Quality of peer-reviewed published reviews: a case study of sugar-sweetened beverages and health.

AJCN 2011 (in press)

Douglas L. Weed, M.D., M.P.H., Ph.D.
Michelle D. Althuis, Ph.D.
Pamela J. Mink, Ph.D.

15 September 2011
IOM Food Forum Meeting
Washington D.C.
Purpose of Reviews

• Scientific statements (claims) regarding causation, including mechanisms
• Scientific statements regarding efficacy (therapeutic and preventive interventions)
• Public health and medical practice recommendations
• Regulatory and practice guidelines (policy)
• Scientific research agendas

Milne, 1993; Weed, 1997; Breslow et al., 1998; Weed et al., 2011
Quality of Reviews

• Concerns arose in the mid-1980’s about methodological quality
• Standard practice was the unsystematic ‘narrative’ review
• Methodological (quality) guidelines for reviews in many journals:
  – Am J Prev Med; Annals Int Med; Annals NY Acad Sci; BMJ; Can Fam Physician; JAMA; J Clin Epidemiol; J Epidemiol Commun Health; JNCI

Methodologic Guidelines for Review Papers
Weed JNCI 1997

• Statement of Purpose
• Search Methods and Inclusion/Exclusion Criteria
• Criteria for Evaluating Validity (Quality) of Studies
• Methods for Summarizing Evidence
• Criteria for Causal Conclusions (or for other stated purposes)
Emergence of a Consensus: Reviews should be Systematic

Limits Bias
Implements Transparency
Implements the Reliability of:
- Causal Conclusions
- Public Health Recommendations
- Clinical Practice Recommendations
- Regulatory Guidelines

State of the Art:
- International Agency for Research on Cancer (IARC)
- US Preventive Services Task Force
- Cochrane Collaboration
- Published Literature (Occasional)

Muir Gray 1994; Mulrow 1994; Oxman 1994; Mullen 2006
Quality of Reviews Remains an Issue

• Many reviews published in epidemiology, clinical oncology, and general medical journals are not systematic.

Breslow et al., 1998; Mulrow, 1987; Bramwell, 1997; Rochon, 2002; Moher 2007; Lundh 2009
Quality of Reviews Remains an Issue

• Many reviews published in epidemiology, clinical oncology, and general medical journals are not systematic.

  Breslow et al., 1998; Mulrow, 1987; Bramwell, 1997; Rochon, 2002; Moher 2007; Lundh 2009

• Are reviews of the relationships between sugar sweetened beverages (SSB) and health outcomes systematic?
Design of a “Review of Reviews” of SSB and Obesity, Metabolic Syndrome, Diabetes, and Coronary Heart Disease

• Purpose: to assess the methodological quality of SSB reviews of epidemiology studies
• Systematic Search and Selection Criteria
• Data Abstraction
• AMSTAR: Tool for Assessing Quality*
• Data Analysis

* Oxman, 2006; Shea, 2007; Shea, 2009
Design of a Systematic “Review of Reviews”

• **Systematic Search and Selection Criteria**
  – 3 Searches: PubMed, Cochrane, and Hand-searching
  – Meta-analyses and reviews: June 2001-June 2011
  – Abstracts reviewed to identify papers for full text review by two authors (MDA, DLW)
  – Outcomes: Obesity, MetSyn, T2DM, CHD
# Search Results

**Figure 1: Results of a search for published reviews/meta-analyses focused predominately on SSB and health outcomes**

<table>
<thead>
<tr>
<th></th>
<th>Obesity Abstracts:</th>
<th>T2DM Abstracts:</th>
<th>MetSyn Abstracts:</th>
<th>CHD Abstracts:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer Search =</td>
<td>423</td>
<td>157</td>
<td>50</td>
<td>334</td>
</tr>
<tr>
<td>Exclusions*:</td>
<td>Not related(^b) = 375</td>
<td>Not related(^b) = 149</td>
<td>Not related(^b) = 45</td>
<td>Not related(^b) = 330</td>
</tr>
<tr>
<td>Not analytical(^c) = 18</td>
<td>Not analytical(^c) = 3</td>
<td>Not analytical(^c) = 2</td>
<td>Not analytical(^c) = 3</td>
<td></td>
</tr>
<tr>
<td>Full Text review =</td>
<td>30</td>
<td>5</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Not primary focus(^d) : -17</td>
<td>Not primary focus(^d) : -3</td>
<td>Not primary focus(^d) : -1</td>
<td>Hand search of ref. list: +1</td>
<td></td>
</tr>
<tr>
<td>Hand search of ref. list: +4</td>
<td>Hand search of ref. list: +1</td>
<td>Hand search of ref. list: +1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review some research(^e): -1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obesity Studies Included:</td>
<td>16</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>T2DM Studies Included:</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MetSyn Studies Included:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CHD Studies Included:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. Reasons for exclusion upon review of the title or published abstract (Pubmed).
b. Publications that were not related to this review. Examples include, but are not limited to, studies of dietary patterns, coffee, tea, or alcohol consumption.
c. Title suggested possible inclusion, but abstract review revealed the publication was not focused on analytic epidemiologic studies (for example, mechanistic studies, descriptive epidemiology, or commentary).
d. Publications that included some discussion of epidemiologic studies of SSB and the outcome of interest, but as *one of many* focuses of the paper.
e. Papers were included if the stated goal or approach implied an attempt to summarize all available literature. One paper was excluded because the stated goal was to review "some of the latest research."

Note: T2DM = Type 2 Diabetes Mellitus; MetSyn = Metabolic Syndrome; CHD = Coronary Heart Disease.
Studies Included (n = 17)

- Hu FB and Malik VS. *Physiol Behav*. 2010;100(1):47-54.
Data Abstraction and Analysis

- Study purpose and population(s)
- Study variables and outcomes
- Number and type of studies summarized
- Method(s) of analysis/interpretation
- Described (by author(s)) as systematic?
- Methods section included?
- AMSTAR: standardized instrument
# AMSTAR questions

Oxman et al., 2006; Shea et al., 2007; Shea et al., 2009

<table>
<thead>
<tr>
<th>Instrument/Question</th>
<th>Amstar Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. &quot;A priori&quot; design?</td>
<td></td>
</tr>
<tr>
<td>2. Duplicate study selection/data abstraction?</td>
<td></td>
</tr>
<tr>
<td>3. Comprehensive literature search?</td>
<td></td>
</tr>
<tr>
<td>4. Publication status used for inclusion?</td>
<td></td>
</tr>
<tr>
<td>5. Listed included/excluded studies?</td>
<td></td>
</tr>
<tr>
<td>6. Characteristics of included studies tabulated?</td>
<td></td>
</tr>
<tr>
<td>7. Scientific quality of included studies assessed?</td>
<td></td>
</tr>
<tr>
<td>8. Scientific quality used appropriately to formulate conclusions?</td>
<td></td>
</tr>
<tr>
<td>9. Appropriate methods used to combine studies?</td>
<td></td>
</tr>
<tr>
<td>10. Likelihood of publication bias assessed?</td>
<td></td>
</tr>
<tr>
<td>11. Conflict of interest included?</td>
<td></td>
</tr>
</tbody>
</table>
Results

• 17 Reviews and Meta-analyses Analyzed
• 65% Documented the search strategy
• 16 reviews of Obesity/weight
  – 11 presented findings combined for children and adults
  – No consistency between reviews in terms of number or types of studies reviewed
• 3 reviews of T2DM, 3 of MetSyn, 2 CHD
  – Dominated by a single group of investigators
  – Fewer than 10 studies summarized for these outcomes
Results: Methods Used

• Narrative format used most frequently: 7/17
• 11/17 did not summarize data from original studies in a tabular format
• 4 meta-analyses: combined studies that differed by geography, age, and gender
Results

• 17 Reviews and Meta-analyses Analyzed

• Overall AMSTAR score: mean 4.4 (out of 11)
  – Std deviation (2.46), Median = 4, Range (0-8.5)
  – 9/17 (53%) scored 4 or lower
  – 4/17 (23.5%) scored 4.5-6.5
  – 4/17 (23.5%) scored 7-8.5

• AMSTAR scores did not improve with year of publication

• AMSTAR scores not related to authors’ conclusions
  – 8 reviews (association exists) mean 4.1
  – 9 reviews (equivocal) mean 4.7 (t-test p-value 0.84)
## Results (including Kappa statistics)

### Table 3. Kappa statistic for agreement for each AMSTAR item and the proportion of "Yes" scores

<table>
<thead>
<tr>
<th>Instrument/Question</th>
<th>Kappa (95% CI)</th>
<th>Reviews scoring &quot;Yes&quot; (N=17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amstar Question</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. &quot;A priori&quot; design?</td>
<td>0.72 (0.35-1.0)</td>
<td>71%</td>
</tr>
<tr>
<td>2. Duplicate study selection/data abstraction?</td>
<td>1.00 (1.0-1.0)</td>
<td>12%</td>
</tr>
<tr>
<td>3. Comprehensive literature search?</td>
<td>1.00 (1.0-1.0)</td>
<td>29%</td>
</tr>
<tr>
<td>4. Publication status used for inclusion?</td>
<td>0.88 (0.66-1.0)</td>
<td>44%</td>
</tr>
<tr>
<td>5. Listed included/excluded studies?</td>
<td>0.77 (0.34-1.0)</td>
<td>15%</td>
</tr>
<tr>
<td>6. Characteristics of included studies tabulated?</td>
<td>0.88 (0.64-1.0)</td>
<td>62%</td>
</tr>
<tr>
<td>7. Scientific quality of included studies assessed?</td>
<td>0.88 (0.64-1.0)</td>
<td>62%</td>
</tr>
<tr>
<td>8. Scientific quality used appropriately to formulate conclusions?</td>
<td>0.66 (0.33-0.99)</td>
<td>50%</td>
</tr>
<tr>
<td>9. Appropriate methods used to combine studies?</td>
<td>0.77 (0.34-1.0)</td>
<td>15%</td>
</tr>
<tr>
<td>10. Likelihood of publication bias assessed?</td>
<td>1.00 (1.0-1.0)</td>
<td>12%</td>
</tr>
<tr>
<td>11. Conflict of interest included?</td>
<td>1.00 (1.0-1.0)</td>
<td>71%</td>
</tr>
</tbody>
</table>
Conclusions

• Reviews published between 2001 and 2011 on SSB and Obesity, T2DM, MetSyn, and CHD scored poorly on AMSTAR (mean score = 4.4).
• Many basic elements of systematic reviews were simply missing, resulting in a lack of transparency.
• Most reviews were unclear about the types of studies, populations, and whether results were from primary or subgroup analyses.
• A mismatch exists between the methodological community (where systematic reviews are advocated) and the practice community.
• Systematic reviews of the SSB-health outcomes issue are needed to better inform policy.
Acknowledgements and Funding

Co-authors

Michelle D. Althuis, Ph.D.
Pamela J. Mink, Ph.D.

Funding

Coca-Cola Company

This project was conceived, designed, analyzed, and interpreted by the authors with a contractual agreement from the funder that the authors had full and final control over all aspects of the project.
Discussion

• Should editors, peer reviewers, and funding agencies insist upon methodologically systematic reviews?

• Two classes of reviews?

“All types of literature reviews can benefit from a systematic approach, with the exception perhaps of a review that aims to argue for a particular position and that would select evidence accordingly.” (Mullen and Ramirez, 2006)