# SWAR 14: Comparison of data extraction techniques (Covidence versus Word-form) for reviewers doing their first systematic review

**Objective of this SWAR**
This SWAR would explore the relative benefits of two techniques for double data extraction in a systematic review, where the reviewers are doing their first systematic review, but with previous knowledge and training on systematic review methodology. The two techniques to be compared are Covidence software and a pre-designed data extraction Word-form. This SWAR would also assess the impact of data extraction errors on the effect estimate obtained through the meta-analysis.

Study area: Data extraction
Sample type: Reviewers doing their first systematic review (e.g. postgraduates)
Estimated funding level needed: Very Low

**Background**
Systematic reviews are robust research and, for questions about the comparative effectiveness of interventions, systematic reviews of randomised trials are regarded as the highest level of evidence (1). Thus, they need to be conducted as accurately and fault-free as possible. The data extraction phase is a critical phase, and the methods used in data extraction need to minimise biases and human errors (2). To extract data, different tools exist to help reviewers in conducting this critical phase (2,3). However, each tool has its drawbacks and benefits, and the choice of the appropriate tool depends on the type and size of the systematic review, and on the needs and resources of the reviewers (2,3,4). The choice of the most suitable tool for data extraction is crucial (4), especially for researchers doing their first systematic review, e.g. postgraduates. There are three main types of data extraction form: paper forms, electronic forms and data systems (4). With the decline in use of paper forms (especially when review teams work in different countries), electronic forms and internet-based tools, e.g. Word-form and Covidence software, are more prominent and will be assessed in this SWAR. For both Word-form and Covidence software, this SWAR will use a double data extraction method, as recommended by Cochrane and other authorities (2,5).

In this implementation of the SWAR, four reviewers will be randomised to one of the following roles when they extract the data from each study in the review: (a) first data extractor and using Covidence; (b) second data extractor and using Covidence; (c) first data extractor and using the Word-form; or (d) second data extractor and using the Word-form. The randomisation will be applied for each included study, giving each reviewer the possibility of being paired with each of the other reviewers. This will be achieved by assigning each of the four reviewers a number from 1 to 4 and producing the 36 different sequences for the numbers 1 to 4. Each study will be randomly matched with one of these sequences, and the reviewers will be assigned to one of the four roles listed above according to that sequence. To ensure allocation concealment, the randomisation will be conducted through a third-party. If there are more than 36 included studies, a second set of sequences will be used.

An online forum for data extraction will be set up and the four reviewers will have separate accounts to access Covidence and the Word-form through this forum. The reviewers will work on one study per day and will receive their allocated role for this study according to the randomisation process when they log in. Each reviewer will extract data independently before making a primary submission to the consensus window. When both reviewers have submitted their results for the study, they will log into the consensus window to discuss and reach consensus. After this, the two reviewers will submit the results as a final submission to an experienced systematic reviewer (third party) and log out. The third-party will then check the final submissions to count the number of errors.

**Interventions and comparators**
Intervention 1: double data extraction through Covidence.
Intervention 2: double data extraction through Word-form

Index Type: Full Review

**Method for allocating to intervention or comparator**
Randomisation

**Outcome measures**
Primary: The third-party will check each item for the presence or absence of the following to measure their proportions:
1. Inaccurate error: when the extracted information is incorrect.
2. Omission error: when the relevant information is present in the study report but has not been extracted.
3. Incompleteness error: when the extracted information is correct but not complete.
4. Total errors: any of the previous errors (as a primary outcome).

The online forum will be used to measure the time needed to complete the data extraction, based on information stored by the forum for the times of (a) log in, (b) primary submission, (c) log in to the consensus window, and (d) final submission. The time outcomes will be calculated as follows:
1. Average time for primary data extraction: mean duration between each reviewer's log in and their primary submission.
2. Time of reaching consensus: duration from opening the consensus window to the final submission.
3. Total time for data extraction per study: mean time for primary extraction and time to reach consensus (as a primary outcome).
Secondary:

**Analysis plans**
Sample size calculations for this SWAR are based on a similar study by Buscemi et al (6), which compared two different methods of data extraction (one reviewer with verification by a second reviewer versus two reviewers working independently) in terms of accuracy and time-saving.

The standard deviation of the total time needed for data extraction per a study in the double data extraction group arm in Buscemi's study (SD= 69.9 minutes, Mean= 135.9 minutes) was used as a common standard deviation to calculate the number of studies to include in this SWAR to detect a mean difference (MD) of 47.5 minutes. This mean difference was chosen to investigate the claim of the Covidence software group, that Covidence produces an reduction of 35% in time needed (7) (i.e. 35% of 135.9 minutes). This MD is compatible with the 49 minutes detected by Buscemi (6). We used paired-t-test (with alpha-value=0.05 and correlation=0) to produce plots of the required number of pairs in our SWAR by power and by effect size. From the two plots, 36 studies are needed for 80% power and MD of 47.4 minutes.

The primary outcome for accuracy (proportions of total errors) was used to calculate the total number of items from all studies needed to detect discordant total errors proportion values (14.54% and 17.69%) using McNemar test. The produced plot of the required number of pairs in our SWAR by power indicates that we need at least 2600 items to have 80% power (with alpha-value=0.05). If we seek 82 items from each study, this will give us 2952 items in total. These discordant proportions are based on Buscemi's study, where there was a difference of 3.2% between the two methods of data extraction (6). A paired t-test will be used for the efficacy outcomes and the McNemar test for the accuracy outcomes.

We will study the impact of the errors on the effect estimate of the meta-analysis by comparing three meta-analysis: (1) using Covidence data; (2) using Word-form data and (3) using third-party corrected data.

**Possible problems in implementing this SWAR**
We expect that implementing this SWAR in different reviews and with different reviewers may lead to different results. Another problem is with masking of the reviewers but we hope that not disclosing the hypothesis of the SWAR to them will help. Finally, this SWAR only compares the two chosen data extraction techniques, so the results may not be generalisable to other techniques.

**References**

1. Oxford Centre for Evidence-based Medicine – Levels of Evidence (March 2009), Produced by Phillips B, Ball C, Sackett D, Badenoch D, Straus S, Haynes B, Dawes M since November 1998. Updated in March 2009 by Howick J.
2. Li T, Higgins JPT, Deeks JJ (editors). Chapter 5: Collecting data. In: Higgins JPT, Thomas J, Chandler J, et al (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019.
3. Elamin MB, Flynn DN, Bassler D, et al. Choice of data extraction tools for systematic reviews depends on resources and review complexity. Journal of Clinical Epidemiology 2009;62(5):506-10.
4. Li T, Vedula SS, Hadar N, Parkin C, et al. Innovations in data collection, management, and archiving for systematic reviews. Annals of Internal Medicine 2015;162(4):287-94.
5. Eden J, Levit L, Berg A, Morton S (editors). Committee on Standards for Systematic Reviews of Comparative Effectiveness Research; Board on Health Care Services. Finding What Works in Health Care: Standards for Systematic Reviews. Washington, DC: National Academies Pr; 2011.
6. Buscemi N, Hartling L, Vandermeer B, et al. Single data extraction generated more errors than double data extraction in systematic reviews. Journal of Clinical Epidemiology 2006;59(7):697-703.
7. Covidence 2020, accessed on 25 July 2020 (https://www.covidence.org).

**Publications or presentations of this SWAR design**

**Examples of the implementation of this SWAR**

People to show as the source of this idea: Ali Mulhem
Contact email address: ali.mulhem@kellogg.ox.ac.uk
Date of idea: 8/JUN/2020
Revisions made by:
Date of revisions: