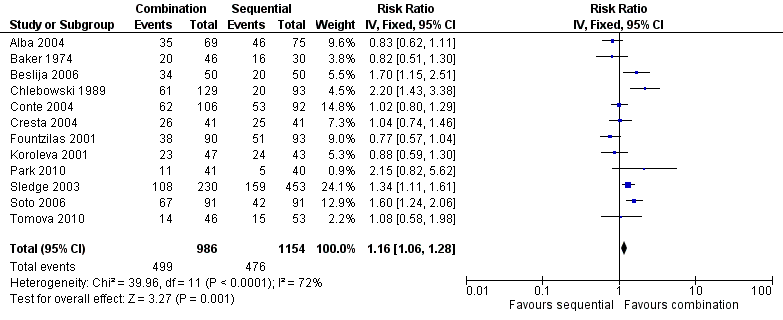
**Data and Methods**

*Cochrane Database of Systematic Review*

The Cochrane Database of Systematic Review (CDSR), which is now published by Cochrane Review Group, is a one of the leading journals publishing systematic review articles about studies on human health intervention and diagnosis to inform diverse sets of stakeholders such as policy makers, patients, health professionals, starting from April 1995[[1]](#footnote-1). The publisher, Cochrane Review Group which was originally found in the U.K. as a non-profit organization, now has 53 different groups[[2]](#footnote-2) with over 30,000 expert volunteers in health science across the world (Cochrane, 2019). Even though the quality was questioned (e.g., Olsen et al. 2001, Büchter et al. 2016), the CDSR has been regarded as one of the most reliable medical resources that guide the evidence-based medical practice (Handoll et al. 2008) since its foundation.

Only reviews with the meta-analyses containing at least more than three studies were selected because we only have the definitive conclusions about the claims when a review examines equal to or more than three studies. Reviews in the CDSR with the meta-analyses can either consider fixed effects or random effects: the fixed effects models assume that all the studies in the review have a common parameter of the effect size, only allowing the random errors within studies while the random effects model allows between-study variances. Each review provides the meta-analysis *p*-values and confidence intervals based on the methodology proposed by Higgins et al., (2011), combining the statistical information from the studies that a review attempts to examine. For instance, review CD008792 published in 2013 evaluates the efficacy of the combination chemotherapy in women with metastatic breast cancer comparing to the sequential application of the same drugs, using 12 studies. Figure 1 shows the overall response rates of the two different therapies from those complied by the authors of the review.

**Figure 1. An Example of a Report from the CDSR [[3]](#footnote-3)**



In total, the number of the IDs of the Cochrane articles for this study is 4,543, which were published before Nov 18, 2017. Each review has its history. Protocols in a preparation stage is first produced by the authors (which sometimes are not published) and they conduct a review after a set of original research articles are identified, which means that a Cochrane review article at least has two elements in its history. If a major updated is conducted later, a newer version will replace the last one, but it maintains the same review ID. Minor amendments such as changes in formatting and an addition of a summary in plain English can also take place, but it does not lead to an updated version of a review. For example, review CD001321 that investigates the efficacy of Cranberries in preventing urinary tract infections was first published in 2003 but updated in 2004, 2008, and 2012. Considering this, I developed a web-scraper to collect all the version histories of 4,543 reviews (which led to 8,212 different versions of the 4,543 reviews) and selected the most recent articles before Nov 18, 2017. The mean of the number of authors of the most recent reviews is 4.59 (std = 2.11, min = 1, max = 41).

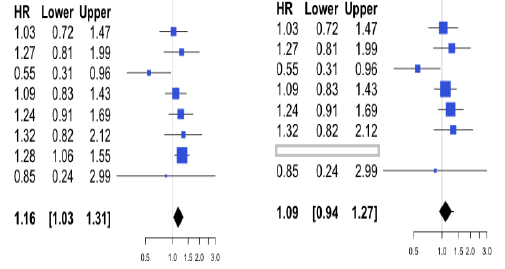
The reference sections of each Cochrane review disclose both the included and excluded original studies.[[4]](#footnote-4) The metadata of these original studies (e.g., the individual studies in Figure 1) in each review, including the names of the authors, the titles, the journals’ names, and the publication year, and the PubMed IDs of original research articles were also collected using a web-scraper. The PubMed ID is of importance here. Each PubMed ID can be used to identify a medical research article in the MEDLINE database, which is co-maintained by the National Library of Medicine (NLM) and the National Center for Biotechnology Information (NCBI) at the National Institutes of Health (NIH) of the United States. The PubMed IDs of each original research article examined by each Cochrane review allowed us to collect the additional data such as the abstracts of the original articles, the references of the original articles, and the list of other original studies published by the authors with the same name, all of which we used to compute the social and intellectual dependencies of the original studies within review articles.

**Measurements**

*Dependent variable: The r-value as a measure replicability*

In this paper, the *r*-value proposed by Shenhav et al. (2015) is used as a measure of replicability, which can be computed based on a meta-analysis. The computation of the *r*-value basically follows the logic of sensitivity analysis in which whether the significance of a claim holds when samples of studies in a systemic review are excluded. By measuring the impact of each study or sets of studies on the overall conclusion of a review, the *r-*value can provide more rigorous evidence of the replicability of scientific claims, especially considering that one study can drive the significance of a meta-analysis with the increased power. Figure 2 graphically shows how the logic can be applied using an example from a review in the CDSR: when study 7 is excluded, the overall conclusion becomes insignificant.

**Figure 2. The Forest Plot from Review CD008792 (Left) and Excluding Study 7 (Right)**

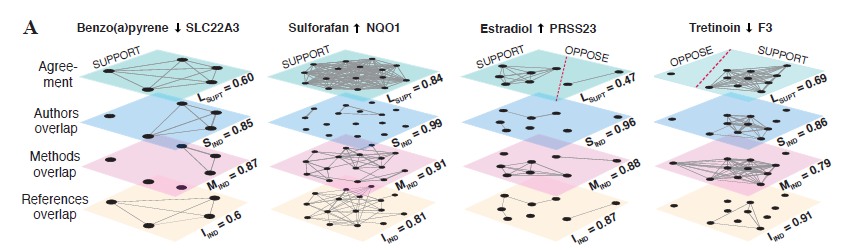


This basic logic can be easily extended beyond the single case exclusion. For example, a researcher can test the effect of the different the number of cases excluded from the original pool of studies. Appendix A provides the detailed computation procedures of the *r*-value.

*Independent variables*

*a) Author, References overlap*

🡪 Draw a similar figure



**References**

Büchter, Roland Brian, and Dawid Pieper. "Most overviews of Cochrane reviews neglected potential biases from dual authorship." *Journal of clinical epidemiology* 77 (2016): 91-94.

Lindson, Nicola, Samantha C. Chepkin, Weiyu Ye, Thomas R. Fanshawe, Chris Bullen, and Jamie Hartmann‐Boyce. "Different doses, durations and modes of delivery of nicotine replacement therapy for smoking cessation." *Cochrane Database of Systematic Reviews* 4 (2019).

Handoll, Helen H., William J. Gillespie, Lesley D. Gillespie, and Rajan Madhok. "The Cochrane Collaboration: a leading role in producing reliable evidence to inform healthcare decisions in musculoskeletal trauma and disorders." *Indian journal of orthopaedics* 42, no. 3 (2008): 247.

Hoff, Peter D., Adrian E. Raftery, and Mark S. Handcock. "Latent space approaches to social network analysis." *Journal of the american Statistical association* 97, no. 460 (2002): 1090-1098.

Higgins, Julian PT, Douglas G. Altman, Peter C. Gøtzsche, Peter Jüni, David Moher, Andrew D. Oxman, Jelena Savović, Kenneth F. Schulz, Laura Weeks, and Jonathan AC Sterne. "The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials." *Bmj* 343 (2011): d5928.

Olsen, Ole, Philippa Middleton, Jeanette Ezzo, Peter C. Gøtzsche, Victoria Hadhazy, Andrew Herxheimer, Jos Kleijnen, and Heather McIntosh. "Quality of Cochrane reviews: assessment of sample from 1998." *Bmj* 323, no. 7317 (2001): 829-832.

1. From 1995 to 2009, CDSR was published quarterly; since then, it has been published in a monthly basis. [↑](#footnote-ref-1)
2. The full list of the review groups can be found in <https://www.cochranelibrary.com/about/cochrane-review-groups>. Note that some topics such as influenza virus can be address by several groups including Acute Respiratory Infections Group, Childhood Cancer Group, etc. [↑](#footnote-ref-2)
3. Dear, Rachel F., Kevin McGeechan, Marisa C. Jenkins, Alexandra Barratt, Martin HN Tattersall, and Nicholas Wilcken. "Combination versus sequential single agent chemotherapy for metastatic breast cancer." *Cochrane Database of Systematic Reviews* 12 (2013). [↑](#footnote-ref-3)
4. The authors first collect the relevant published research and sort them with some evaluation criteria to safeguard the scientific validity or to guarantee fair comparisons. The authors decide to exclude studies when they do not meet the criteria or if the study designs are inconsistent with other studies. A reader can find detailed inclusion criteria in each review. In addition to this, note that each review only includes one of studies duplicated by the same authors. For example, the reference section of the review CD001321 includes the two studies conducted by Salo et al. (2010) and Salo et al. (2011). Since the two studies are conducted by the same group of the authors, the review CD001321 only cites Salo et al. (2010). But when the PubMed IDs were identified, the metadata of every original article in the reference sections was collected regardless of whether it is included or excluded, and whether it is directly cited. [↑](#footnote-ref-4)