How Do We Know an Intervention Works?

Paul G. Shekelle
Greater Los Angeles VA Healthcare System
February 4, 2019

Disclosures

- **Research Funding**
  - Agency for Healthcare Research and Quality
  - Department of Veterans Affairs
  - Hospital for Special Surgery
- **Royalties**
  - Up to Date

Two criteria for concluding an intervention works

- Confidence that there is a temporally-appropriate association between an intervention and an outcome unconfounded by other factors
- A mechanism for how the intervention improves the outcome

Roles for Mechanistic Evidence

- To extend evidence of effect from one set of interventions/populations to related interventions/populations
- To support causal conclusions about an intervention and outcome

Roles for Mechanistic Evidence

- To extend evidence of effect from one set of interventions/populations to related interventions/populations
- To support causal conclusions about an intervention and outcome
- Two examples from my own work

Mechanistic evidence to support extending direct evidence to other populations and outcomes

- Key Question: What is the evidence for effect of bariatric surgery compared to nonsurgical treatment of diabetes in patients with a BMI of 30-35 kg/m2?
- Study inclusion criteria were RCTs of patients with diabetes and BMI 30-35 kg/m2, treated with bariatric surgery or non-surgical treatment, measuring outcomes of weight, glucose control, quality of life, and adverse events
- Literatures search: Zero RCTs of bariatric surgery for this population (diabetic patients with BMI 30-35)
- What to do?
  - Turn in “empty” review, i.e., no eligible studies, therefore no evidence available, therefore “very low” or “insufficient” certainty of evidence?
  - Use other evidence to reach a conclusion

Melinda Maggard-Gibbons, MD, MSHS; Margaret Maglione, MPP; Masha Livhits, MD; et al JAMA. 2013;309(21):2250-2261

Other Evidence

- 2 RCTs enrolled diabetic patients with a mean BMI of 37
- 1 RCT enrolled patients with a BMI of 30-35, and 38% had impaired fasting blood glucose
- Total N=290
- 1 observational controlled study of patients with diabetes and BMI 30-34.9
- Numerous case series (uncontrolled) studies of patients with diabetes and BMI in the proper range
- Many RCTs of the effectiveness of nonsurgical therapies of diabetes, mostly with patients whose mean BMI was in the proper range.

Evidence outside our original inclusion criteria

- 2 RCTs enrolled diabetic patients with a mean BMI of 37
- 1 RCT enrolled patients with a BMI of 30-35, and 38% had impaired fasting blood glucose
- Total N=290
  - **Evidence showed much higher proportions of patients achieving target HbA1c control in patients treated with bariatric surgery as compared to nonsurgical therapy, and much greater weight loss as well.**
- 1 observational controlled study of patients with diabetes and BMI 30-34.9
- Numerous case series (uncontrolled) studies of patients with diabetes and BMI in the proper range
- Many RCTs of the effectiveness of nonsurgical therapies of diabetes, mostly with patients whose mean BMI was in the proper range.

Evidence outside our original inclusion criteria

- 2 RCTs enrolled diabetic patients with a mean BMI of 37
- 1 RCT enrolled patients with a BMI of 30-35, and 38% had impaired fasting blood glucose
- Total N=290
  - **Evidence showed much higher proportions of patients achieving target HbA1c control in patients treated with bariatric surgery as compared to nonsurgical therapy, and much greater weight loss as well.**
- 1 observational controlled study of patients with diabetes and BMI 30-34.9
  - **Much greater decrease in HgbA1c and weight loss in patients treated with surgery compared to nonsurgical therapy.**
- Numerous case series (uncontrolled) studies of patients with diabetes and BMI in the proper range
- Many RCTs of the effectiveness of nonsurgical therapies of diabetes, mostly with patients whose mean BMI was in the proper range.

Evidence outside our original inclusion criteria

- 2 RCTs enrolled diabetic patients with a mean BMI of 37
- 1 RCT enrolled patients with a BMI of 30-35, and 38% had impaired fasting blood glucose
- Total N=290
- 1 observational controlled study of patients with diabetes and BMI 30-34.9
- Numerous case series (uncontrolled) studies of patients with diabetes and BMI in the proper range
  - **Large percentages of patients able to achieve blood glucose control without medications**
- Many RCTs of the effectiveness of nonsurgical therapies of diabetes, mostly with patients whose mean BMI was in the proper range.

Evidence outside our original inclusion criteria

- 2 RCTs enrolled diabetic patients with a mean BMI of 37
- 1 RCT enrolled patients with a BMI of 30-35, and 38% had impaired fasting blood glucose
- Total N=290
- 1 observational controlled study of patients with diabetes and BMI 30-34.9
- Numerous case series (uncontrolled) studies of patients with diabetes and BMI in the proper range
  - **Large percentages of patients able to achieve blood glucose control without medications**
- Many RCTs of the effectiveness of nonsurgical therapies of diabetes, mostly with patients whose mean BMI was in the proper range.
  - **On average modest effects on HgbA1c (about 1-2 percentage points) and minimal effect on weight.**

How mechanistic evidence helped reach conclusions

- So, we reasoned that the best explanation for the results from this "other evidence" is that bariatric surgery has an effect on blood glucose control through both a weight loss pathway and (in the case of Roux-en-Y Bypass) a more immediate effect on glucose uptake and metabolism. Additionally, we reasoned that the thresholds used to define the target population are arbitrary cut points on a continuum (the difference between a BMI of 33 and 37 is one of degree, that glucose control is a continuum with IGT being a less severe derangement than diabetes).
- We therefore reached the conclusion of moderate certainty evidence that bariatric surgery provides better control of diabetes than non-surgical therapy in patients with a BMI of 30-35 kg/m2.

Melinda Maggard-Gibbons, MD, MSHS; Margaret Maglione, MPP; Masha Livhits, MD; et al JAMA. 2013;309(21):2250-2261

Mechanistic evidence to support causal conclusions about an intervention and outcome

- Key Question: What is the evidence for effect of colchicine, NSAIDs, and oral steroids in treatment of acute gout?

Shekelle PG, Newberry SJ, FitzGerald JD, Motala A, O'Hanlon CE, Tariq A, et al. Management of Gout: A Systematic Review in Support of an American College of Physicians Clinical Practice Guideline. Ann Intern Med. ;166:37–51.

Eligibility Criteria

- To be included as evidence, a study had to:
  - Enroll adult patients with a diagnosis of acute gout
  - Assess a FDA-approved medication
  - Report patient clinical outcomes
  - Be an RCT if it reported effectiveness outcomes

Results

- Key question #1 Treatment of Acute Gout. Among 49 eligible studies, we identified:
  - Colchicine
    - 2 Placebo-controlled RCTs
  - Steroids
    - 0 Placebo-controlled RCTs
  - NSAIDS
    - 1 Placebo-controlled RCT (n=30)

Results

- Key question #1 Treatment of Acute Gout. Among 49 eligible studies, we identified:
  - Colchicine
    - 2 Placebo-controlled RCTs
  - Steroids
    - 0 Placebo-controlled RCTs
  - NSAIDS
    - 1 Placebo-controlled RCT (n=30)
  - What to do?
    - Conclude “very low” or “insufficient” certainty of evidence?
    - Use other evidence to reach a conclusion

ClinicalEvidence

Gout

Search date September 2010

Martin Underwood

ABSTRACT

INTRODUCTION: Gout affects about 5% of men and 1% of women, with up to 80% of people experiencing a recurrent attack within 3 years. METHODS AND OUTCOMES: We conducted a systematic review and aimed to answer the following clinical questions: What are the effects of treatments for acute gout? What are the effects of treatments to prevent gout in people with prior acute episodes? We searched: Medline, Embase, The Cochrane Library, and other important databases up to September 2010 (Clinical Evidence reviews are updated periodically; please check our website for the most up-to-date version of this review). We included harms alerts from relevant organisations such as the US Food and Drug Administration (FDA) and the UK Medicines and Healthcare products Regulatory Agency (MHRA). RESULTS: We found 16 systematic reviews, RCTs, or observational studies that met our inclusion criteria. We performed a GRADE evaluation of the quality of evidence for interventions. CONCLUSIONS: In this systematic review, we present information relating to the effectiveness and safety of the following interventions: colchicine, corticosteroids, corticotropin (ACTH), non-steroidal anti-inflammatory drugs (NSAIDs), sulfinpyrazone, xanthine oxidase inhibitors, advice to lose weight, advice to reduce alcohol intake, and advice to reduce dietary intake of purines.

QUESTIONS

INTERVENTIONS

TREATING ACUTE GOUT

Likely to be beneficial

Unknown effectiveness

PREVENTION OF RECURRENCE

Unknown effectiveness

ClinicalEvidence

Gout

Search date September 2010

Martin Underwood

- We don’t know whether NSAIDs reduce pain and tenderness in an acute attack of gout, although they are commonly used in clinical practice. They are associated with increased risks of gastrointestinal, and possible cardiovascular, adverse effects.
- Colchicine may be more effective than placebo at improving symptoms in acute gout. Its use is limited by the high incidence of adverse effects; although these may be reduced with low-dose colchicine regimens.
- We don’t know whether intra-articular or parenteral corticosteroids, or corticotrophin (ACTH), improve symptoms in acute gout.

What else do we know about acute gout and its treatment ?

- An acute attack is caused by an inflammatory response

What else do we know about acute gout and its treatment ?

- An acute attack is caused by an inflammatory response
  - **Colchicine, NSAIDs, and corticosteroids are all anti-inflammatory treatments**

What else do we know about acute gout and its treatment ?

- An acute attack is caused by an inflammatory response
  - Colchicine, NSAIDs, and corticosteroids are all anti-inflammatory treatments
- NSAIDs are FDA-approved based on RCTs for temporary pain relief of acute painful conditions

What else do we know about acute gout and its treatment ?

- An acute attack is caused by an inflammatory response
  - Colchicine, NSAIDs, and corticosteroids are all anti-inflammatory treatments
- NSAIDs are FDA-approved based on RCTs for temporary pain relief of acute painful conditions
  - **Gout flare is an acute painful condition**

What else do we know about acute gout and its treatment ?

- An acute attack is caused by an inflammatory response
  - Colchicine, NSAIDs, and corticosteroids are all anti-inflammatory treatments
- NSAIDs are FDA-approved based on RCTs for temporary pain relief of acute painful conditions
  - Gout flare is an acute painful condition
- Corticosteroids are proven effective in the acute inflammation phase of other inflammatory arthritides

What else do we know about acute gout and its treatment ?

- An acute attack is caused by an inflammatory response
  - Colchicine, NSAIDs, and corticosteroids are all anti-inflammatory treatments
- NSAIDs are FDA-approved based on RCTs for temporary pain relief of acute painful conditions
  - Gout flare is an acute painful condition
- Corticosteroids are proven effective in the acute inflammation phase of other inflammatory arthritides
  - **1950 Nobel Prize in medicine was awarded to Philip S. Hench for the treatment of acute RA with cortisone**

What else do we know about acute gout and its treatment ?

- An acute attack is caused by an inflammatory response
  - Colchicine, NSAIDs, and corticosteroids are all anti-inflammatory treatments
- NSAIDs are FDA-approved based on RCTs for temporary pain relief of acute painful conditions
  - Gout flare is an acute painful condition
- Corticosteroids are proven effective in the acute inflammation phase of other inflammatory arthritides
  - 1950 Nobel Prize in medicine was awarded to Philip S. Hench for the treatment of acute RA with cortisone
- There are 6 RCTs that compare NSAIDs vs. corticosteroids with no placebo group

What else do we know about acute gout and its treatment ?

- An acute attack is caused by an inflammatory response
  - Colchicine, NSAIDs, and corticosteroids are all anti-inflammatory treatments
- NSAIDs are FDA-approved based on RCTs for temporary pain relief of acute painful conditions
  - Gout flare is an acute painful condition
- Corticosteroids are proven effective in the acute inflammation phase of other inflammatory arthritides
  - 1950 Nobel Prize in medicine was awarded to Philip S. Hench for the treatment of acute RA with cortisone
- There are 6 RCTs that compare NSAIDs vs. corticosteroids with no placebo group
  - **All 6 trials report equivalency in outcomes. Therefore, either both are about equally effective or equally ineffective.**

We used this "other evidence" in our decisions about strength of evidence

Conclusion for acute gout Rx	Strength of Evidence
Colchicine reduces pain	High
NSAIDs reduce pain	High
Corticosteroids reduce pain	High

The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute?

Jeremy Howick[1] • Paul Glasziou[1] • Jeffrey K Aronson[2]

[1] Centre for Evidence-Based Medicine, Rosemary Rue Building, Old Road Campus, University of Oxford, Oxford OX3 7LF

[2] Department of Primary Health Care, University of Oxford, Oxford

Correspondence to: Jeremy Howick. E-mail: Jeremy.howick@dphpc.ox.ac.uk

DECLARATIONS

Competing interests
None declared

Funding
None

Ethical approval
Not applicable

Guarantor
JH

Contributorship
This paper was a truly collaborative effort that resulted from a series of meetings attended by all three authors. JH produced the initial draft and was in charge of revising subsequent drafts. PG provided insights about the Mother's Kiss example, and

'A main cause of philosophical disease – a one-sided diet: one nourishes one's thinking with only one kind of example.' Ludwig Wittgenstein

Introduction: when non-RCT evidence is sufficient to conclude that the intervention caused the outcome

High quality randomized controlled trials (RCTs) (concealed allocation, relevant groups blinded and sufficiently powered, etc.) will usually provide sufficient evidence to establish that a particular treatment caused an outcome. Yet sufficiently well-conducted RCTs are rare.[1] Trials can be under-powered,[2] or unsuccessfully blinded,[3,4] and often suffer from many undetected biases. The results of most RCTs are therefore often insufficient to establish causation. At the same time, RCTs are often not required to establish causation.[5] Treatments including the Heimlich manoeuvre, cardiac defibrillation and parachutes to prevent death[6] have never been tested in RCTs, yet their effectiveness is surely strongly supported by evidence.

Evidence-grading systems that place randomized trials at the top of a hierarchy[7–13] will deliver before concluding causation. We investigated and revised the Bradford Hill 'guidelines for causation', in order to refine our intuitions about whether to believe that intervention is effective. Our intention is not to debunk previous attempts to grade evidence, but rather to contribute to their natural evolution and development.

Revising Bradford Hill's guidelines

We believe that Bradford Hill's guidelines form a useful tool as they stand. Nevertheless, they can be modified in ways that make them easier to use. For instance, some of the guidelines, such as 'specificity' can safely be omitted while others, such as 'experiment' and 'strength' can be combined; still others, such as 'biological plausibility' can benefit from a more detailed analysis. Moreover, the guidelines have an inherent structure that is unclear in the original exposition. We propose that the guidelines be organized into the following three categories:

(1) *Direct evidence* from studies (randomized or non-randomized) that a probabilistic association between intervention and

The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute?

Jeremy Howick[1] • Paul Glasziou[1] • Jeffrey K Aronson[2]

[1] Centre for Evidence-Based Medicine, Rosemary Rue Building, Old Road Campus, University of Oxford, Oxford OX3 7LF

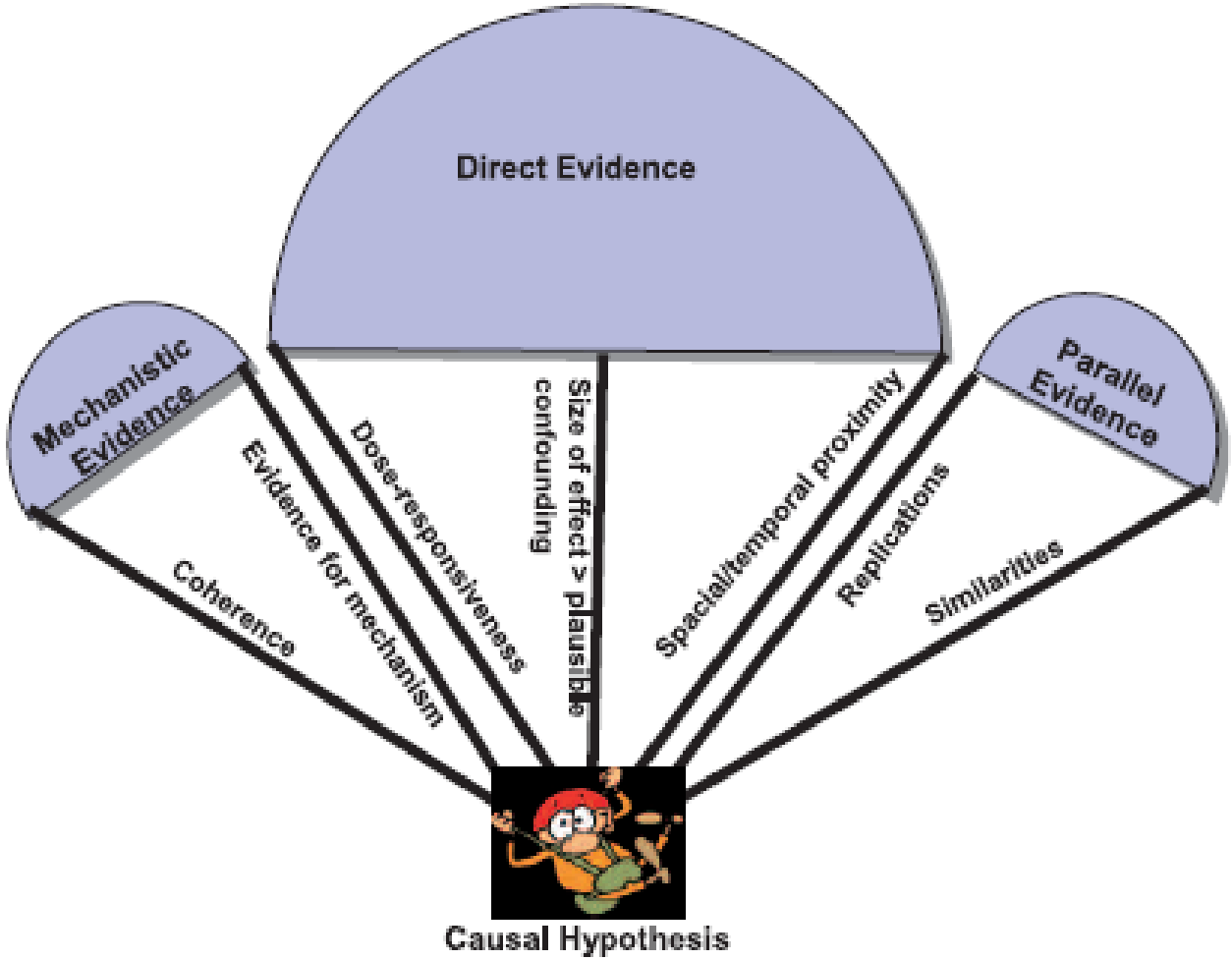
[2] Department of Primary Health Care, University of Oxford, Oxford

Correspondence to: Jeremy Howick. E-mail: Jeremy.howick@dphpc.ox.ac.uk

(1) *Direct evidence* from studies (randomized or non-randomized) that a probabilistic association between intervention and outcome is causal and not spurious;
(2) *Mechanistic evidence* for the alleged causal process that connects the intervention and the outcome;
(3) *Parallel evidence* that supports the causal hypothesis suggested in a study, with related studies that have similar results.

The evolution of evidence hierarchies: what can Bradford Hill's 'guidelines for causation' contribute?

Jeremy Howick, Paul Glasziou, Jeffrey K Aronson J R Soc Med 2009: 102: 186–194



Evidence-based mechanistic reasoning

Jeremy Howick[1] • Paul Glasziou[1] • Jeffrey K Aronson[2]

[1] Centre for Evidence-Based Medicine, Department of Primary Health Care, University of Oxford

[2] Department of Primary Health Care, University of Oxford

Correspondence to: Jeremy Howick. E-mail: Jeremy.howick@dphpc.ox.ac.uk

DECLARATIONS

Competing interests
None declared

Funding
None

Ethical approval
Not applicable

Guarantor
JH

Contributorship
JH prepared the first manuscript based on his philosophical research on mechanisms. PG provided the examples of 'high quality' mechanistic reasoning. JKA introduced the idea of mechanistic 'stages' and

An evidential role for mechanisms

Systematic reviews of high quality randomized trials generally count as the 'best evidence'.[1] However, well-conducted randomized trials are sometimes unavailable,[2,3] unfeasible,[4] unethical[5] or unnecessary.[6,7] In such cases other forms of evidence must be considered. Many EBM proponents accept mechanistic reasoning ('pathophysiologic rationale') for generalizability,[1,8] hypothesis generation,[9] ruling out implausible hypotheses,[10,11] and for supporting efficacy in the absence of other 'stronger' forms of evidence. Yet because mechanistic reasoning has often led us astray,[12,13] most EBM proponents are justifiably sceptical about using mechanistic reasoning as evidence for efficacy.

We suggest that the scepticism about the value of mechanistic reasoning should not extend to high quality mechanistic reasoning. Just as poor quality randomized trials (that are unblinded,[14–16] underpowered or biased,[17] that employ unconcealed allocation,[15,16] or otherwise biased) will not provide high quality evidence for efficacy, so poor quality mechanistic reasoning will be unreliable. In this theoretical exploration we suggest that mechanistic reasoning involving a not incomplete

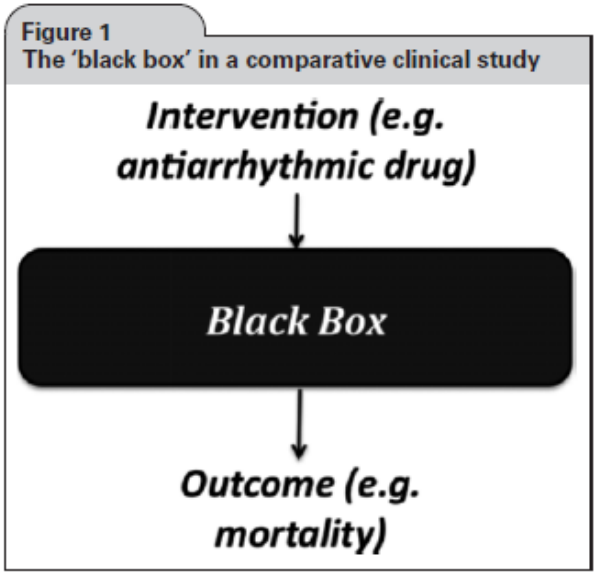


Figure 1
The 'black box' in a comparative clinical study

intervention and the clinically relevant outcome. For example, a randomized trial of antiarrhythmic drugs versus placebo suggested that the drugs unexpectedly increased mortality by 3.3%.[19] This conclusion did not rely on an explanation of how they did so – that remained a 'black box' (Figure 1).

Mechanistic reasoning involves looking inside the 'black box', and relies on knowledge of the

Potential Pitfalls to Mechanistic Reasoning

- Howick, Glasziou, and Aronson argue that analogous to RCTs, mechanistic reasoning can be high quality or poor quality. Examples where mechanistic reasoning has failed due to poor quality should not result in a conclusion that mechanistic reasoning is an invalid method, any more than an RCT that failed due to poor quality means that the RCT is an invalid method.
- Howick, Glasziou, and Aronson identify two categories of poor quality mechanistic reasoning
  - Empty or partial mechanisms
  - Probabilistic and complex nature of mechanisms

Jeremy Howick, Paul Glasziou, Jeffrey K Aronson, J R Soc Med 2010: 103: 433–441

The role of mechanistic reasoning in evidence frameworks

- Mechanistic evidence is a requirement to reach conclusions that an intervention works
- In many, if not most, situations mechanistic evidence takes a subsidiary role to empirical evidence such as that coming from RCTs.
- However, there are situations where mechanistic evidence is sufficiently strong that when combined with uncontrolled clinical observations conclusions of intervention effectiveness have high certainty.
- Characteristics of such situations include:
  - The mechanism is strongly supported in all steps, and the number of intermediate steps is small
  - The time between the intervention and the change in the outcome is short
  - The difference in the outcome between the treated and untreated state is obvious.
  - Alternative explanations for the difference in outcome are far less plausible

THE END