Mechanisms and the evidence hierarchy

Federica Russo

Dipartimento di Studi Umanistici, Università di Ferrara

Topoi
DOI 10.1007/s11245-013-9220-9

Mechanisms and the Evidence Hierarchy

Brendan Clarke · Donald Gillies · Phyllis Illari · Federica Russo · Jon Williamson

Abstract Evidence-based medicine (EBM) makes use of explicit procedures for grading evidence for causal claims. Normally, these procedures categorise evidence of correlation produced by statistical trials as better evidence for a causal claim than evidence of mechanisms produced by other methods. We argue, in contrast, that evidence of mechanisms needs to be viewed as complementary to, rather than inferior to, evidence of correlation. In this paper we first set out the case for treating evidence of mechanisms alongside evidence of correlation in explicit protocols for evaluating evidence. Next we provide case studies which exemplify the ways in which evidence of mechanisms complements evidence of correlation in practice. Finally, we put forward some general considerations as to how the two sorts of evidence can be more closely integrated by EBM.

B. Clarke · D. Gillies · P. Illari (✉)
Science and Technology Studies, University College London, London, UK
e-mail: phyllis.illari@ucl.ac.uk

B. Clarke
e-mail: b.clarke@ucl.ac.uk

D. Gillies
e-mail: donald.gillies@ucl.ac.uk

F. Russo
Dipartimento di Studi Umanistici, Università degli Studi di Ferrara, Ferrara, Italy
e-mail: federica.russo@unife.it

J. Williamson
Philosophy, SECL, University of Kent, Canterbury, UK
e-mail: j.williamson@kent.ac.uk

Published online: 03 December 2013

Keywords Mechanism · Difference-making · Evidence · Evidence of mechanism · Evidence in medicine · Evidence-based medicine

1 Introduction

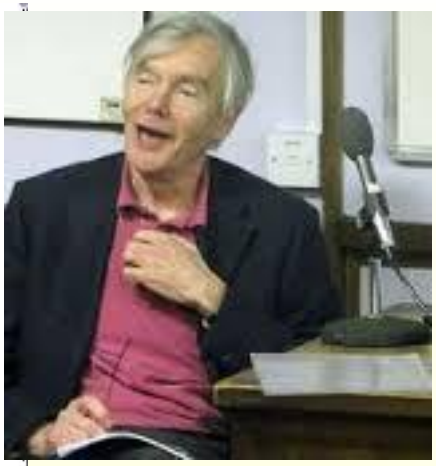
Sackett et al. (1996) characterise evidence-based medicine (EBM) as follows:

> Evidence based medicine is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients.

In order to make decisions about patient care, one typically needs to diagnose—to determine the most probable cause of the patient's symptoms—and treat—to determine which treatment intervention is most likely to alleviate the diagnosed causes. Thus one needs to establish what causes what and one needs to apply this causal knowledge to new patients. This paper is concerned with methods for establishing and using causal claims, particularly in EBM.

The EBM movement has transformed the way in which evidence is gathered and evaluated in medicine. Medical researchers and those charged with making treatment and public health decisions now tend to be guided by explicit *evidence hierarchies*. An evidence hierarchy ranks evidence for a causal claim. Table 1, for example, depicts an evidence hierarchy advocated by the UK National Institute for Health and Care Excellence for evaluating treatment effectiveness, while Table 2 is a corresponding hierarchy for evaluating diagnostic claims (NICE 2006). More recently, NICE advocated the GRADE system depicted in Table 3, which highlights the main point of commonality between the plethora of evidence hierarchies that abound in the literature: randomised trials (RCTs) are ranked more

Springer



Overview

The background

- Evidence-based medicine
- The ‘Russo-Williamson’ Thesis

Mechanisms and evidence evaluation

- Integration of evidence
- Guidelines for using evidence of mechanisms

The background

EVIDENCE-BASED MEDICINE

Evidence-based medicine de-emphasizes intuition, unsystematic clinical experience, and pathophysiological rationale as sufficient grounds for clinical decisions making and stresses the examination of evidence from clinical research.

(Evidence-based working group, 1992)

Evidence based medicine is the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients.

(Sackett et al., 1996)

What is ‘best evidence’?

Evidence hierarchies

Developed by e.g.:

NICE, Oxford Centre for Evidence-based medicine, …

A common claim:

RCTs are better than observational studies,

which are better than any other type of evidence

and better than mechanistic reasoning

Oxford Centre for Evidence-Based Medicine 2011 Levels of Evidence

Question	Step 1 (Level 1*)	Step 2 (Level 2*)	Step 3 (Level 3*)	Step 4 (Level 4*)	Step 5 (Level 5)
How common is the problem?	Local and current random sample surveys (or censuses)	Systematic review of surveys that allow matching to local circumstances**	Local non-random sample**	Case-series**	n/a
Is this diagnostic or monitoring test accurate? (Diagnosis)	Systematic review of cross sectional studies with consistently applied reference standard and blinding	Individual cross sectional studies with consistently applied reference standard and blinding	Non-consecutive studies, or studies without consistently applied reference standards**	Case-control studies, or "poor or non-independent reference standard"	Mechanism-based reasoning
What will happen if we do not add a therapy? (Prognosis)	Systematic review of inception cohort studies	Inception cohort studies	Cohort study or control arm of randomized trial*	Case-series or case-control studies, or poor quality prognostic cohort study**	n/a
Does this intervention help? (Treatment Benefits)	Systematic review of randomized trials or *n*-of-1 trials	Randomized trial or observational study with dramatic effect	Non-randomized controlled cohort/follow-up study**	Case-series, case-control studies, or historically controlled studies**	Mechanism-based reasoning
What are the COMMON harms? (Treatment Harms)	Systematic review of randomized trials, systematic review of nested case-control studies, *n*-of-1 trial with the patient you are raising the question about, or observational study with dramatic effect	Individual randomized trial or (exceptionally) observational study with dramatic effect	Non-randomized controlled cohort/follow-up study (post-marketing surveillance) provided there are sufficient numbers to rule out a common harm. (For long-term harms the duration of follow-up must be sufficient.)**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning
What are the RARE harms? (Treatment Harms)	Systematic review of randomized trials or *n*-of-1 trial	Randomized trial or (exceptionally) observational study with dramatic effect			
Is this (early detection) test worthwhile? (Screening)	Systematic review of randomized trials	Randomized trial	Non -randomized controlled cohort/follow-up study**	Case-series, case-control, or historically controlled studies**	Mechanism-based reasoning

*** Level may be graded down on the basis of study quality, imprecision, indirectness (study PICO does not match questions PICO), because of inconsistency between studies, or because the absolute effect size is very small; Level may be graded up if there is a large or very large effect size.**

**** As always, a systematic review is generally better than an individual study.**

How to cite the Levels of Evidence Table

OCEBM Levels of Evidence Working Group*. "The Oxford 2011 Levels of Evidence".

Oxford Centre for Evidence-Based Medicine. http://www.cebm.net/index.aspx?o=5653

** OCEBM Table of Evidence Working Group = Jeremy Howick, Iain Chalmers (James Lind Library), Paul Glasziou, Trish Greenhalgh, Carl Heneghan, Alessandro Liberati, Ivan Moschetti, Bob Phillips, Hazel Thornton, Olive Goddard and Mary Hodgkinson*

Fallibility of statistics and philosophical qualms

Sample size, sample bias, confounding

RCTs: can they trump any other sort of trials?

- But randomisation often fails …

Are the merits of meta-analyses justified?

- But they may lead to inconsistent results

…

RCTs

Arguments developed so far
concentrate on
the top of the hierarchy.

We concentrate on the bottom part.

Mechanisms

The background

THE RUSSO-WILLIAMSON THESIS

The thesis

To establish a causal claim, one normally needs to establish two things:

that a cause makes a difference to the effect, and

that there is a mechanism from cause to effect

Russo and Williamson

Interpreting causality in the health sciences, ISPS 2007

Generic vs. single-case causality. The case of autopsy. EJPS 2011

Epistemic causality and evidence-based medicine. HPLS 2011

EnviroGenomarkers. The interplay between difference-making and mechanisms. MedSt 2012

Disambiguation

Mechanistic evidence vs

evidence *of* mechanisms

Difference-making evidence vs

evidence *of* difference-making

Evidence vs

evidence-*gathering* methods

Illari, P. Disambiguating the Russo-Williamson Thesis, ISPS, 2011

What mechanism?

What mechanism ought to support a causal claim?

Fully-known? Confirmed? Plausible?

Gillies D. The Russo-Williamson thesis and the question of whether smoking causes heart disease, in *Causality in the* Sciences, 2011

What ‘kind’ of mechanism?

Biological, biochemical, socio-economic, …?

F. Russo, Causal webs in epidemiology, *Paradigmi*, 2011

M. Kelly and F. Russo, The integration of social and biological mechanisms of disease causation, in preparation

Scope of RWT

A thesis

- About the epistemology of causality
  - What sources of evidence allows us to establish causal claims

- With methodological implications
  - What evidence-gathering methods to use to establish causal claims

Mechanism and evidence evaluation

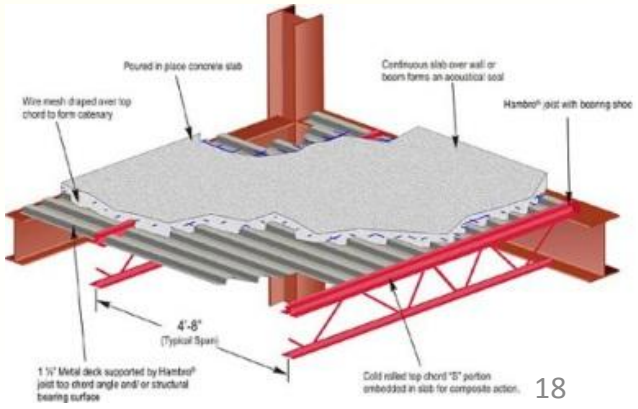
INTEGRATION OF EVIDENCE

The analogy of reinforced concrete

Evidence: integration, not substitution

- Difference making helps with masking
- Mechanisms helps with confounding
  - Integration helps solve more problems, and better
- Difference making and mechanisms help each other with their respective weaknesses

The more integrated, the merrier



Bradford Hill's guidelines

1. Strength of association
2. Temporality
3. Consistency
4. Theoretical plausibility
5. Coherence
6. Specificity in the causes
7. Dose response relationship
8. Experimental evidence
9. Analogy

Different aspects involved

Our observations reveal an association between two variables, perfectly clear-cut and beyond what we would care to attribute to the play of chance. What aspects of that association should we especially consider before deciding that the most likely interpretation of it is causation?

Hill (1965)

Not conditio sine qua non

Here then are nine different viewpoints from all of which we should study association before we cry causation. What I do not believe—and this has been suggested—is that we can usefully lay down some hard-and-fast rules of evidence that must be obeyed before we accept cause and effect. **None of my nine viewpoints can bring indisputable evidence for or against the cause and effect hypothesis and none can be required as a sine qua non.** What they can do, with greater or less strength, is to help us to make up our minds on the fundamental question—is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?

Hill (1965)

Mechanisms help

Design and interpretation of studies

Physiological knowledge is not only indispensible in explaining disease, but is also necessary to good clinical observation. For example, I have seen observers surprised into describing as accidents certain thermal phenomena which occasionally result from nerve lesions; if they had been physiologists, they would have known how to evaluate morbid symptoms which are really nothing but physiological phenomena.

Bernard 1856

External validity

External validity of treatments

- To whom the results apply?

The external validity of policy action

- Intervening on the same mechanism?
- Altering the causal structure?

From the population to the single case

The reference class problem

- Objective homogeneity
- Epistemic homogeneity

Personalised treatment?

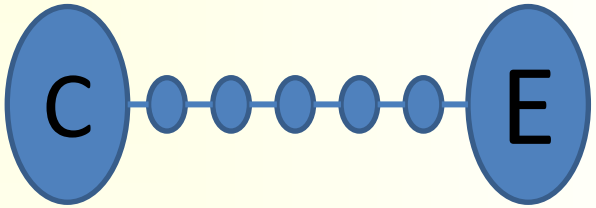
How to evaluate evidence of mechanisms?

Mechanisms and evidence evaluation

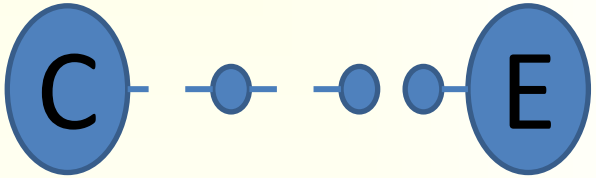
GUIDELINES FOR USING EVIDENCE OF MECHANISMS

Categories of evidence *of* mechanism

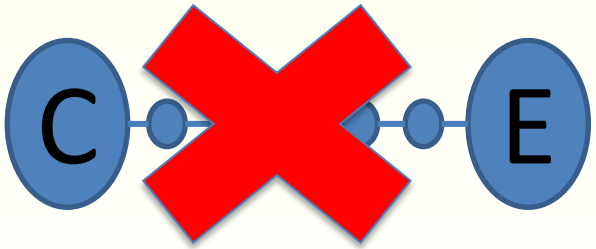
1. That there is a specific linking mechanism



2. That there is some kind of linking mechanism or other

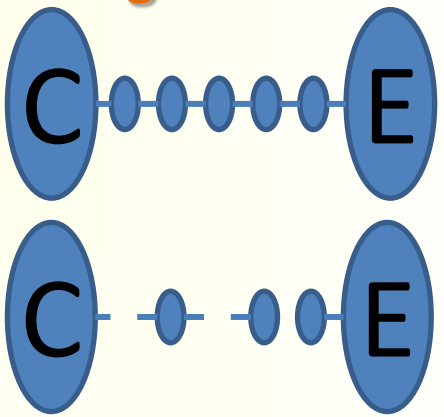


3. That there is no linking mechanism

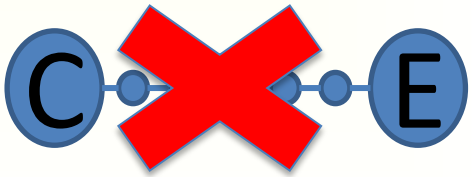


What evidence of mechanism is

1. Evidence of the existence and nature of the entities and activities of a linking mechanism, and their organization.
   - In vitro evidence
   - Animal experiments
   - Analogous mechanisms
   - Autopsy
   - Simulation
   - Even RCTs…



2. Evidence that suggests that a linking mechanism does not or could not exist.
   - Well established knowledge
     - Energy constraints on biochemical mechanisms
     - Comparative studies



Quality of evidence of mechanism

Desiderata for quality assessment:

We do not want to provide a rigid hierarchy or a ticklist, but something more fluid.

We want to move away from the idea that there is some baseline method which everything can be judged in relation to (as RCTs currently function).

We want to provide an assessment of mechanisms ultimately to be integrated with an assessment of the complementary evidence of difference-making.

Quality of evidence of mechanism

Pluses

Each independent method for detection of entity/interaction

Each independent research group confirming the result

More entities in the mechanism found

More links in the mechanism established

Analogous mechanisms known

Robust, reproducible in different conditions

Minuses

Single method used for detection of entity/interaction

Single research group confirming the result

Fewer entities in the mechanism found

Fewer links in the mechanism established

No analogous mechanisms known

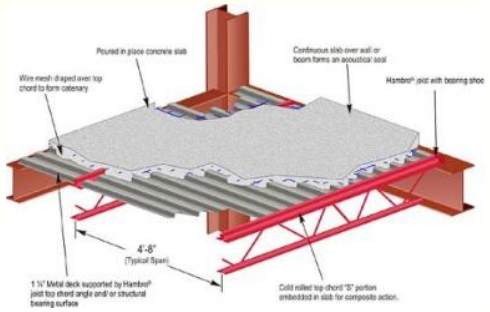
Fragile, not reproducible in slightly varying conditions

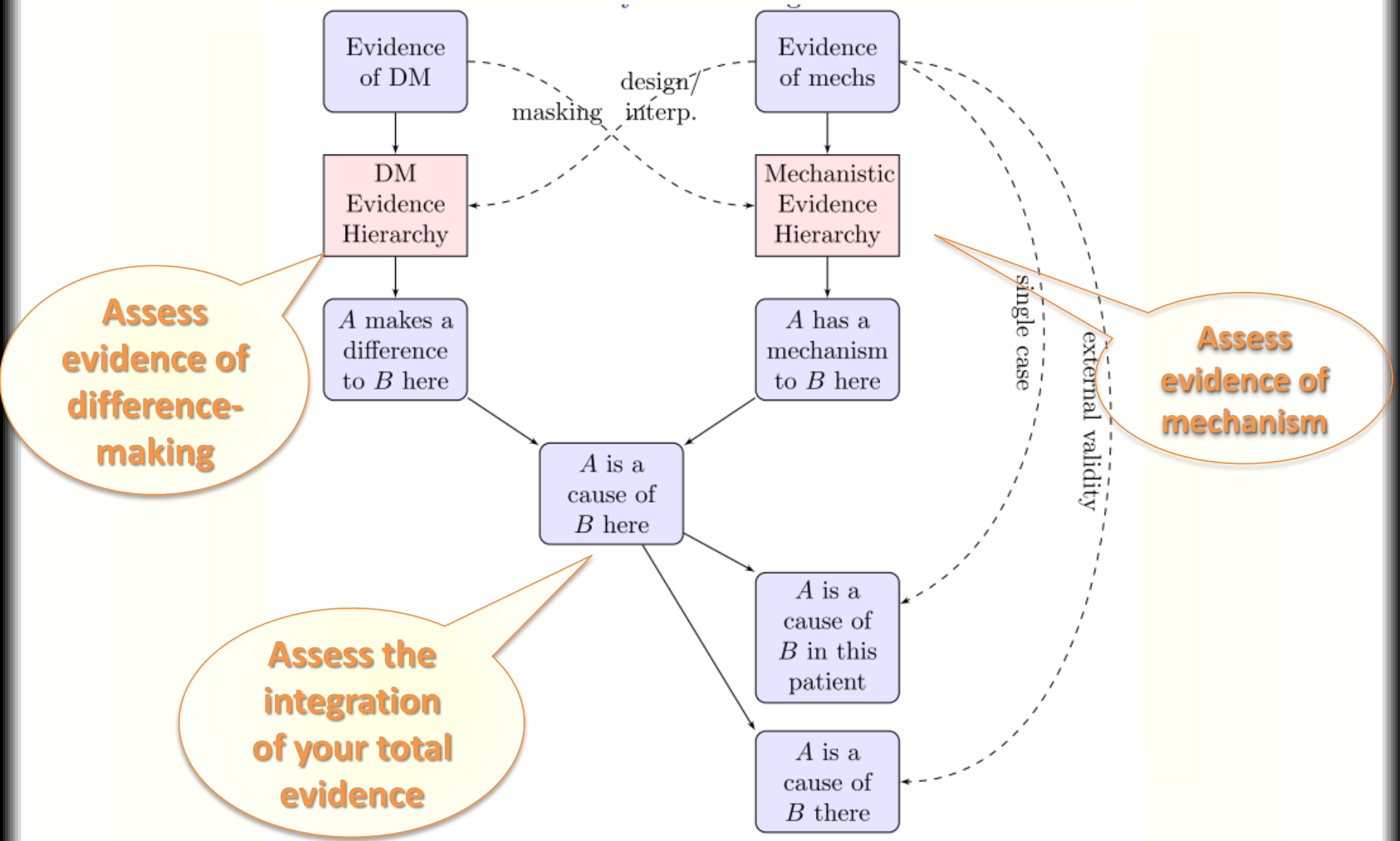
Integration in practice

Currently, observational studies can be upgraded to the level of an RCT, in principle.

In practice, they are not.

Our model allows the integration of observational with good mechanistic evidence.




Evidence of DM
Evidence of mechs
masking
design/ interp.
DM Evidence Hierarchy
Mechanistic Evidence Hierarchy
A makes a difference to B here
A has a mechanism to B here
A is a cause of B here
A is a cause of B in this patient
A is a cause of B there
single case
external validity
Assess evidence of difference-making
Assess evidence of mechanism
Assess the integration of your total evidence

TO SUM UP AND CONCLUDE

What we claim, what we don't

A thesis about what evidence is needed for causal assessment

A thesis about evaluating evidence

But it is not a rigid tick-list

The same item of evidence can be evidence of both difference making and of mechanisms

'Normally' does not imply no exceptions

Evidence of mechanism does not imply we know the mechanism in full detail

Mechanisms do not replace RCTs

Mechanisms are not infallible

Mechanisms are not 'stories'. We talk about evidence.

The more integrated, the merrier

Evidence of difference-making and of mechanisms

- Bernard, Hill, …
- Many cases in history of medicine

Regain generality of causal reasoning

- Seek ‘help’ from different available sources of evidence
- No gold standards, but best integrated practices

Bernard, C. (1856). An introduction to the study of experimental medicine. Macmillan, New York, 1927 edition.

Bradford Hill, A. (1965). The environment and disease: association or causa- tion? Proceedings of the Royal Society of Medicine, 58:295–300.

Broadbent, A. (2011). Inferring causation in epidemiology: mechanisms, black boxes, and contrasts. In Illari, P. M., Russo, F., and Williamson, J., editors, Causality in the sciences, pages 45–69. Oxford University Press, Oxford.

Campaner, R. (2011). Understanding mechanisms in the health sciences. The- oretical Medicine and Bioethics, 32:5–17.Campaner, R. and Galavotti, M. C. (2012). Evidence and the assessment of causal relations in the health sciences. International Studies in the Philosophy of Science, 26(1):27–45.

Clarke, B. (2011). Causality in medicine with particular reference to the viral causation of cancers. PhD thesis, Department of Science and Technology Studies, University College London, London.

Claveau, F. (2012). The Russo-Williamson theses in the social sciences: Causal inference drawing on two types of evidence. Studies in History and Philosophy of Biological and Biomedical Sciences, DOI 10.1016/j.shpsc.2012.05.004.

Darby, G. and Williamson, J. (2011). Imaging technology and the philosophy of causality. Philosophy & Technology, 24(2):115–136.

Gillies, D. A. (2011). The Russo-Williamson thesis and the question of whether smoking causes heart disease. In Illari, P. M., Russo, F., and Williamson, J., editors, Causality in the Sciences, pages 110–125. Oxford University Press, Oxford.

Howick, J. (2011). Exposing the vanities—and a qualified defence—of mechanis- tic evidence in clinical decision-making. Philosophy of Science, 78(5):926–940. Proceedings of the Biennial PSA 2010.

Illari, P. M. (2011). Mechanistic evidence: Disambiguating the Russo- Williamson thesis. International Studies in the Philosophy of Science, 25:139– 157.

Russo, F. and Williamson, J. (2007). Interpreting causality in the health sci- ences. International Studies in the Philosophy of Science, 21(2):157–170.Russo, F. and Williamson, J. (2011a). Epistemic causality and evidence-based medicine. History and Philosophy of the Life Sciences, 33(4):563–582.Russo, F. and Williamson, J. (2011b). Generic versus single-case causality: the case of autopsy. European Journal for Philosophy of Science, 1(1):47–69.Russo, F. and Williamson, J. (2012). EnviroGenomarkers: the interplay between mechanisms and difference making in establishing causal claims. Medicine Studies, 3(4):249–262.

Weber, E. (2009). How probabilistic causation can account for the use of mecha- nistic evidence. International Studies in the Philosophy of Science, 23(3):277– 295.