CAUSAL AND ASSOCIATIONAL LANGUAGE IN OBSERVATIONAL HEALTH RESEARCH: A SYSTEMATIC EVALUATION  BY NOAH A. HABER ET MANY AL.  HTTPS://DOI.ORG/10.1093/AJE/KWAC137

Ionica Smeets, July 5th, 2023

“Statistics Communication and (in)numeracy” reading group
“Some author guidelines explicitly prohibit the use of causal language in studies other than randomized controlled trials (RCTs), often justified by the inaccurate, but common, belief that causal inference is only possible with RCTs.”
“It is not clear what “counts” as causal language, with no clear standards and few attempts to define and categorize what constitutes causal language.”
This study systematically examined the linking language used in studies with a main exposure and outcome in the high-profile medical and epidemiologic literature.

<table>
<thead>
<tr>
<th>Identify</th>
<th>Identify the linking words and phrases used to describe relationships between exposures and outcomes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Generate</td>
<td>Generate estimates of the strength of causality stated or implied by the linking phrases and sentences.</td>
</tr>
<tr>
<td>Examine</td>
<td>Examine the prevalence of action recommendations that would require causal inference to have been made</td>
</tr>
<tr>
<td>Examine</td>
<td>Examine disconnects between causal implications in linking sentences and action implications</td>
</tr>
</tbody>
</table>
DATA SELECTION

Journals (peer-reviewed, medical, ‘important’)

Individual studies (on humans, main research question must be about causal relation, quantitative study)
Articles were randomly assigned to 3 of 18 screening reviewers, with 2 independent reviewers and 1 arbitrating reviewer.

Identify main outcome and exposure, linking sentence and action recommendations.
<table>
<thead>
<tr>
<th>Rating</th>
<th>Linking Sentence</th>
<th>Action Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA</td>
<td>The linking sentence does not imply in any way that a causal relationship was identified.</td>
<td>No action recommendation exists.</td>
</tr>
<tr>
<td>None</td>
<td>The linking sentence might imply that a causal relationship was identified, but it is unclear.</td>
<td>The action recommendation may be made appropriately had a causal relationship been identified.</td>
</tr>
<tr>
<td>Weak</td>
<td>The linking sentence mostly implies that a causal relationship was identified, but it is unclear.</td>
<td>The action recommendation most likely could only be made appropriately had a causal relationship been identified.</td>
</tr>
<tr>
<td>Moderate</td>
<td>The linking sentence clearly implies that causality had been identified.</td>
<td>The action recommendation could only be made appropriately had a causal relationship been identified.</td>
</tr>
</tbody>
</table>
45.7% of the articles used ‘associate’ as a linking word

Only 0.8% of studies use the word ‘cause’
Ratings among reviewers ($n = 47$) for causal implication of root words were highly heterogeneous
Figure 2. Summary scores for the degree of causal implication in linking sentences and action recommendations in a study of causal and associational linking language in observational research and health evaluation literature, showing the frequency of key strength of causal implication metrics for the 1,170 non-randomized-control trial studies in our sample, as indicated by the arbitrating reviewer. A–C) The strength of causal implication ratings for the language ratings in the abstract, discussion, and popout sections; D–F) the strength of causal implication ratings for the action recommendations in the abstract, discussion, and popout sections.
THE ALIGNMENT OF ADVICE AND KEY LINKING PHRASE
“Our results suggest that “Schrödinger’s causal inference”—where studies avoid stating (or even explicitly deny) an interest in estimating causal effects yet are otherwise embedded with causal intent, inference, implications, and recommendations—is common in the observational health literature.”
“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.”
PubMed Search Hits ($n = 31,014$)

Screened (Process Continued Until Journal/Study Type Quotas Were Met) ($n = 5,507$)

Accepted by Both Reviewers ($n = 1,170$)

Conflicting Reviewers ($n = 1,236$)

Arbitrator Accepted ($n = 556$)

Articles Accepted for Potential Full Review ($n = 1,726$)

Assigned to Reviewers for Full-Text Review ($n = 1,275$ articles)

Final Article Review Data Set ($n = 1,275$ Articles $\times$ 3 Reviewers $= 3,825$ Reviews)

Never Screened (Quotas Met) ($n = 25,507$)

Both Reviewers Rejected ($n = 3,101$)

Arbitrator Did Not Accept ($n = 680$)

Screened Articles Set Aside to Serve as Potential Replacements ($n = 451$)

Replaced During Review ($n = 70$)

Never Reviewed ($n = 381$)