Causal inference

Lecture by **Neil Pearce**
Department of Medical Statistics
London School of Hygiene and Tropical medicine

modified by **Manolis Kogevinas**
Dept. of NCDs & Environment, ISGlobal, Barcelona

with some slides from **Jorn Olsen,** EEPE course, Florence

OMEGA-NET Course, Barcelona, 1 April 2022

What is a cause?

Types of causes

How do we assess causality?

(Old and current debates)

Causes in epidemiology

- We run a discipline, which aims at identifying causes of diseases
- Our discipline rests upon the axioms that diseases have causes. Some of these causes are human-made and can be partly or fully eliminated. Therefore, some diseases are preventable

Causes in epidemiology

- We have a concept of causation - without a concept we would have no language

- A problem is that the epidemiologic concept of causation is very different from a common sense concept of causation.

Causes in epidemiology

- Some smokers get lung cancer, but most do not. If they do, they get the disease 20 to 60 years after they started smoking

- When I press the switch, the light is on

- Can we think about a concept of causation that would fit both situations?

Causes in epidemiology

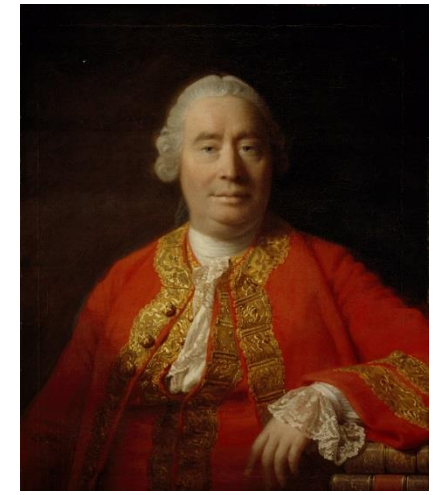
- What about if you said this? If I press the switch, sometimes the light turns on, sometimes it does not. If it turns on, it happens with 20 years delay. Who would take you seriously? We say this all the time!

- Do smokers take part in a different lottery than non-smokers – the probabilistic concept? Mackie and Rothman have developed a ”weak” Hume model of causation that fits both situations

Causes in epidemiology

- How can we explain delayed effects?
- Why do causes operate in a probabilistic manner?
- How can we explain that effects may have more than one cause?

What is a cause?



…we may define a cause to be *an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second*

[induction!]

[Hume. An Enquiry Concerning Human Understanding, 1748]

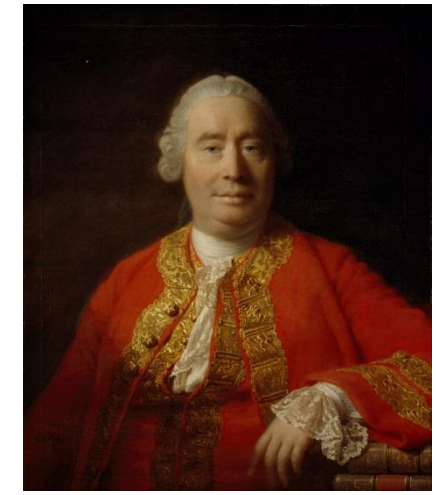
Not many examples of E  D associations in medicine.

Causes in the ”strong-Hume” sense could be seen as what Rothman describe as **necessary** and **sufficient causes**

Several examples in medicine, but they are usually cyclically defined by us:

- COVID-19
- tuberculosis
- painter’s syndrome
- AIDS etc.

What is a cause?



…we may define a cause to be *an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second*

[induction!]

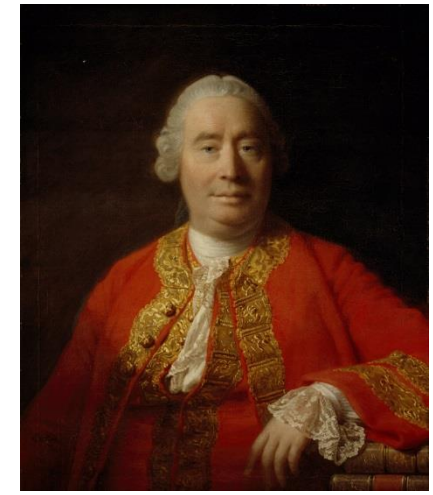
then he continues:

or in other words if the first object had not been, the second had never existed

[counterfactuals!]

[Hume. An Enquiry Concerning Human Understanding, 1748]

What is a cause?



We may define a cause to be an object, followed by another… where, if the first object had not been, the second had never existed [Hume, 1748]

An association may be classed as presumptively causal when it is believed that, had the cause been altered, the effect would have been changed [MacMahon and Pugh, 1967]

NB: these two definitions are different: the first definition assumes that the ‘exposure’ is present or not present (or present at different levels), but does not say how this occurs; the second definition assumes that the ‘exposure’ has somehow been altered

- The “strong-Hume” concept of causation does not reflect that several causes are needed to activate an effect.

- Even in fully deterministic systems with only one causal pathway, several causes are needed to activate an effect.

- Supply of fuel → power plant → transport of electricity, fully fitted electrical cords, a functioning light bulb → pressing the switch → the light is on!

Causality

What is a cause?

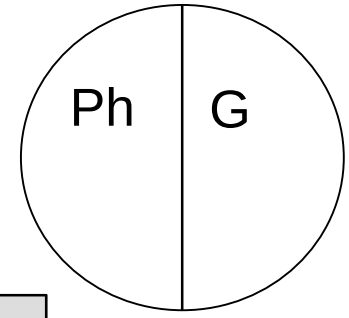
Types of causes

How do we assess causality?

Old debates

Current debates

Causes of Phenylketonuria (PKU)



		PKU gene: No	PKU gene: Yes
Phenylalanine in the diet	High	Zero risk	High (100%) risk
	Low	Zero risk	Zero risk

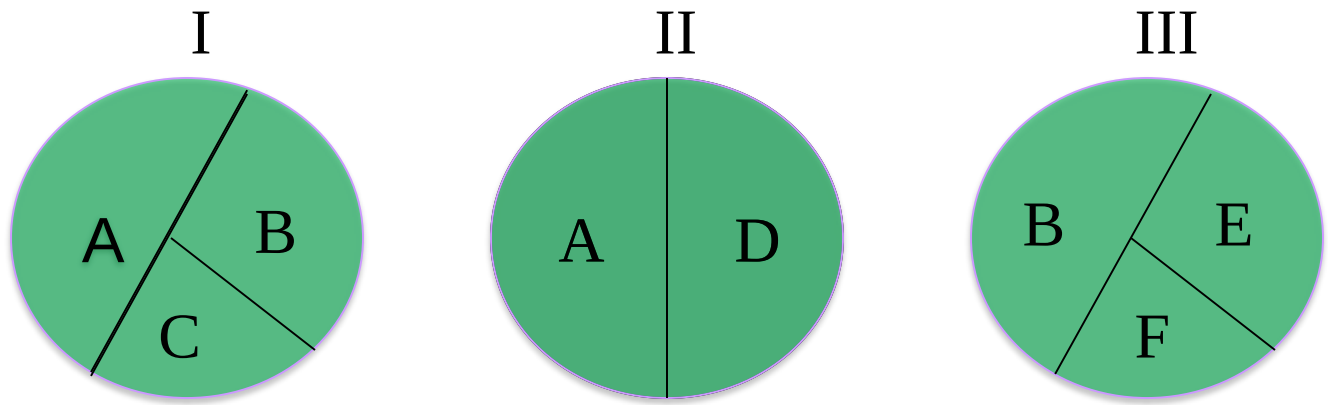
Phenylketonuria (PKU), is **a rare inherited disorder that causes phenylalanine (an amino acid) to build up in the body**. PKU is caused by a defect in the gene that helps create the enzyme needed to break down phenylalanine.
Without the enzyme necessary to process phenylalanine, a dangerous buildup can develop when a person with PKU eats foods that contain protein or eats aspartame, an artificial sweetener. This can eventually lead to serious health problems.

What is the difference between these two causes?

- One is a state (i.e. having the PKU gene)
- One is an action (i.e. eating food high in phenylalinine)

Causes act conditionally upon other causes and in combination they become a sufficient causal field.

Each component cause is only sufficient given the other component causes exist.

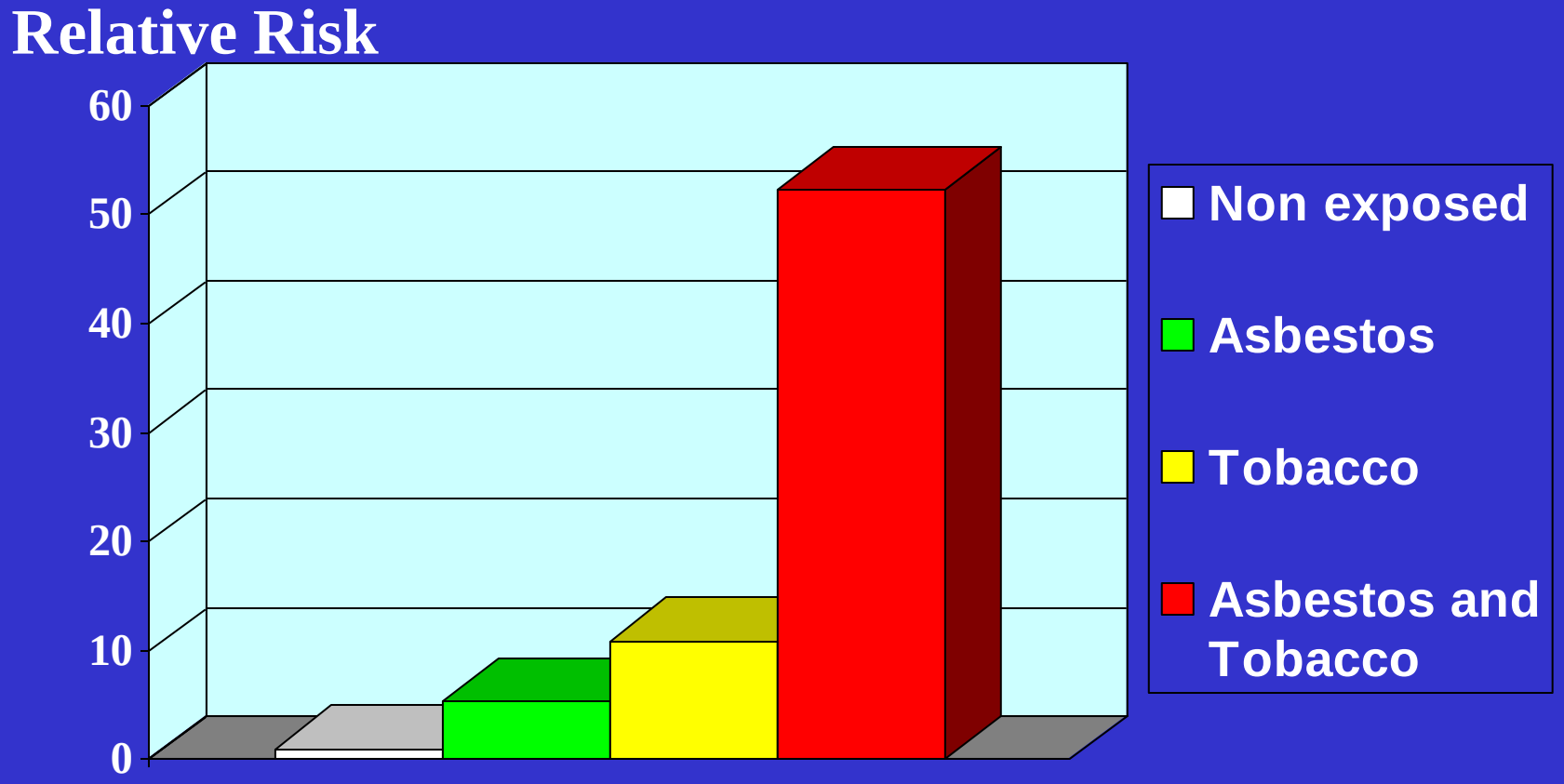


Causes are only **necessary** for a subdomain of diseases (I, II or III) and they are only **sufficient** given the other causes are in place.

The strength of a given association depends upon the prevalence of the other component causes.

- Component causes are insufficient but necessary within a causal field.
- Causal fields are un-necessary (if there are more) but sufficient

Interaction between exposure to asbestos fibres, tobacco smoke and ling cancer risk



(Selikoff et al 1979)

The causal field model **explains why attributable fractions and etiologic fractions sum up to more than 100%**

-
- Causal inference is limited in epidemiology to the comparisons we can make.
- Lung Cancer would look like a genetic disease if we all smoked 20 cigarettes a day
- Excess fractions or attributable fractions are not equal to etiologic fractions because they do not take into consideration cases that are accelerated by the exposure.

Causality

What is a cause?

Types of causes

How do we assess causality?

Old debates

Current debates

Bradford Hill's viewpoints

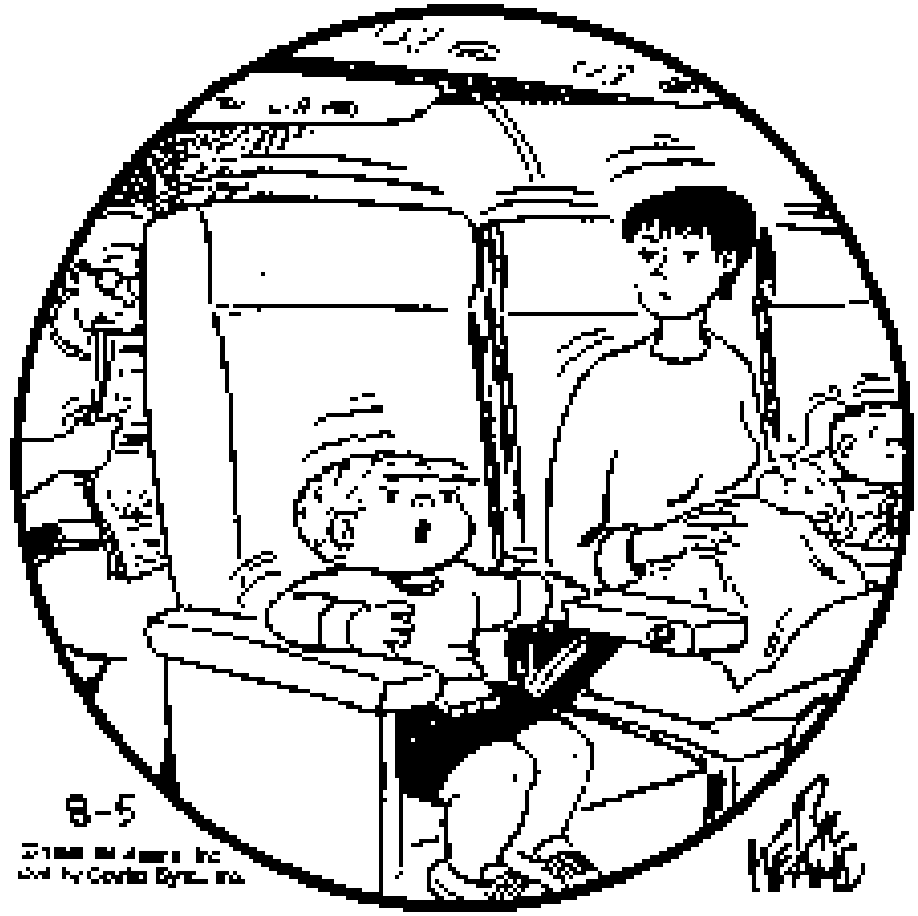
Nine "viewpoints" for assessing causality

1. Strength
2. Consistency
3. Specificity
4. Temporality
5. Biological gradient
6. Plausibility
7. Coherence
8. Experiment
9. Analogy

"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 can 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?"

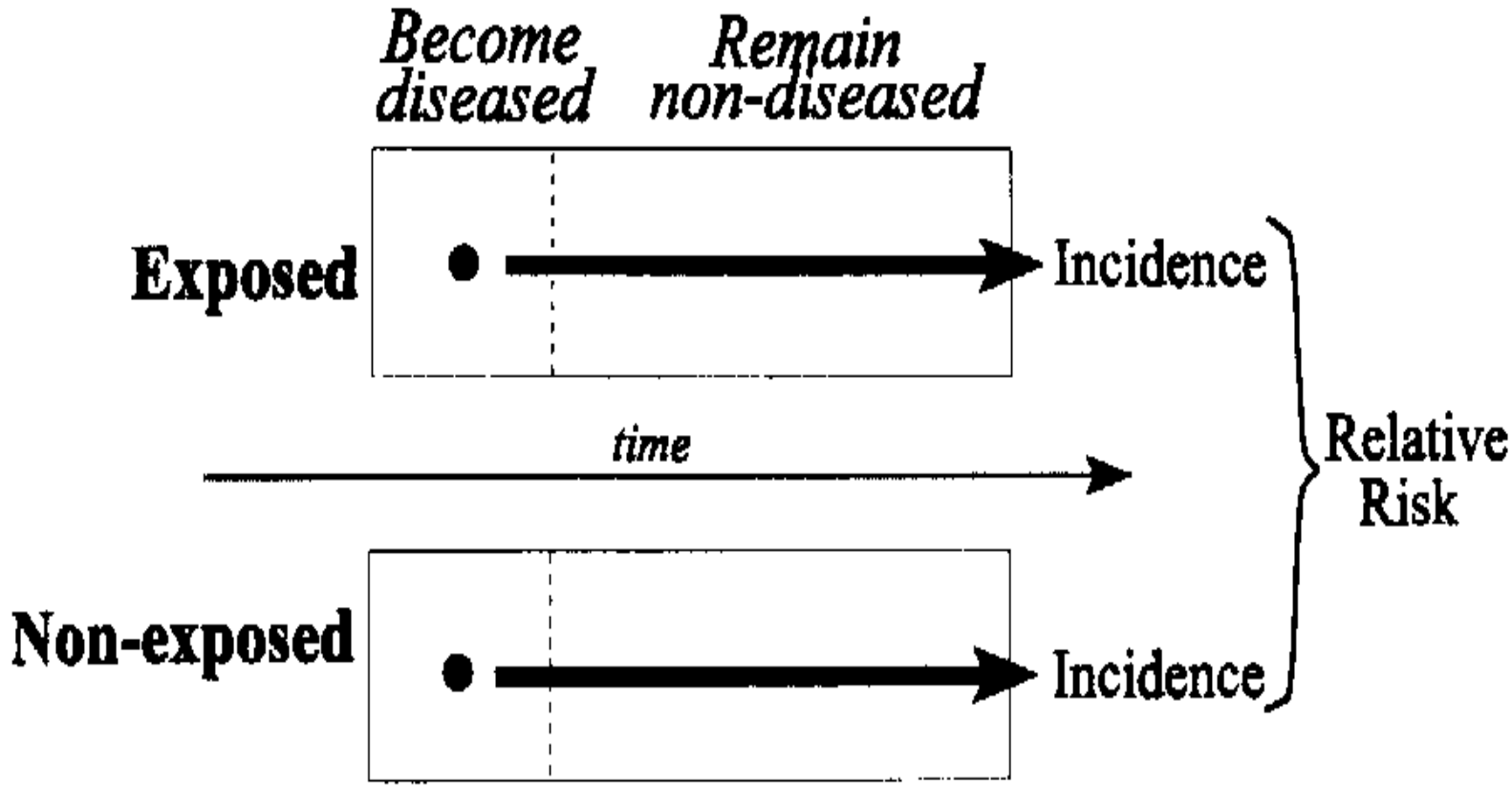
❖ These criteria, guidelines are useful but only guidelines-only temporarily is a required condition – a sine qua non requirement

THE FAMILY CIRCUS



"I wish they didn't turn on that seatbelt sign so much! Every time they do, it gets bumpy."

Cohort study design (but we evaluate temporality in all major designs)



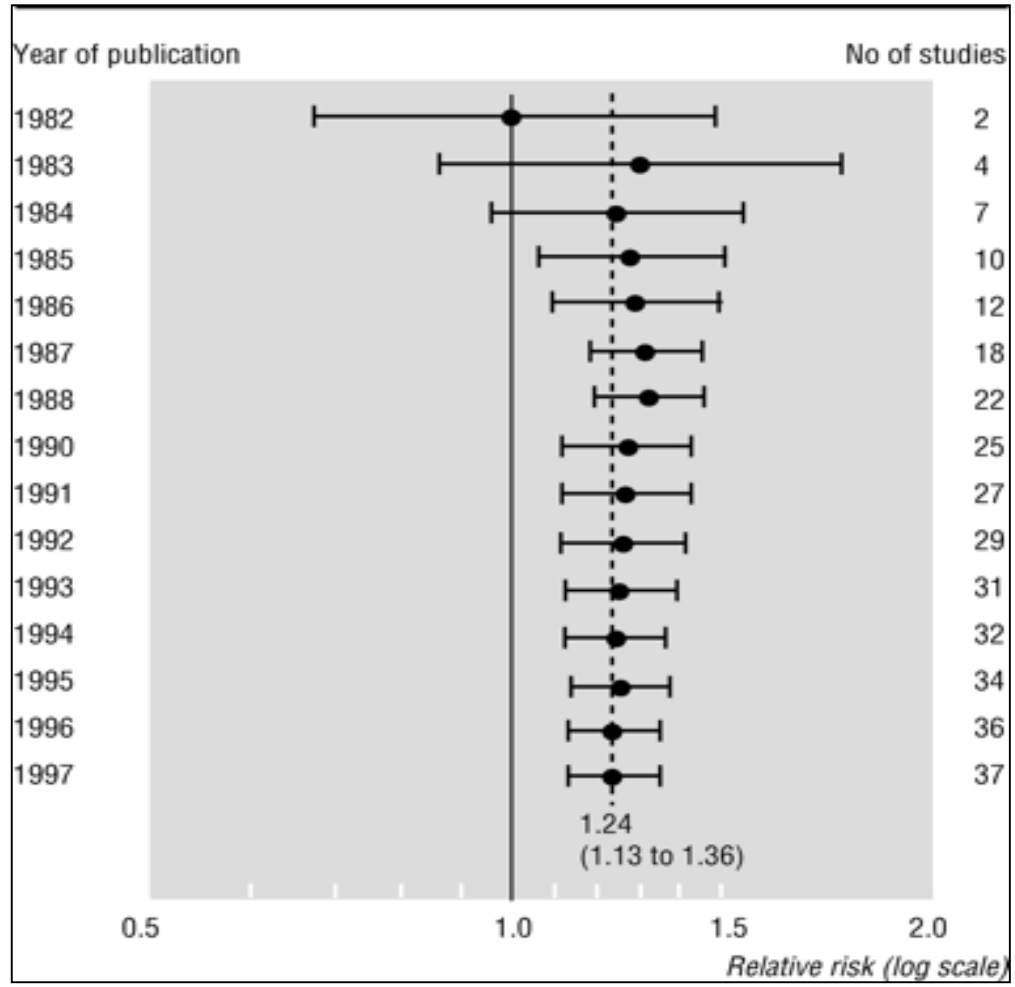
Strength

The relative risk for smoking (active) and lung cancer is around 10 (ie 1000% increased risk)

The relative risk of secondhand smoke and lung cancer (among non-smokers) is around 1.2 (ie 20% increased risk).

Do we trust one association more than the other?

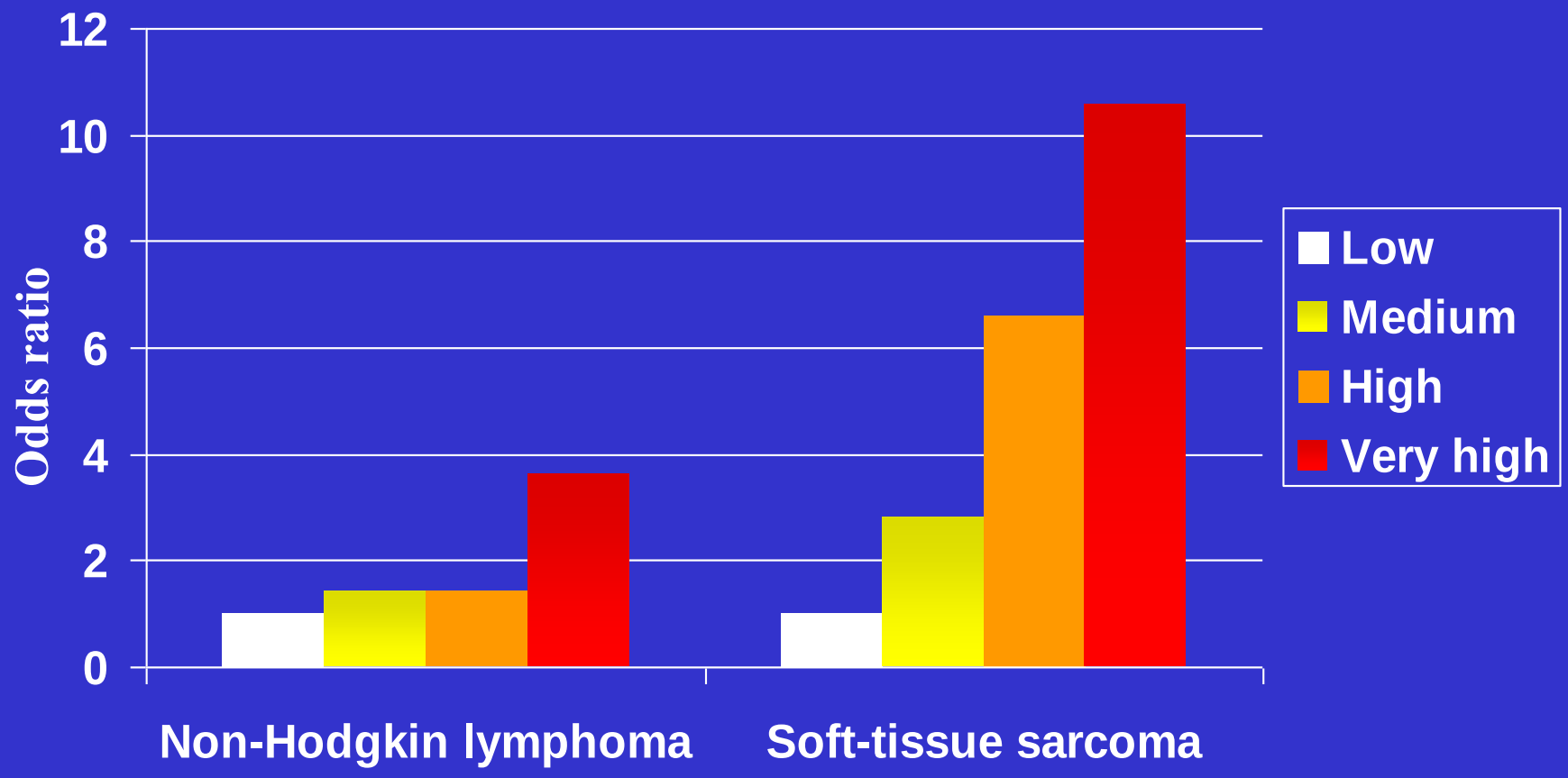
Does the fact that secondhand smoke causes a smaller increase in risk invalidate the causal nature of the association?



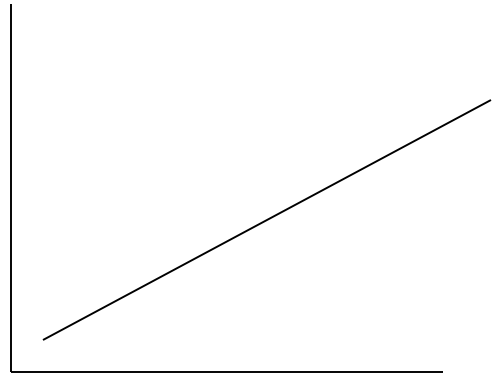
Relative risk of lung cancer among non-smokers who lived with smokers compared with others who did not live with smokers (Hackshaw BMJ 1997)

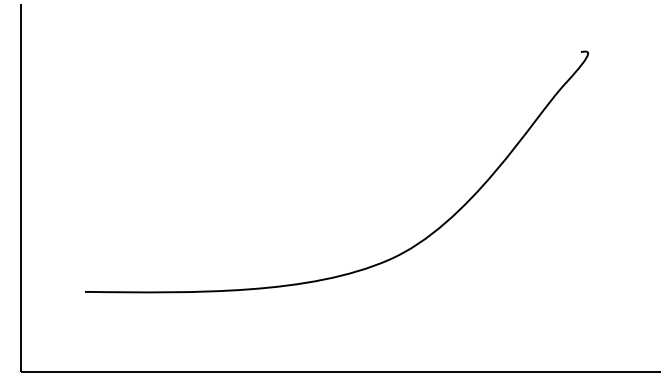
Biological gradient (dose-response)

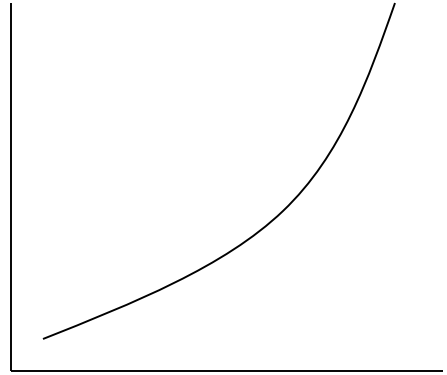
Odds Ratios for non-Hodgkin lymphoma and soft-tissue sarcoma. Nested case-control study within the IARC cohort study on workers exposed to dioxns (TCDD)

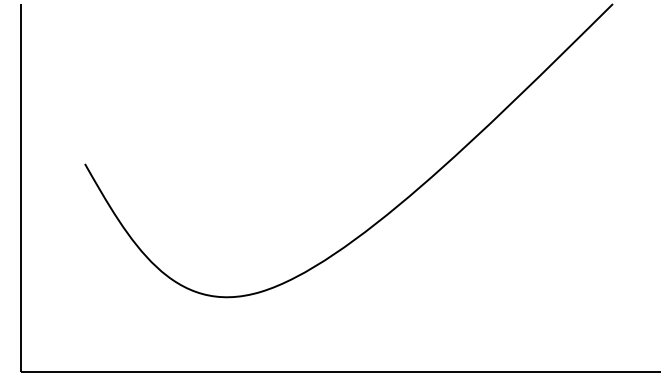


(Kogevinas et al 1995, Epidemiology)



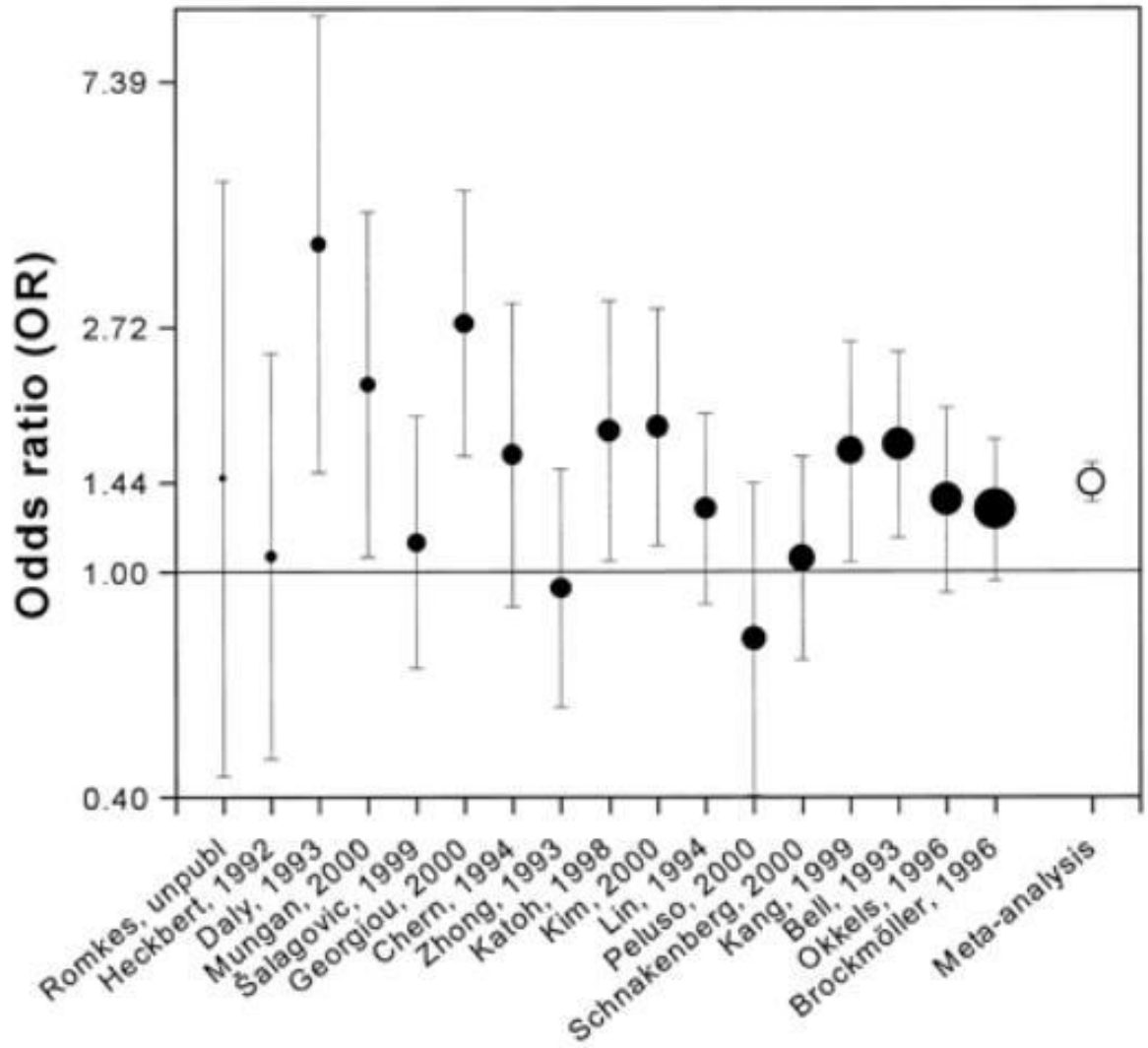






Consistency

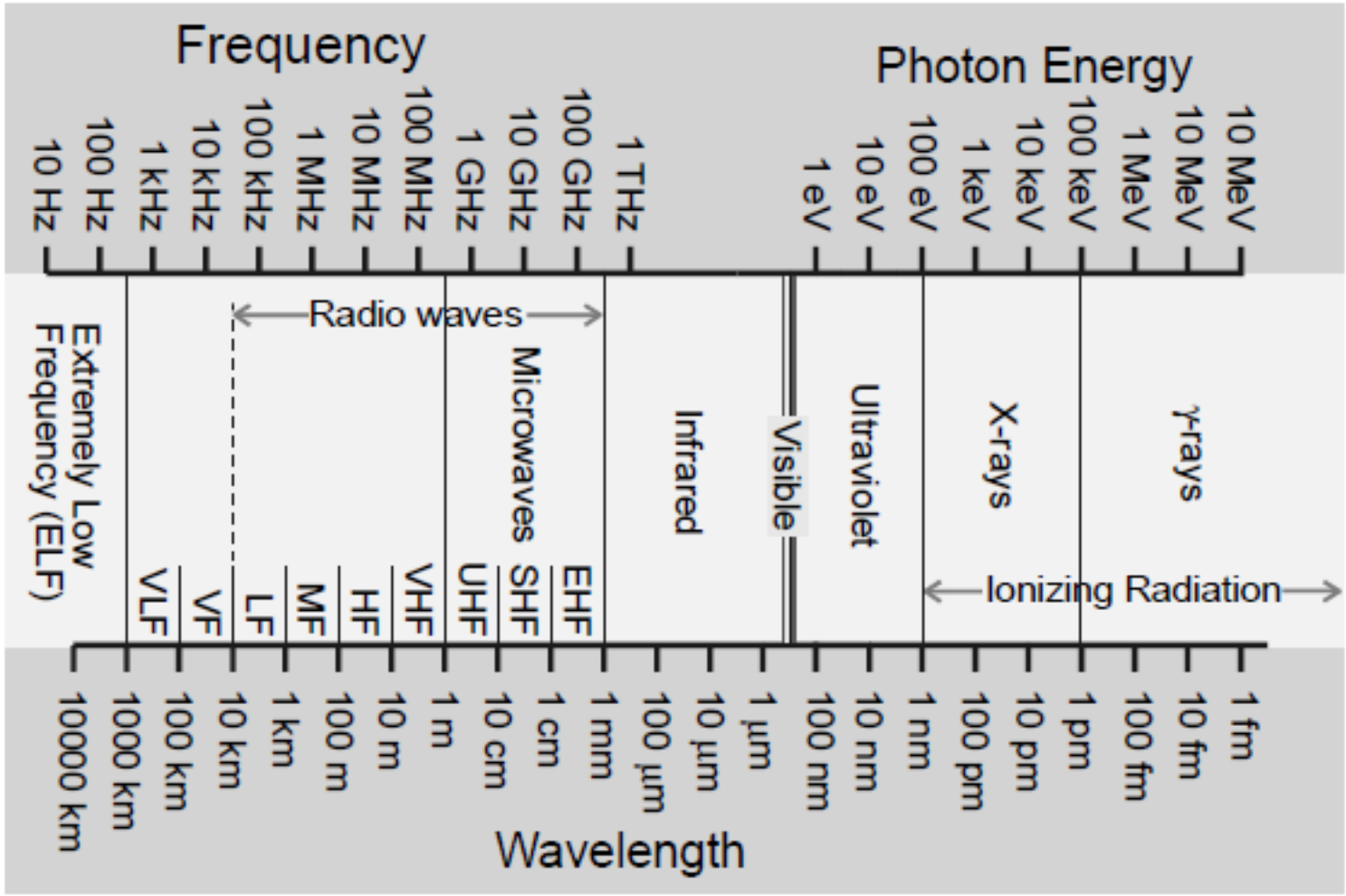
Odds ratio and 95%CI for deletion (null-null) of *GSTM1 and urinary bladder cancer*. Size of the circles represent the number of cases. Y-axis is logarithmic (Engel, AJE 2002)



Plausibility

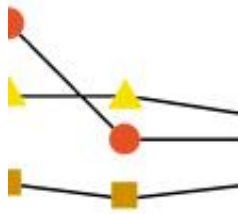
- the case of smoking and lung cancer
- the case of second-hand smoke and lung cancer
- the case of ELF EMF (extremely low frequency electromagnetic fields) and child leukaemia

Frequencies – electromagnetic fields



Pooled analysis of studies on child leukaemia and exposure to low frequency EMF (Ahlbom, Br J Cancer 2000)

Exposure	Relative Risk (95% CI)
Non-exposed (<0,1μT)	1.0
0.1-<0.2 μT	1.1 (0.9-1.3)
0.2-<0.4 μT	1.1 (0.8-1.5)
Above 0.4 μT	2.0 (1.3-3.1)



Two different approaches to investigating and assessing causality

- Do one ‘really good’ study and dismiss the other evidence
- Consider all of the evidence together
  - Pluralism
  - Triangulation
  - Understanding of mechanisms

How do we assess causation?

- By putting all of the evidence together
- There is rarely, if ever, a single study which establishes causality

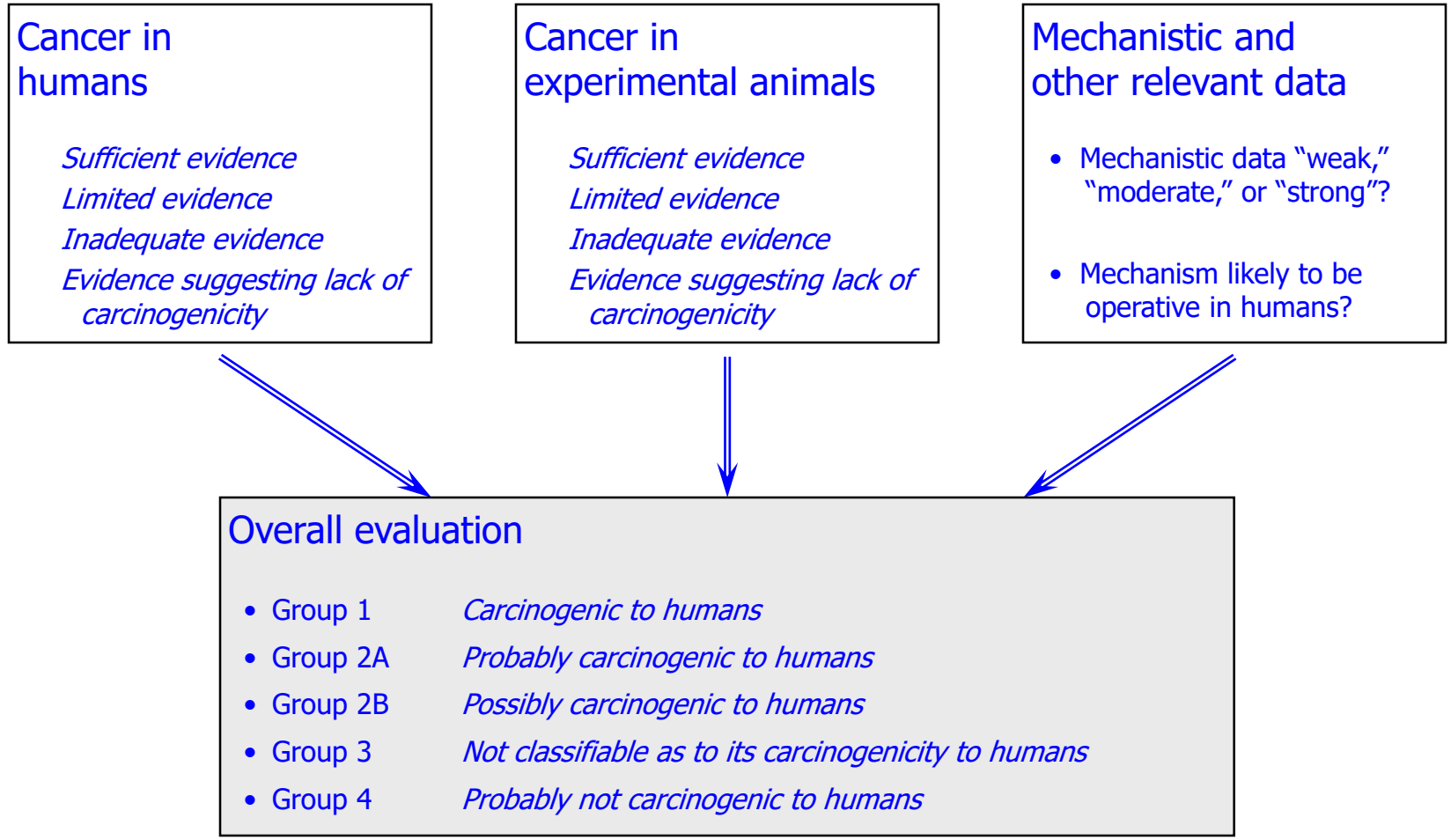
How do we assess causation?

- Even a randomized trial cannot establish causality on it's own, because it always relies on auxiliary assumptions/evidence (e.g. full/unbiased participation, lack of misclassification, lack of contamination of the comparison group, etc)
- This applies in all science – the Duhem/Quine thesis – a theory always relies on auxiliary hypotheses

Developments from Bradford-Hill

- Ruling out alternatives
  - E.g. the time trends in lung cancer in the 20th century rule out the possibility that there is a gene which both causes smoking and (independently) causes lung cancer
- Interlocking arguments from different areas of science
  - E.g. the IARC Monographs integrate information from animal, human and mechanistic studies
- Triangulation
  - E.g. negative controls, estimating effects in different populations with different confounding structures

Putting all the evidence together: the IARC Monographs (since 1971)



Causality

What is a cause?

Types of causes

How do we discover a cause?

How do we assess causality?

Old debates

Current debates

Observational epidemiology vs randomized controlled trials: back to the 1970s

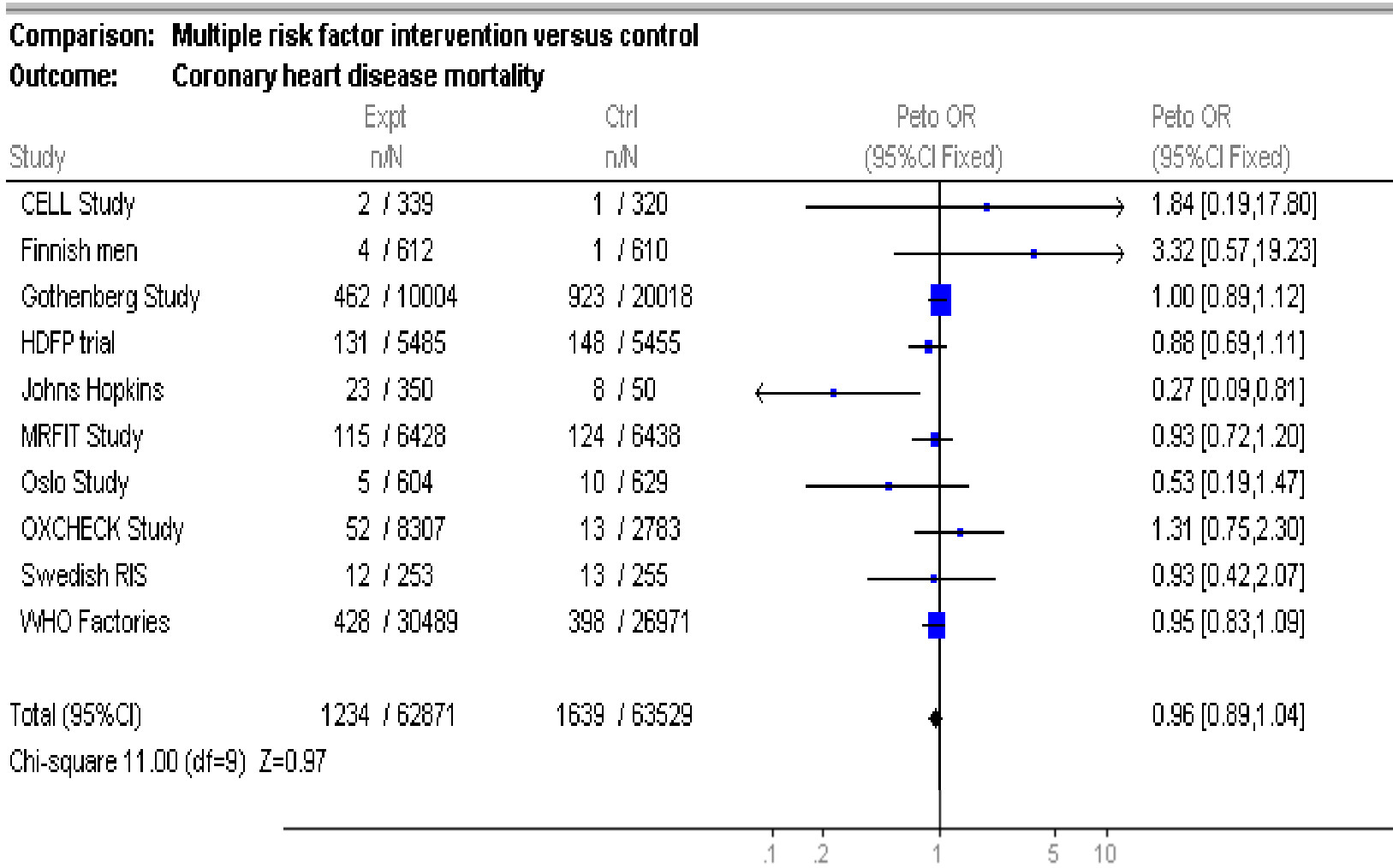
"To understand causal effects, epidemiologists should put more effort into organizing large-scale randomized trials instead of traditional observational studies, which are inevitably crippled by confounding and other biases."

"Where the phenomenon is large and complex, the challenge is to break it into parts that can be meaningfully tested via experiments of demonstrable internal and external validity."

Limitations of randomized trials

- Only certain questions can be asked
  - States cannot be studied directly (e.g. sex, obesity)
  - Many actions (e.g. climate change, smoking) are impossible or difficult to randomize
- Only simplistic questions get asked
  - Most interventions on poverty do not address structural inequalities
- The intervention doesn’t work
  - Many interventions are not applied properly and/or there is ‘spillover’ to the non-intervention group
- The intervention works but it is not clear why
  - Often it is not clear which component of the intervention accounts for its success

Systematic review of multiple risk factor interventions: effect on CHD mortality



Ebrahim and Davey Smith BMJ 1997 and Cochrane Library

The false hierarchy of study designs

Premier league	RCTs
Championship (but desperate for promotion!)	‘Causal inference’ Mendelian Randomization
Division 1	Cohort studies
Division 2	Case-control studies
Vauxhall Conference	Cross-sectional studies
Amateurs	Ecologic studies, case series

The false hierarchy of study designs

- None of these designs provides definitive evidence on its own; all depend on ‘auxiliary information’
- Causal inference always involves putting all of the evidence together
- ‘On the average’ a premier league team will defeat a ‘lower league’ team, but there are important exceptions
- Many questions cannot be tested in RCTs (or in observational studies that look like RCTs)
- Other study designs can provide key information

Causality

What is a cause?

Types of causes

How do we discover a cause?

How do we assess causality?

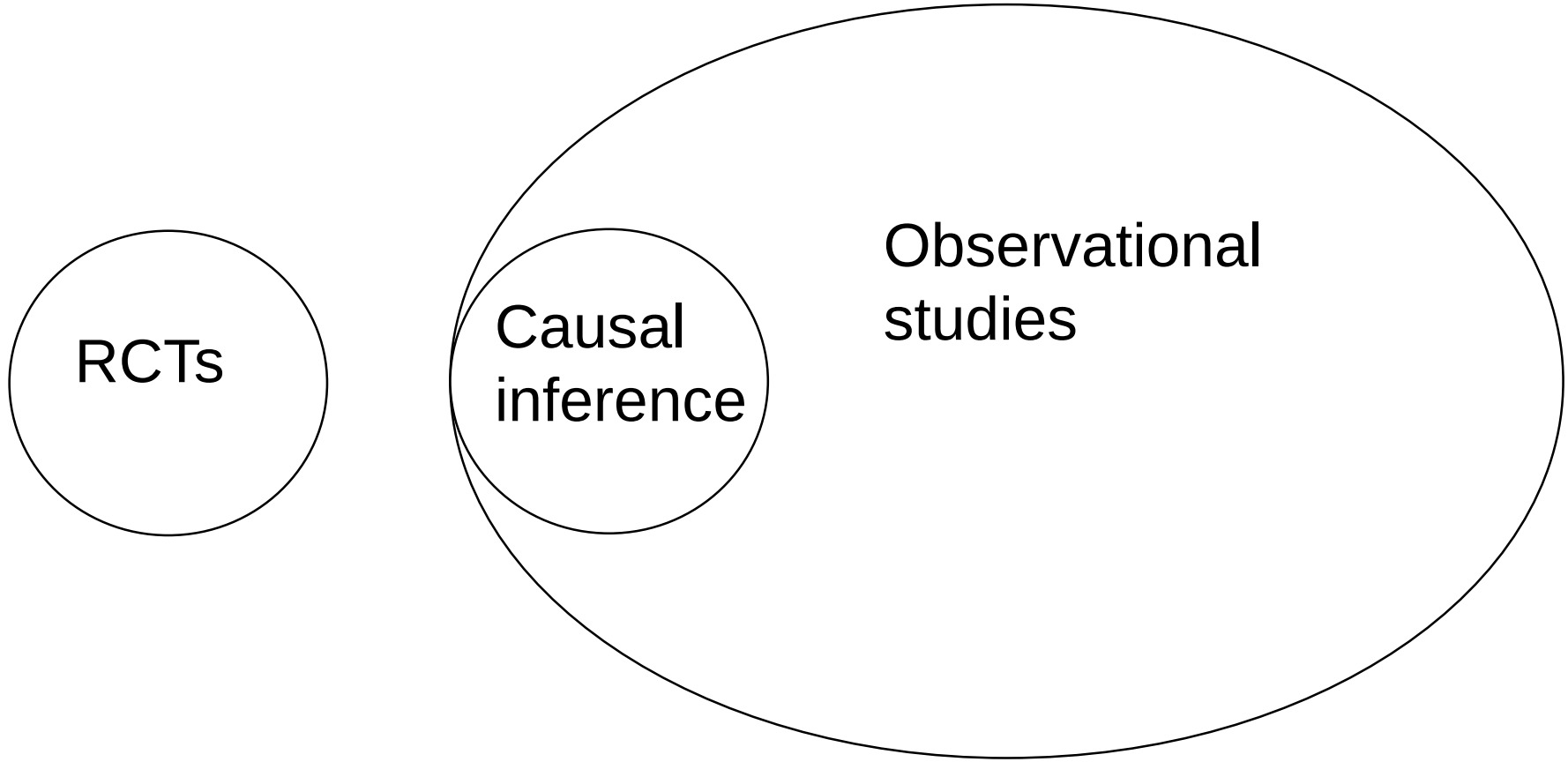
Old debates

Current debates

The debate is back!

Old version: RCTs versus observational studies

New version: ‘Causal inference’ versus other types of observational studies



The revenge of the algorithms

A risk of bias instrument for non-randomized studies of exposures: A users' guide to its application in the context of GRADE

Rebecca L. Morgan[a], Kristina A. Thayer[b], Nancy Santesso[a], Alison C. Holloway[c], Robyn Blain[d], Sorina E. Eftim[d], Alexandra E. Goldstone[d], Pam Ross[d], Mohammed Ansari[e], Elie A Akl[a,f], Tommaso Filippini[g], Anna Hansell[h,i,j], Joerg J. Meerpohl[k], Reem A. Mustafa[a,l], Jos Verbeek[m], Marco Vinceti[g,n], Paul Whaley[o], Holger J. Schünemann[a,p,*], GRADE Working Group

[a] Department of Health Research Methods, Evidence, and Impact, McMaster University, Health Sciences Centre, Room 2C14, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada
[b] Integrated Risk Information System (IRIS) Division, National Center for Environmental Assessment (NCEA), Office of Research and Development, US Environmental Protection Agency, Building B (Room 211i), Research Triangle Park, NC 27711, USA
[c] Department of Obstetrics and Gynecology, McMaster University, Health Sciences Centre, Room 3N52A, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada
[d] ICF International Inc., 9300 Lee Highway, Fairfax, VA, USA
[e] School of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, ON K1H 8M5, Canada
[f] Department of Internal Medicine, Faculty of Health Sciences, American University of Beirut, P.O. Box: 11-0236, Riad-El-Solh Beirut 1107 2020, Lebanon
[g] Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Italy
[h] MRC-PHE Centre for Environment and Health, Imperial College London, St Mary's Campus, Praed St, Paddington, London W2 1PG, UK
[i] Public Health Directorate, Imperial College Healthcare NHS Trust, St Mary's Hospital, Paddington, London, W2 1PG, UK
[j] Centre for Environmental Health and Sustainability, University of Leicester, George Davies Building, University Road, Leicester LE1 7RH, UK
[k] Institute for Evidence in Medicine (for Cochrane Germany Foundation), Medical Center - University of Freiburg, Breisacher Strasse 153, 79110 Freiburg, Ger
[l] Division of Nephrology and Hypertension, Department of Medicine, University of Kansas Medical Center, 3901 Rainbow Boulevard, Kansas City, KS 66160,
[n] Finnish Institute of Occupational Health, Cochrane Work, Neulaniementie 4, 70701 Kuopio, Finland
[o] Department of Epidemiology, Boston University School of Public Health, Boston, MA, USA
[p] Lancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK
[*] Department of Medicine, McMaster University, Health Sciences Centre, Room 2C14, 1280 Main Street West, Hamilton, ON L8S 4K1, Canada

ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions

Jonathan AC Sterne,[1] Miguel A Hernán,[2] Barnaby C Reeves,[3] Jelena Savović,[1,4] Nancy D Berkman,[5] Meera Viswanathan,[6] David Henry,[7] Douglas G Altman,[8] Mohammed T Ansari,[9] Isabelle Boutron,[10] James R Carpenter,[11] An-Wen Chan,[12] Rachel Churchill,[13] Jonathan J Deeks,[14] Asbjørn Hróbjartsson,[15] Jamie Kirkham,[16] Peter Jüni,[17] Yoon K Loke,[18] Theresa D Pigott,[19] Craig R Ramsay,[20] Deborah Regidor,[21] Hannah R Rothstein,[22] Lakhbir Sandhu,[23] Pasqualina L Santaguida,[24] Holger J Schünemann,[25] Beverly Shea,[26] Ian Shrier,[27] Peter Tugwell,[28] Lucy Turner,[29] Jeffrey C Valentine,[30] Hugh Waddington,[31] Elizabeth Waters,[32] George A Wells,[33] Penny F Whiting,[34] Julian PT Higgins[35]

Non-randomised studies of the effects of interventions are critical to many areas of healthcare evaluation, but their results may be biased. It is therefore important to understand and appraise their strengths and weaknesses. We developed ROBINS-I

such as cohort studies and case-control studies in which intervention groups are allocated during the course of usual treatment decisions, and quasi-randomised studies in which the method of allocation falls short of full randomisation. Non-randomised studies can provide evidence additional to that available from randomised trials about long term outcomes, rare events, adverse effects and populations that are typical of real world practice.[1,2] The availabil-

What does it mean in practice?

Good aspects

- More explicit formulation of the counterfactual contrast
- Designing epidemiological studies to (where possible and appropriate) mimic as closely as possible what the results that have been obtained from the corresponding RCT
- Better understanding of epidemiological theory and practice for studies which involve causes of this type (i.e. actions)

Regrettable aspects

- Restriction of epidemiology to studying causes which fit this paradigm (i.e. actions), and rejection of studies of states
- Scoring of studies according to the RCT paradigm, and rejection of studies with low scores

Why is it a problem?

- In many ways this is a re-run of the old debate about RCTs versus observational studies (‘old wine in new bottles’)
- We used to be told that we could only establish causality with RCTs; now we are being told that we can only establish causality with RCTs, or with observational studies which closely mimic RCTs
- Exclusion and/or neglect of other causes which are states, even though many of these are important clinically (e.g. dynamic states such as obesity, hypercholesterolemia, high blood pressure) or socioeconomically (e.g. ‘fixed’ states such as sex)
- Suggests that causality can be established with a single ‘perfect’ study, and that we should always strive to do such a study rather than taking a more comprehensive (triangulation) approach
- Neglect or elimination of many of the things that make epidemiology unique and important (the population perspective)

Uproar after research claims red meat poses no health risk

One expert says findings by international experts represent 'egregious abuse of evidence'

▲ Critics of the study say many of the participants were young and unlikely to succumb to illness during the trial period. Photograph: Getty

New research that claims red and processed meat is probably not harmful to our health has caused controversy among experts who maintain people should cut down.

The World Health Organization has classified red and processed meats as

blic health bodies worldwide urge people to limit their

The New York Times

Opinion

Don't Let a Killer Pollutant Loose

The Trump administration is moving to ease standards on a particularly deadly air contaminant.

By John Balmes

Dr. Balmes is a medical professor and member of the California Air Resources Board.

April 14, 2019

Kim Ryu

PM 2.5 kills people. There has been little dispute that microscopic particulate matter in air pollution penetrates into the deepest parts of the lungs and contributes to the early deaths each year of

"Rather than relying on the weight-of-evidence approach that the EPA has traditionally used to infer causation, [the Clean Air Science Advisory Committee] wants to rely on studies that use… 'manipulative causality'. This theory restricts epidemiologic evidence that may be considered acceptable to assess causality to results from intervention studies or studies that have been analysed with the use of causal inference statistical methods". [NEJM 2019; 381:8]

The argument of last resort: "There is no alternative"

- The alternative (actually the mainstream) is what we have been doing for the last 200 years (e.g. Bradford-Hill)
- These traditional methods can be improved (but not replaced) by more recent developments (e.g. triangulation, Mendelian Randomization, causal inference methods)
- We did manage to establish causality (e.g. smoking and lung cancer) prior to modern 'causal inference' methods, and without requiring RCTs

The argument of last resort: “There is no alternative”

- Causal inference in epidemiology

  ‘Causal inference’ began in 1986; ‘there is no alternative’ to the methods that have been developed since then

- Sex in England:

  Sexual intercourse began
  In nineteen sixty-three
  (which was rather late for me) -
  Between the end of the "Chatterley" ban
  And the Beatles' first LP

  [Philip Larkin]

What should we do?

- Risk of bias tools can be useful in providing a ‘checklist’ of possible biases in individual studies, or groups of similar studies
- They only provide part of the information needed for evidence synthesis
- A ‘low scoring’ study may provide crucial evidence in the context of triangulation and/or systematic reviews
- Risk of bias assessment cannot be done in the abstract, but requires knowledge of the context, and the broader evidence that is to be synthesized
- Risk of bias assessments should not be used to reject ‘low scoring’ studies which may still provide useful information for evidence synthesis

Solutions 1: triangulation

International Journal of Epidemiology, 2017, 1–21
doi: 10.1093/ije/dyw314
Original Article

Original Article

Triangulation in aetiological epidemiology

Debbie A. Lawlor,[1,2,*] Kate Tilling[1,2] and George Davey Smith[1,2]

[1]MRC Integrative Epidemiology Unit at the University of Bristol, Bristol, UK and [2]School of Social and Community Medicine, University of Bristol, Bristol, UK

*Corresponding author. MRC IEU, University of Bristol, Oakfield House, Oakfield Grove, Bristol BS8 2BN, UK. E-mail: d.a.lawlor@bristol.ac.uk

Accepted 3 October 2016

Abstract

Triangulation is the practice of obtaining more reliable answers to research questions through integrating results from several different approaches, where each approach has different key sources of potential bias that are unrelated to each other. With respect to

Triangulation is the practice of obtaining more reliable answers to research questions through integrating results from several different approaches, where each approach has different key sources of potential bias that are unrelated to each other… We emphasize the importance of being explicit about the expected direction of bias within each approach, whenever this is possible, and seeking to identify approaches that would be expected to bias the true causal effect in different directions.

Solutions 2: assess specific sources of bias – don't use algorithms

American Journal of Epidemiology

Vol. 188, No. 9
DOI: 10.1093/aje/kwz131
Advance Access publication: May 30, 2019

Special Article

The Problem With Mechanistic Risk of Bias Assessments in Evidence Synthesis of Observational Studies and a Practical Alternative: Assessing the Impact of Specific Sources of Potential Bias

David A. Savitz*, Gregory A. Wellenius, and Thomas A. Trikalinos

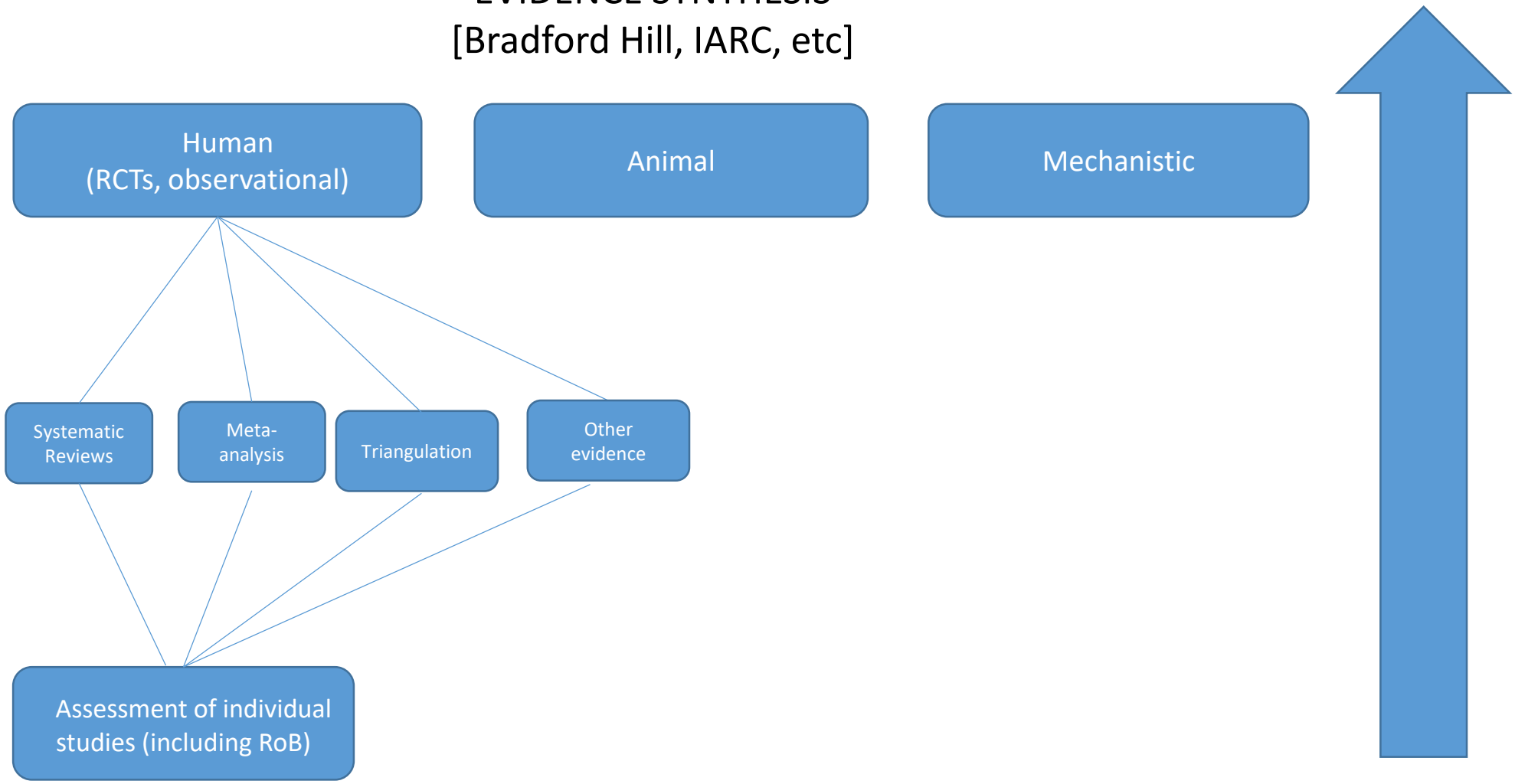
* Correspondence to David A. Savitz, Department of Epidemiology, Brown University School of Public Health, 121 South Main Street, Box G-S121, Providence, RI 02912 (e-mail: david_savitz@brown.edu).

Initially submitted August 4, 2018; accepted for publication May 17, 2019.

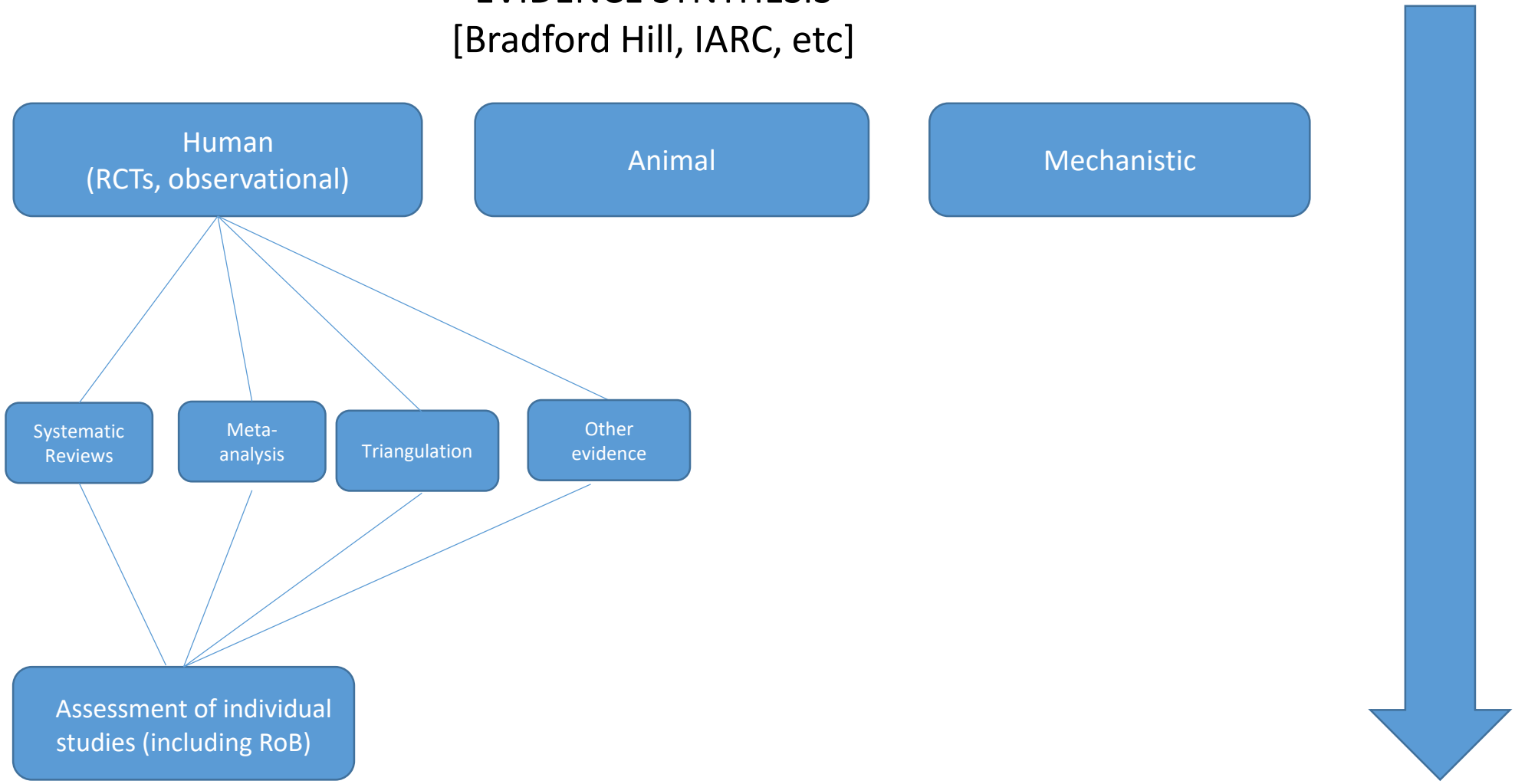
The trustworthiness of individual studies is routinely characterized in systemic reviews by evaluating risk of bias, often by mechanistically applying standardized algorithms. However, such instruments prioritize the repeatability of the process over a more thoughtful and informative but necessarily somewhat more subjective approach. In mechanistic risk of bias assessments, the focus is on determining whether specific biases are present, but these assessments do not

Mechanistic risk of bias assessments focus on assessing whether specific biases are present but fail to provide insights into the direction, magnitude, and relative importance of individual biases. Instead, risk of bias assessments should focus on identifying a small number of the most likely influential sources of bias… classifying each specific study based on how effectively it has addressed each potential bias, and determining whether results differ across studies in relation to susceptibility to each hypothesized source of bias.

EVIDENCE SYNTHESIS
[Bradford Hill, IARC, etc]



EVIDENCE SYNTHESIS
[Bradford Hill, IARC, etc]



What should we do/teach instead?

- Epidemiology is the science of the distribution and determinants of disease in human populations
- Discovering a cause usually involves a complex process with a wide variety of evidence
- Causal inference also involves a wide variety of evidence
- There is no ‘best’ study design, and rarely a single definitive study
- Traditional methods can be improved or supplemented, but not replaced, by more recent methodological developments

Current controversies in causal inference

Neil Pearce
Department of Medical Statistics
London School of Hygiene and
Tropical medicine