:314-318



# The perfect intervention

- Cost-effective: benefits outweigh harms and costs
- Predictable: we will know how and when it works best
- Sustainable: can be embedded within existing structures and routines

# Whether an intervention works in *theory* as well as in practice



UNIVERSITY OF LEEDS



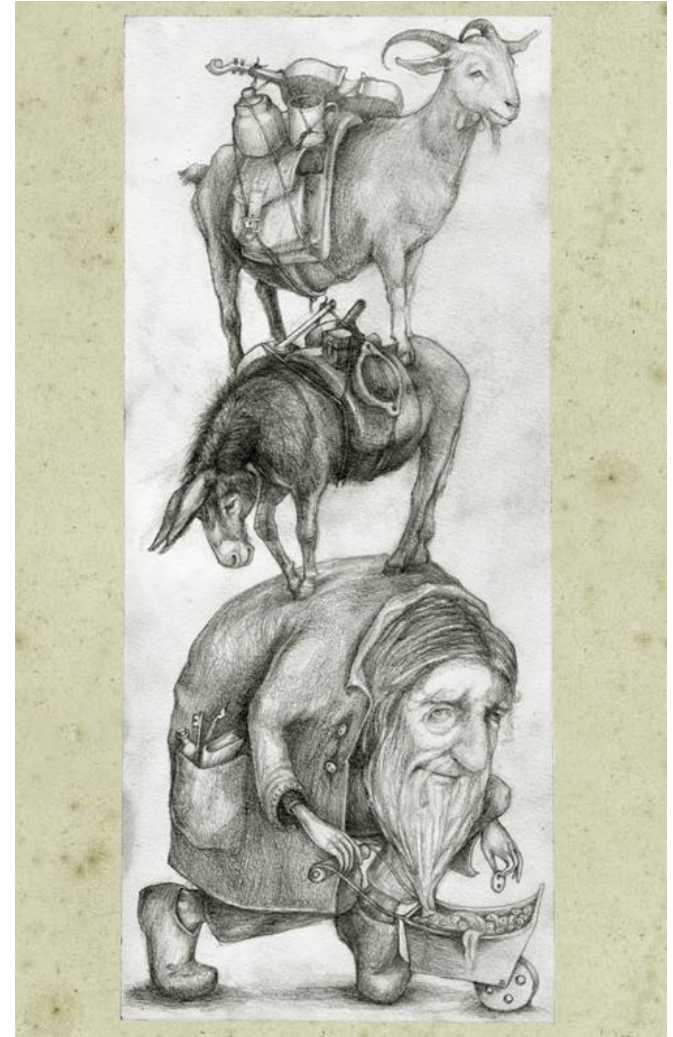
Theory-driven evaluations can enhance generalisability, i.e. inform predictions about effects across different targeted behaviours, interventions and contexts

... thereby building a cumulative understanding of the nature of change

# Why reaching for the clouds isn't so bad



UNIVERSITY OF LEEDS



# Why reaching for the clouds isn't so bad



UNIVERSITY OF LEEDS

- Effect sizes are in range, if not better, than many recommended clinical treatments
- Effect sizes can be worthwhile in relation to costs of interventions
- Effects of quality improvement interventions can be cumulative and complementary
- There is scope for enhancing effect sizes through tailoring of interventions



# A great leap of faith...

UNIVERSITY OF LEEDS



Or small steps?

- Most advancement is incremental and painstaking
- Need for rigorous evidence of effectiveness if promulgating or adopting a given intervention
- Coupled with a better understanding of 'active ingredients' of behaviour and behaviour change



- Beware of the hype
  - **Don't let fashion undermine evidence-based implementation**
- If you want to prove that something works, use a weak study design
  - **Or, use a fair test of effectiveness**
- Reach for the clouds
  - **Modest effects can still be important**
- Make it easier to do the right rather than the wrong thing
  - **Modest effects can be enhanced – although often easier said than done**
- Consider whether an intervention works in *theory* as well as in practice
  - **Understanding how things work allows us to make predictions**
- There are no rules
  - **Evidence and theory are prone to change**

# Identifying and choosing priorities for quality improvement



UNIVERSITY OF LEEDS

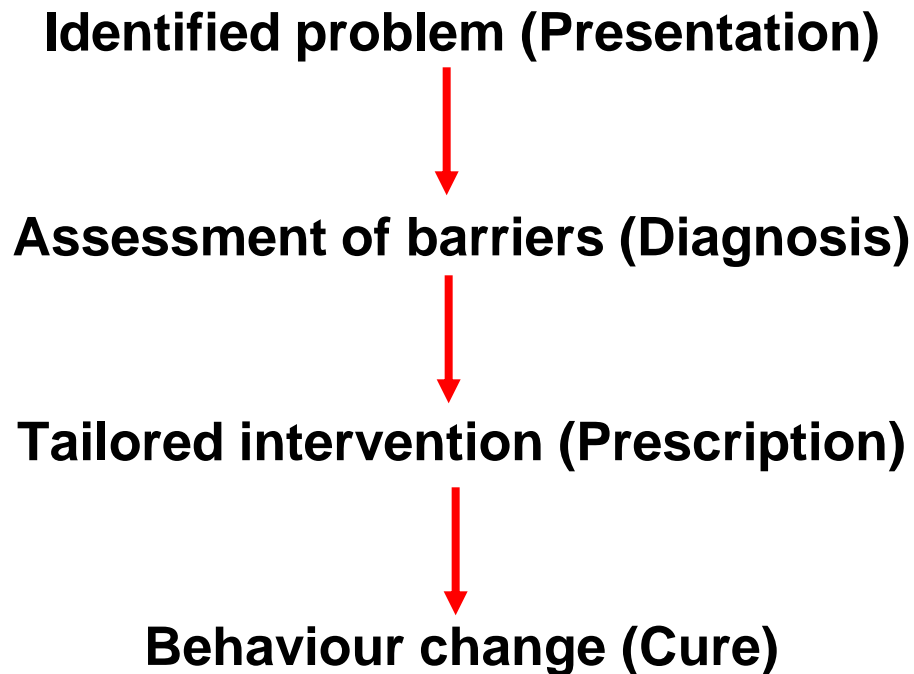
- Given the limited resources available for any health service activity – including quality improvement – we need to prioritise which issues to tackle
- What criteria would you suggest for prioritising quality improvement issues?

# Examples of research and evaluation questions



UNIVERSITY OF LEEDS

<i>Question</i>	<i>Possible design or approach</i>
Why is there a gap between evidence and practice?	Qualitative or quantitative studies
Is this QI intervention effective?	Randomised trial, quasi-experimental design
How does this QI intervention work?	Qualitative or quantitative process evaluation
Has practice changed following our QI intervention?	Before and after study



# Thinking through barriers to and enablers of change



UNIVERSITY OF LEEDS

1 ...

2 ...

3 ...

4 ...

5 ...



Behavioural theory offers a basis for understanding clinical practice and thereby a rationale for strategies to change behaviour

- Consensus process identified main constructs from 33 psychological theories and grouped them into 12 domains
- Subsequent semi-structured interview schedule



- Framework for studying the implementation
- Means to develop strategies for effective implementation
- 14 theoretical domains
  - Component constructs for each domain
  - Interview questions

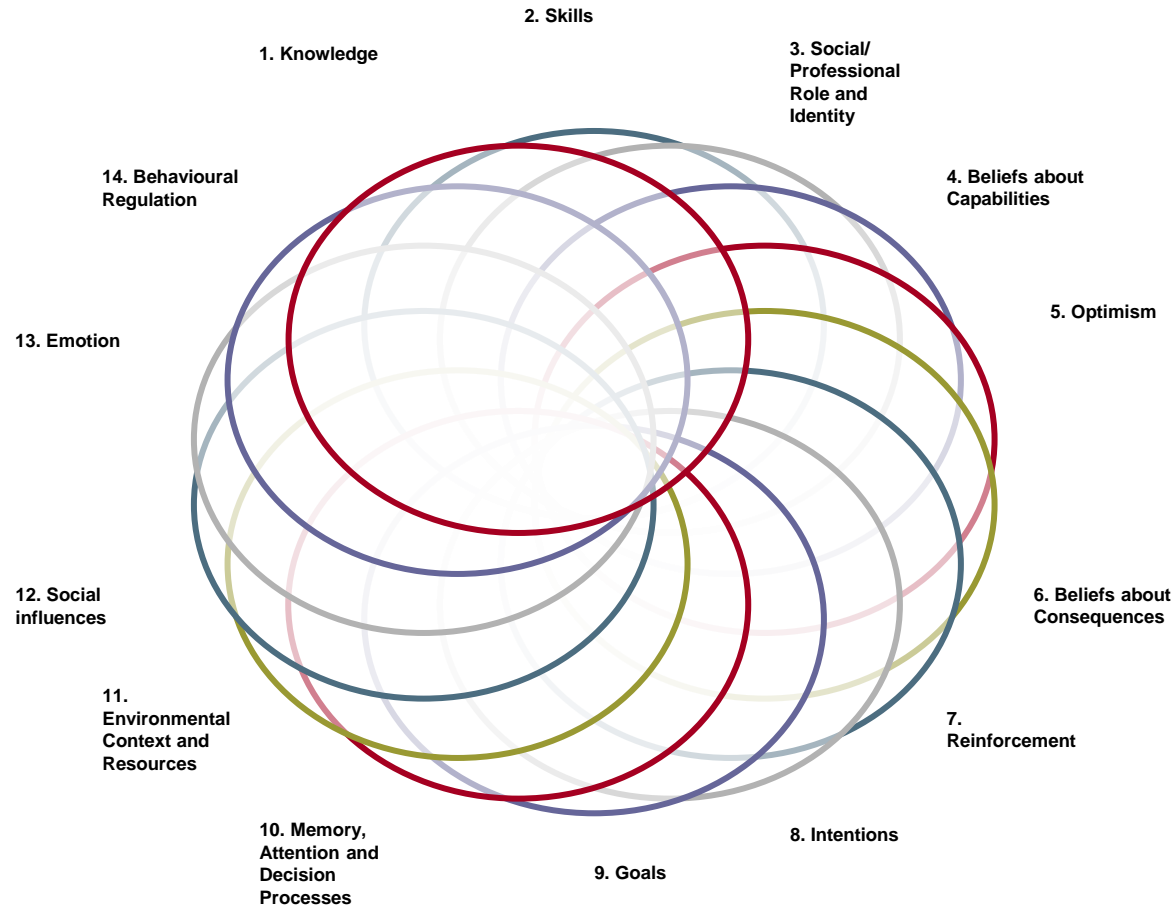
- Consensus process identified main constructs from 33 psychological theories and grouped them into 12 domains subsequently updated to 14 (in press)
- Subsequent Theory Based Implementation Interview (TBII) schedule
- Each domain can subsequently be linked to one or more behaviour change techniques

Michie et al. *Qual Saf Health Care* 2005;14:26-33

Michie et al. *Applied Psych* 2008;57:660-680



# Theoretical Domains Framework

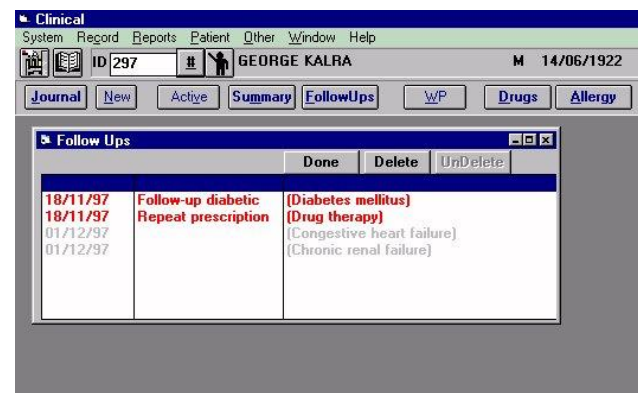


# Linking domains to behaviour change strategies



UNIVERSITY OF LEEDS

- Recommendation: **introducing and monitoring of beta-blockers for heart failure**
- Domain: **Memory and attention processes (limited recall of all steps needed to initiate prescribing, such as ensuring patients have no contraindications to treatment, stepping up dosages, and clinical and biochemical monitoring)**
- Content: **Prompts and triggers**
- Delivery: **On-screen computerised protocol and reminders**

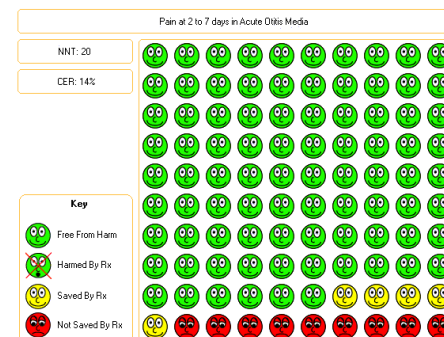


# Linking domains to behaviour change strategies



UNIVERSITY OF LEEDS

- Recommendation: Introducing and monitoring of beta-blockers for heart failure
- Domain: Attitude
- Content: Persuasive communication to challenge sceptical attitudes, e.g. for every 22 patients treated with beta-blockers for a year, one life will be saved
- Delivery: If targeting multiple practices, computerised educational messages; or if targeting particular practices, 1:1 feedback from a respected clinician



# Potential costs and harms of guidelines



UNIVERSITY OF LEEDS

- Health service costs
  - Direct, e.g. prescribing
  - Indirect, e.g. equipment, people
- Opportunity costs
  - Rationing by default
- Failure to base recommendations on valid evidence
  - Purported benefits not realised

What to measure?

- Should criteria be explicit or implicit?
- Should criteria relate to the structure, process or outcomes of care?
- Do the criteria have sufficient impact to lead to improvements in care?
- What level of performance is appropriate to aim for?

How to measure?

- Is the required information available?
- How is an appropriate sample of patients identified?
- How big should the sample be?
- How to choose a representative sample?
- How will the information be collected?
- How will the information be interpreted?



- Structure as the environment in which health care is provided,
- Process as the method by which health care is provided,
- Outcome as the consequence of the health care provided.

- **Patient reported outcome measures**
  - Quality of Life/Health-related Quality of Life
  - Patient satisfaction/experience
- **Externally reported outcomes measures**
  - Mortality
  - Observed morbidity eg amputation rates
- **Intermediate eg biomedical measures**
  - Blood pressure
  - Cholesterol level



# Examples of health indicators

- Doctors per thousand population (structure)
- Waiting times (access)
- Use of perioperative antibiotics (process)
- Vaccination rates (process/intermediate outcome)
- Post operative mortality (outcome)
- Wound infection (outcome)

# Twelve features of a good indicator (ERPHO 2004)



UNIVERSITY OF LEEDS

1. Purposeful – face validity - clear / intuitive / logical
2. Valid – construct validity - does it measure what we want it to measure?
3. Possible to construct (and deconstruct) – right data, right level...
4. Comparable over time periods, between places, and person groups (same denominators/boundaries..)
5. Timely – they must inform, such that timely action is possible
6. Balanced within sets of indicators

# Twelve features of a good indicator (ERPHO 2004)



UNIVERSITY OF LEEDS

7. Consistent/linked between sets of indicators
8. Do not induce perverse incentives and unintentional consequences
9. Statistically robust - they should detect (only) meaningful differences
10. Evidence of a link to the desired outcome - where the indicator is related to infrastructure or process, there should be evidence of a causal link to the desired outcome
11. Are they indicating a point (e.g. an average) or a distribution/difference (e.g. an inequality)?
12. Are they integrated with other organisational and policy driven initiatives?



# Having realistic expectations

- Process of care or clinical end-point?
- What effect size is good enough?

- A range of quality improvement strategies exists BUT...
- We cannot confidently predict which intervention works best for a given targeted behaviour and context HOWEVER...
- We could consider some evidence-informed principles, SUCH AS...
  - Weighing up potential benefits, harms and costs in selecting priorities
  - Being clear about what change is required
  - Using frameworks to help diagnose the causes of quality problems
  - Using an explicit rationale to select and develop intervention strategies

# We know how to do it?



UNIVERSITY OF LEEDS

*Audit and feedback can be effective in improving professional practice. When it is effective, the effects are generally small to moderate.*

Jamtvedt G, Young JM, Kristoffersen DT, O'Brien MA, Oxman AD.  
Audit and feedback: effects on professional practice and health care outcomes.  
*Cochrane Database of Systematic Reviews* 2006, Issue 1

# We know how to do it?



UNIVERSITY OF LEEDS

Questions about how best to do audit for diabetes in primary care

- Does it work for diabetes in primary care?
- Does it work across all dimensions of care?
- Should feedback be comparative or anonymised?
- How intensive should feedback be?
- Who should deliver the feedback?
- What other activities should accompany feedback?
- What to do about the poorest performers?

Varying only five elements of audit and feedback produces 288 combinations

Foy R, Eccles M, Jamvedt G, et al. BMC HSR 2005;5:50

The Improved Clinical Effectiveness through Behavioural Research Group (ICEBeRG). Implementation Science 2006;1:4

Doing nothing is not an option... So what can we do to improve the quality of care?



UNIVERSITY OF LEEDS

Draw upon the evidence base and guiding principles:

- Weighing up potential benefits, harms and costs in selecting priorities
- Being clear about what change is required
- Using frameworks to help diagnose the causes of quality problems
- Using an explicit rationale to select and develop intervention strategies



## Possibility that QI can cause harm

- 12 out of 66 studies on improving diabetes care examined rates of hypoglycaemia
- Hypoglycaemia was more common in intervention groups in 7 out of 12 reports



- The costs of some interventions may actually outweigh any potential benefits or anticipated cost savings of a change in practice
- E.g. educational outreach visits versus passive dissemination

Illustrative domain	Illustrative questions
Beliefs about capabilities (Self-efficacy)	<p><i>How difficult or easy is it for them to do x? (Prompt about internal and external capabilities/constraints)</i></p> <p><i>What problems have they encountered?</i></p> <p><i>What would help them?</i></p>
Social influences (Norms)	<p><i>To what extent do social influences facilitate or hinder x? (prompt about peers, managers, other professionals, patients, relatives)</i></p>

# Explicit rationales for selection of intervention strategies



UNIVERSITY OF LEEDS

Beliefs about capabilities (Self-efficacy)	<i>Graded task</i>  <i>Modelling</i>
Social influences (Norms)	<i>Persuasive communication, e.g. comparative feedback</i>

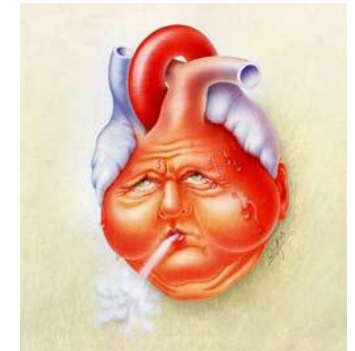
# So what?



UNIVERSITY OF LEEDS

- Allows a better informed judgement about which recommendations are likely to be problematic in advance can guide the level of effort / resources needed to support implementation
- Forms a basis for explicitly linking implementation strategies to different recommendations

A trial shows that an intervention increases the use of beta-blockers for heart failure in primary care



Could the same intervention be applied to:

- Reduce prescribing of antibiotics for upper respiratory tract infections?
- Increase initiation of anticoagulation for atrial fibrillation?

# Thinking through the nature of change required



UNIVERSITY OF LEEDS

<b>Targeted behaviour</b>	<b>Complex/ simple</b>	<b>Habitual/ new</b>
Beta-blockers in heart failure	Complex	New
Avoidance of antibiotics for URTI	Simple	Habitual
Anticoagulation for atrial fibrillation	Complex	New

Foy R, MacLennan G, Grimshaw J, Penney G, Campbell M, Grol R. J Clin Epidemiol 2002; 55: 717-22



- Process of care or clinical end-point?
- What effect size is good enough and realistic?

Doing nothing is not an option... So what can we do to improve the quality of care?



UNIVERSITY OF LEEDS

Draw upon the evidence base and guiding principles:

- Weighing up potential benefits, harms and costs in selecting priorities
- Being clear about what change is required
- Thinking through the nature of change required
- Using frameworks to help diagnose the causes of quality problems
- Think through the likely active ingredients of an intervention
- Having realistic expectations



# What does it all mean?

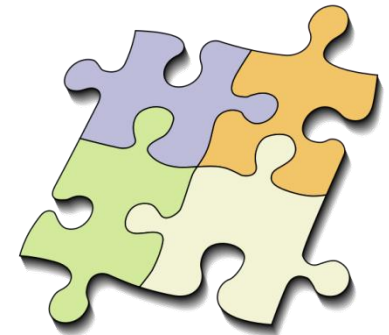
- Characteristics of recommendations can influence the effects of implementation strategies
  - And variations in modifying effect may depend upon the implementation strategy
  - And hard to disentangle recommendation characteristics from targeted groups
- Limited evidence to inform NICE implementation strategies
  - Value of precise wording
  - Incompatibility with professional norms and values is not necessarily a sign of a 'bad' recommendation nor predictor of poor implementation
- ... but a lot of theory

# Where do we go from here?



UNIVERSITY OF LEEDS

- NICE
  - Build in process for judging characteristics of selected key guideline recommendations
  - Judgement about difficulty of implementation and likely ‘best fit’ and proportionate implementation strategies
- Research
  - Development of a typology of recommendation characteristics
  - Evaluation of subsequently tailored implementation strategies



Liz Glidewell  
Faculty of Medicine and Health  
University of Leeds

G12 Charles Thackrah Building  
101 Clarendon Road  
Leeds, United Kingdom  
LS2 9LJ

[l.glidewell@leeds.ac.uk](mailto:l.glidewell@leeds.ac.uk)  
[www.leeds.ac.uk/lihs](http://www.leeds.ac.uk/lihs)

# Implementation Science: unofficial Impact Factor 2.87



UNIVERSITY OF LEEDS



Implementation  
Science

## Co-Editors in Chief

Martin Eccles, Brian Mittman

[Implementationscience.editors@ncl.ac.uk](mailto:Implementationscience.editors@ncl.ac.uk)

## Scope

All aspects of research relevant to the scientific study of methods to promote the uptake of research findings into routine healthcare in both clinical and policy contexts

[www.implementationscience.com](http://www.implementationscience.com)

- Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M: Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008, 337:a1655
- Grol R: Beliefs and evidence in changing clinical practice. *BMJ* 1997, 315:418-421
- Shojania KG, Grimshaw JM: Evidence-based quality improvement: the state of the science. *Health Affairs* 2005, 24:138-150.
- Berwick D. The science of improvement. *JAMA* 2008; 299: 1182-4
- Auerbach AD, Landefeld CS, Shojania KG. The tension between needing to improve care and knowing how to do it. *NEJM* 2007;357:608-13
- Foy R, Ovretveit J, Shekelle PG, Pronovost PJ, Taylor SL, Dy S, Hempel S, McDonald KM, Rubenstein LV, Wachter RM. The role of theory in research to develop and evaluate the implementation of patient safety practices. *BMJ Qual Saf* 2011;20:453-459



## References continued

- Donabedian A. Evaluating the quality of medical care. *Millbank Memorial Fund Quarterly* 1996; 44: 166.
- Sheldon T. The healthcare quality measurement industry: time to slow the juggernaut? *QSHC* 2005; 14: 3 – 4.
- Epstein AM. The outcomes movement: will it get us where we want to go? *N Engl J Med* 1990; 323: 266–70.
- Shanks J, Frater A. Health status, outcome and attributability. *Quality in Health Care* 1993; 2: 259.



- Lilford R, Mohammed M, Spiegelhalter D, Thomson R.
- Use and misuse of process and outcome data in managing performance of acute medical care: avoiding institutional stigma.  
Lancet 2004; 363: 1147-54.
- Mohammed A Mohammed, K K Cheng, Andrew Rouse, Tom Marshall. Bristol, Shipman, and clinical governance: Shewhart's forgotten Lessons.  
Lancet 2001; 357: 463–67.