Causal and Associational Language in Observational Health Research: A Systematic Evaluation

Jose A. Calvache MD MSc PhD   (@jacalvache)

On behalf of authors team (48 members)

Department of Anesthesiology, Universidad del Cauca, Colombia

Department of Anesthesiology, Erasmus MC Rotterdam, The Netherlands

Global Evidence Summit

Using evidence. Improving lives.

Declaration of Conflict of interest

I have no actual or potential conflict of interest in relation to this presentation and I am acting of behalf of the entire authors team.

All errors are the sole responsibility of the authors of the following published article, and no funders had any role in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. The researchers were independent from funders, and all authors, external and internal, had full access to all of the data (including statistical reports and tables) in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis.

Global Evidence Summit

Background

- Health sciences research often explores the association between exposure and outcome, **aiming to infer causal effects** even with nonrandom assignment.
  - Causal inference is a core task of epidemiology (explanation)
  - Researchers and editors resort to "euphemisms" or "partial use" of causal language to maintain clarity and plausibility (**or to avoid complexities**)
  - Users often misunderstand the used language (i.e. media, news, among others)
  - "Just say association" like some journals rules

- Ambiguity in defining causal language without established standards is leading to
  - Potential disconnects between the authors' intentions, aims, methods, conclusions, and the perceptions of the results by research consumers and decision-makers
  - Over or undervalue results of research with potential misuse of evidence
  - Poor communication of research results

Background

"We found that **X** may be **related with Y**. This evidence suggest **we should reduce consuming X**"

- Exposure: **X**

- Outcome: **Y**

- Linking phrase: *related with* ➝ **Causal linking words / phrases**
  - ✓ Implies X influences Y, and/or
  - ✓ Levels or changes in Y, if any, would be attributable to X

- Modifying phrase: *may be*

- Action implication: "*We should reduce consuming X*" ➝ **Causal action implication**
  - ✓ A claim made about how to use the results of the research that requires that causality had been inferred

Global Evidence Summit

Objectives

We examined the linking language used in studies with a main exposure and outcome in the high-profile medical and epidemiologic literature.

- Our objectives were to:
  1. **identify the linking words** and phrases used to describe relationships between exposures and outcomes,
  2. generate estimates of the strength of causality stated or implied by the linking phrases and sentences using a guided subjective assessment process,
  3. examine the prevalence of action recommendations **that would require causal inference** to have been made, and
  4. examine disconnects between causal implications in linking sentences and action implications.

Methods

Selection of studies

- Searched and screened 1.170 articles

- Medical/public health/epidemiology literature from 2010 to 2019

- Three reviewers rated the degree of causality implied in abstracts and full text
  - Relevant training and standardization
  - Huge collaboration and transparency

- Assessment
  - linking language and
  - action recommendations
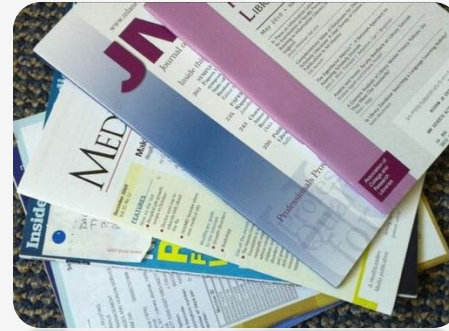
18 high profile journals

Primary research articles examine the causal and/or noncausal association between one main exposure concept and one main outcome concept.
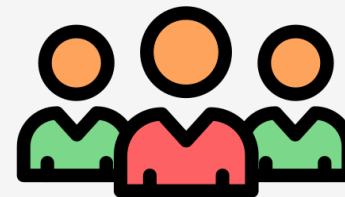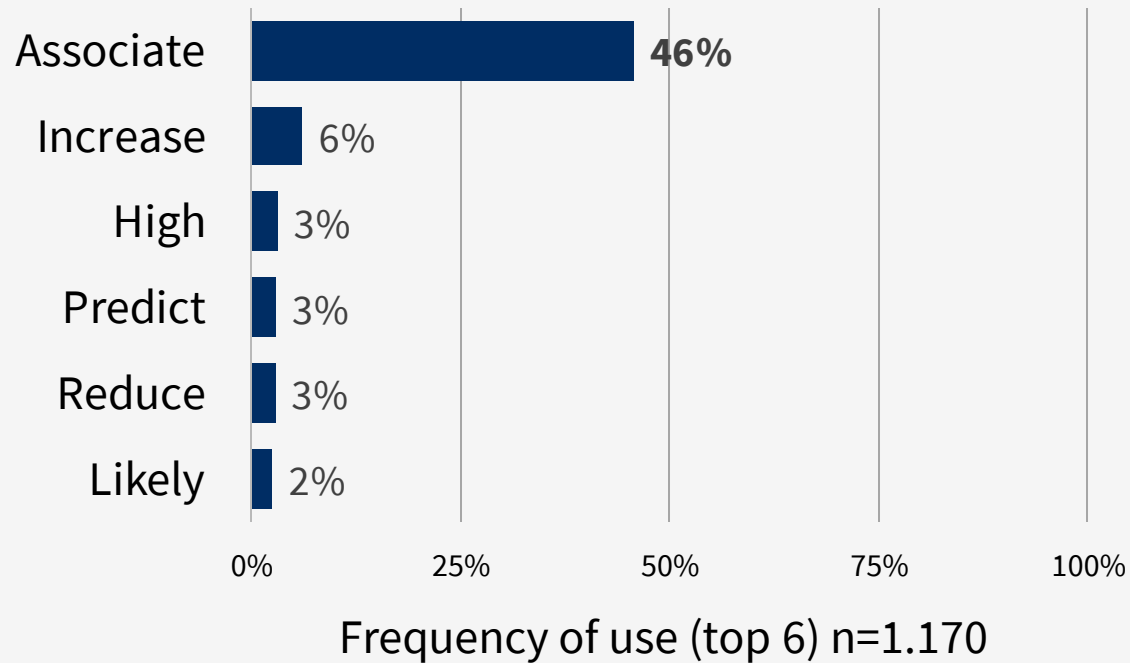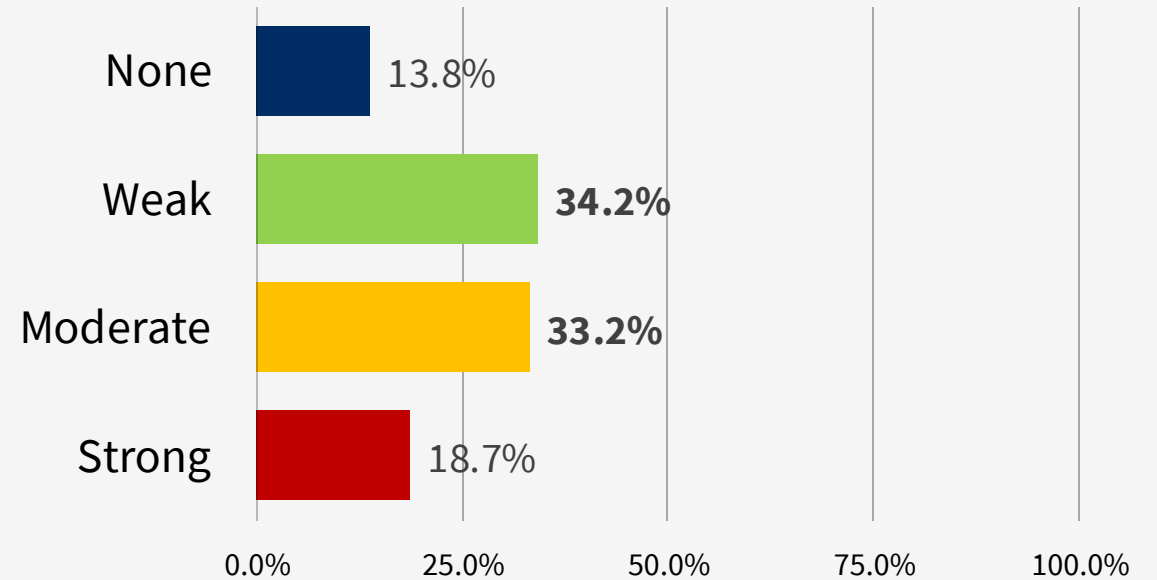
Abstract

Full text

Results

Linking words
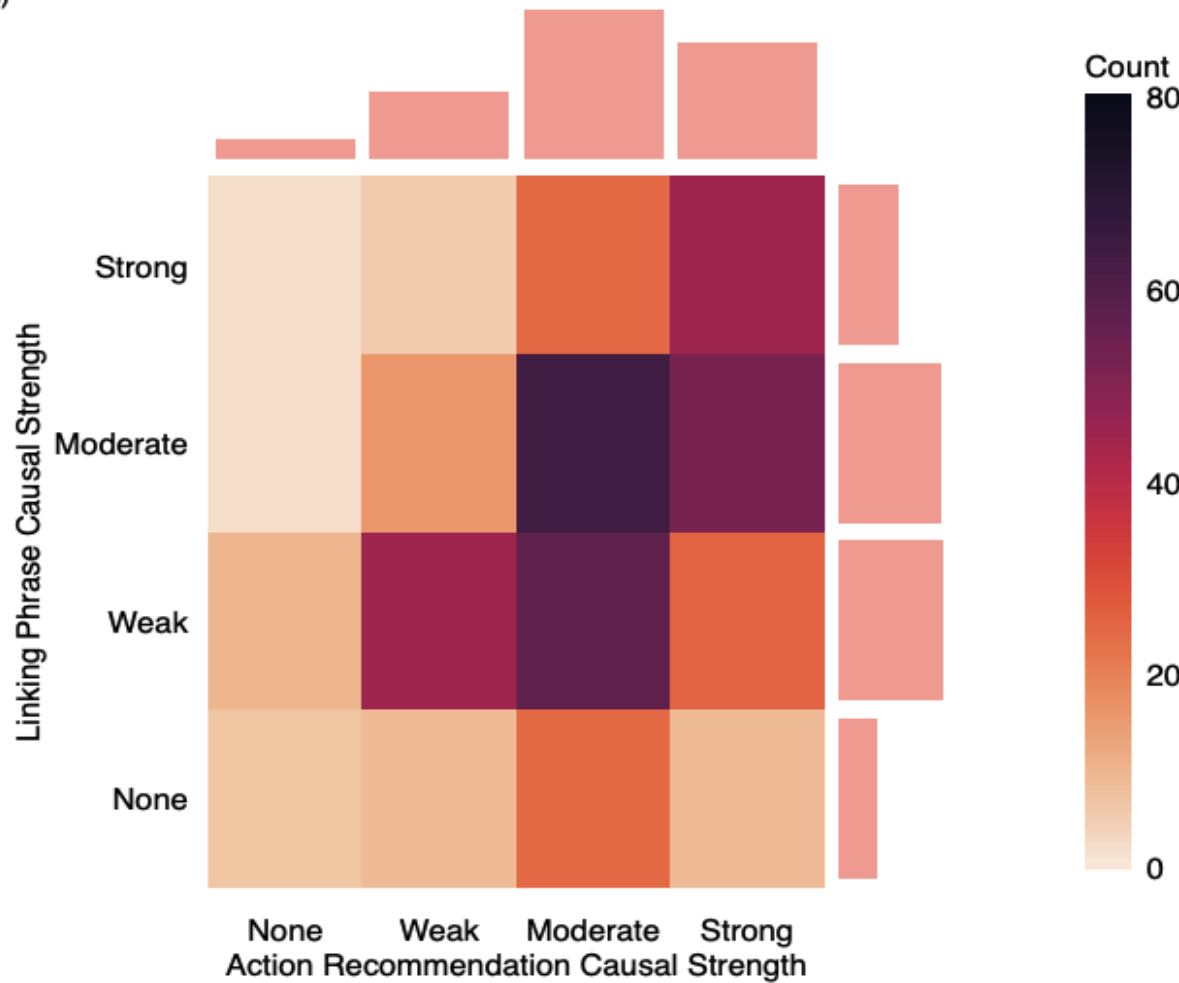

Frequency of use (top 6) n=1.170

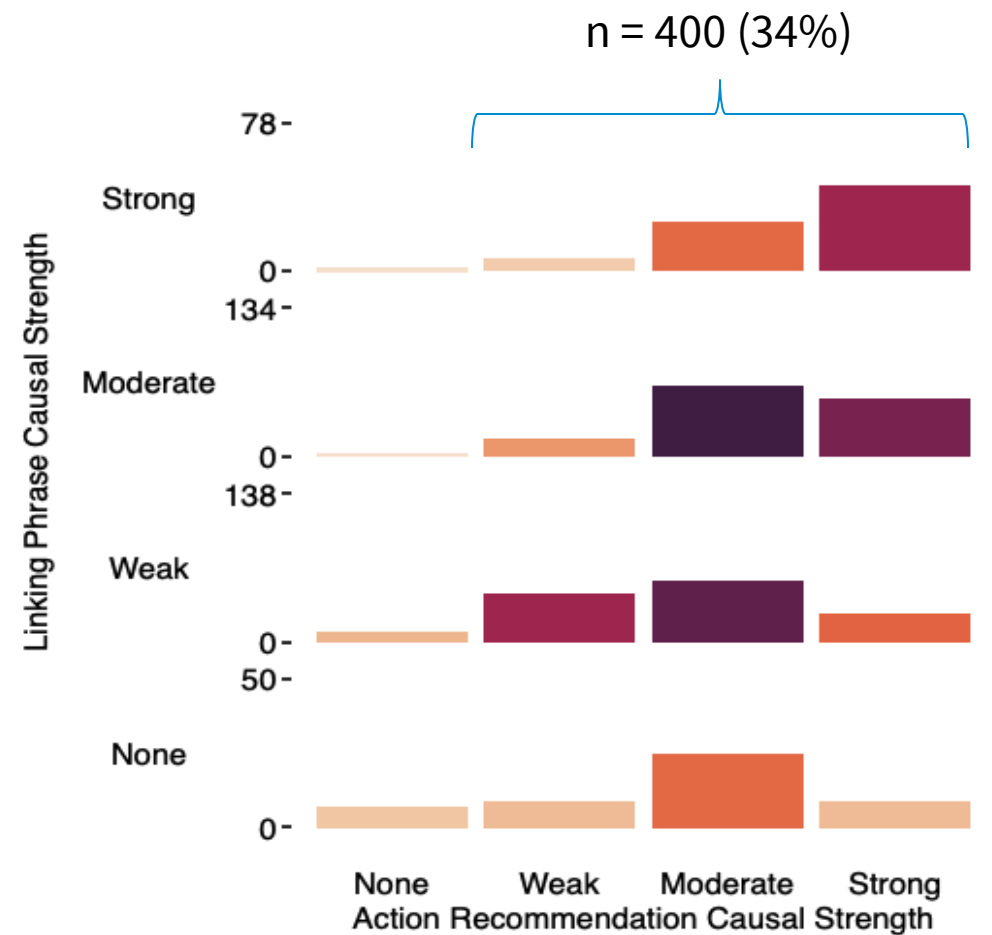**Causal implication of the results (in the abstract)**

- In 9 cases (0.8%) the linking words was cause
- No major differences in language used in RCTs (also "associate" is the most common)
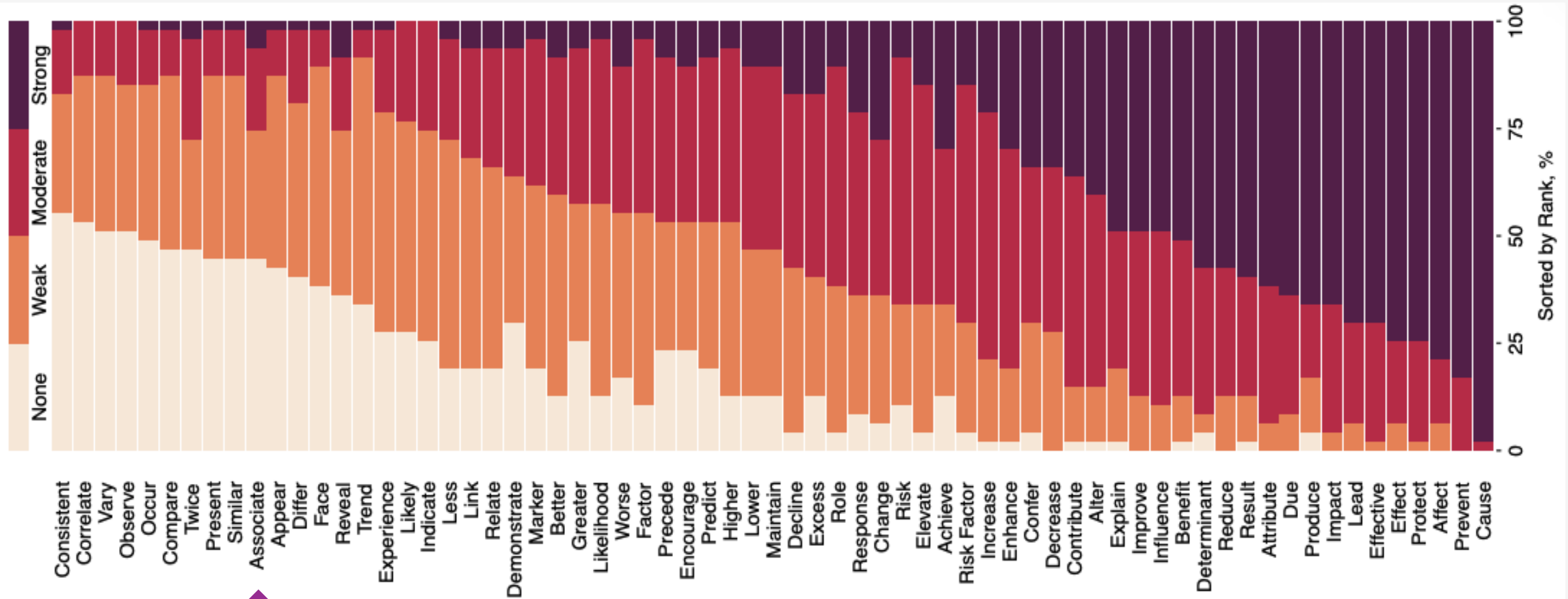
Linking phrases and action recommendations

*Action recommendations were identified in 34% of abstracts and in 60% of discussion sections

Strength of causal implication ratings for the most common linking words

Reviewer ratings for causal strength

Potential causal interest

- Most studies in our sample provided at least some indication of potential causal interest.

- While only 4% of studies presented formal causal models, most offered some discussion of the theoretical nature of the causal relationship between exposure and outcome or

- Among those that discussed theory,
  - 60% moderately or strongly indicated a theoretical causal relationship between exposure and outcome;
  - 25% of studies had a disclaimer statement regarding causality; and
  - 69% explicitly mentioned variations of the word "confound" (many control for)

Discussion (main findings summary)

1. Most common linking word was "associate"

2. Most studies used language that moderately or strongly implied causality

3. Vast majority of action implications imply that causality had been inferred

4. Action recommendations had stronger causal language than linking sentences

5. Causal implications commonly made through indirect means (discussion and confounding)

Substantial disconnect in language and implications

Causal implication in linking language and action recommendations are misaligned

Global Evidence Summit

"Schrödinger's causal inference

- This research undercuts the assumption that avoiding "causal" words leads to clarity of interpretation in medical research.

  - It is likely that the rhetorical standard of "just say association" has meant that many researchers no longer fully believe that the word "association" just means association.

  - We do not know the influence of journal editors, reviewers, authors, or academic community to the implicit and explicit rules of causal language
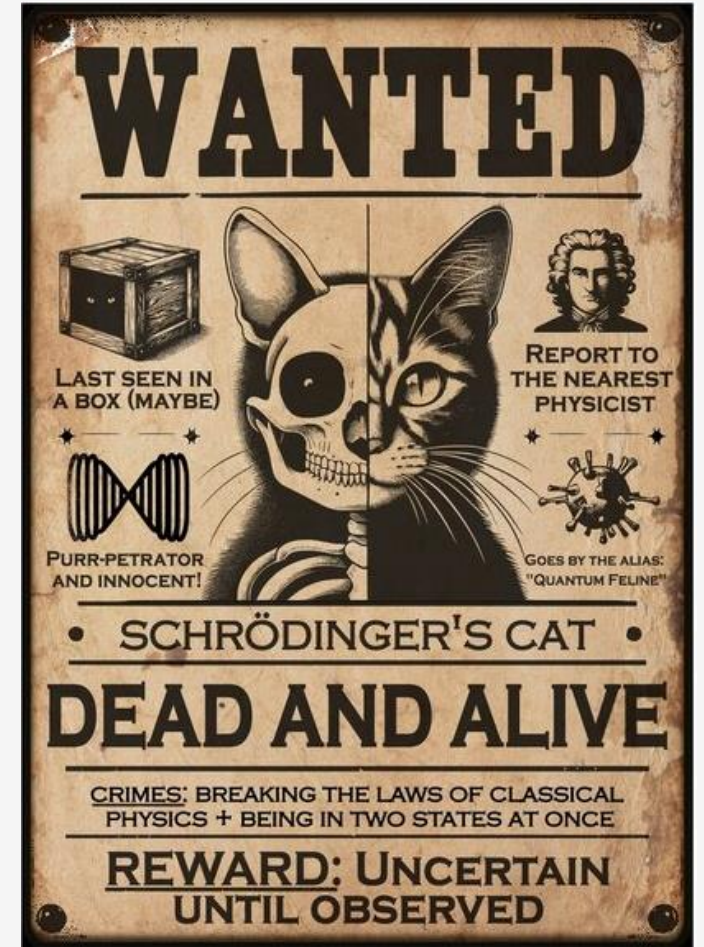
Draw causal inferences while claiming that causal inferences cannot be made !

Final remarks

- The practice of avoiding causal language linking exposures and outcomes appears to add little if any clarity

- Rather than policing which words we use to describe relationships between exposures and outcomes, we recommend focusing on how researchers, research consumers, and reviewers can better identify and assess causal inference study designs and assumptions.

- Quantitative empirical research should clearly state its target estimand to clarify the research question, including

  – explicitly stating when such estimands are causal.

  – Authors, reviewers, and editors should focus on being clear about what questions are being asked (37, 38), what decisions are being informed, and the degree to which we are and are not able to achieve those goals.

Fox MP, Edwards JK, Platt R, et al. The critical importance of asking good questions: the role of epidemiology doctoral training programs. Am J Epidemiol. 2020;189(4):261–264.

Causal and Associational Language in Observational Health Research: A Systematic Evaluation

Noah A. Haber*, Sarah E. Wieten, Julia M. Rohrer, Onyebuchi A. Arah, Peter W. G. Tennant, Elizabeth A. Stuart, Eleanor J. Murray, Sophie Pilleron, Sze Tung Lam, Emily Riederer, Sarah Jane Howcutt, Alison E. Simmons, Clémence Leyrat, Philipp Schoenegger, Anna Booman, Mi-Suk Kang Dufour, Ashley L. O'Donoghue, Rebekah Baglini, Stefanie Do, Mari De La Rosa Takashima, Thomas Rhys Evans, Daloha Rodriguez-Molina, Taym M. Alsalti, Daniel J. Dunleavy, Gideon Meyerowitz-Katz, Alberto Antonietti, Jose A. Calvache, Mark J. Kelson, Meg G. Salvia, Camila Olarte Parra, Saman Khalatbari-Soltani, Taylor McLinden, Arthur Chatton, Jessie Seiler, Andreea Steriu, Talal S. Alshihayb, Sarah E. Twardowski, Julia Dabravolskaj, Eric Au, Rachel A. Hoopsick, Shashank Suresh, Nicholas Judd, Sebastián Peña, Cathrine Axfors, Palwasha Khan, Ariadne E. Rivera Aguirre, Nnaemeka U. Odo, Ian Schmid, and Matthew P. Fox

@jacalvache

jacalvache@unicauca.edu.co

Using evidence. Improving lives.