Abstract-Paper-Registry Concordance Project

Methods:

Journal and trial selection

1) We will identify the 5 psychiatry journals between 2014 through 2016 with the highest impact-factor based on the 1-year Science Citation Index Expanded (SCIE) impact factor using in the 2017 Journal Citation Reports (JCR), published by Clarivate Analysis. We will exclude journals that have not published at least 10 RCTs over this period to focus on journals that are publishing trials regularly. If a journal is excluded, the next journal on the 2017 JCR that meets journal inclusion criteria will be used.
2) We will identify randomized controlled trials published between the years of 2014 and 2016 (inclusive) in the identified top 5 journals. To identify articles, we will use the following search strategy: each of the 5 identified journals' names will each be used as a search term on Pubmed.gov with the specifier "[Journal]". Filters for a start date of January 1, 2014 and an end date of December 31, 2016 and "Randomized Controlled Trial" on PubMed.gov will be used.
3) Titles and abstracts will be independently screened for inclusion and exclusion criteria by two readers. Disagreements will be resolved by consensus or by a third reader if consensus is not reached. Reasons for exclusion will be scored.
4) Inclusion criteria: To be included, a report must be the results of a randomized clinical trial studying an intervention and containing at least one control group including a randomization process. The trial must be registered on ClinicalTrials.gov or have an English-language record on one of the primary registries in the World Health Organization (WHO) registry network (https://www.who.int/ictrp/network/primary/en/).
5) To match the published report with the registry entry, we will first use the identifier identified in the paper or abstract. If no identifier is present but the registry of pre-registration is identified, we will use the first, final and corresponding authors' last name(s) to search the registry. If still not present, we will use the article title to search the identified registry to identify a corresponding record.
6) We will exclude abstracts without published papers, articles without accompanying abstracts, letters to the editor, editorials, observational studies, duplicate reports of a study, follow-up studies, meta-analyses, studies without a randomization process, studies without at least one control group, studies without an identified registry within the paper, studies with an identified registry but where the corresponding record could not be found. If the study identifies that results of the trial have previously been published elsewhere or are follow-up results from another trial, the study will be excluded.

Power and sample size calculation

No formal power calculation is required for the descriptive statistics planned. No formal statistical difference tests are planned *a priori*. We plan to analyze randomized controlled trials published between 2014 and 2016. The time period was selected to enable an assessment of a high portion of the likely high-impact, recently published literature, balanced with resource constraints.

Data abstraction
Included papers will be reviewed by two readers and data will be extracted from each independently using a standard form. Kappa scores will be calculated for the agreement between the two raters for each of the primary and three secondary aims. Disagreements will be resolved by consensus or by a third reader if consensus is not reached.

The data extraction form was created by RK and JW beforehand. It was tested on trials that were published in 2018 (and therefore ineligible for this study), with modifications made as needed.

Statistical testing
Descriptive statistics will be used to summarize the characteristics of the included and excluded studies. The primary and secondary outcomes of this observational study will be summarized using descriptive statistics. Kappa scores will be calculated for the agreement between the two raters for each of the primary and three secondary aims.

Study registration
This study is not eligible to be registered on ClinicalTrials.gov, EU-CTR, or PROSPERO as per the inclusion criteria available on these websites. To our knowledge, there is not currently a broadly adopted registry for meta-research studies. Therefore, we uploaded a time-stamped protocol to Github on March 27, 2019. Any changes to the protocol after uploading will be clearly identified in the final paper.

Primary aim in bold:
PO1: What is the percentage of papers with complete concordance in primary outcome registration and reporting between the registry, the paper and the abstract?

Explanation: This is a composite outcome, requiring that all of SO1, SO2, SO3 show no discrepancies. (This requires complete concordance between measurement tools/measures and timepoints and that all primary outcomes are clearly identified as primary. It also requires no primary outcomes be included in one location without being in the other two.)

Secondary outcomes
SO1: In what percentage of papers is there complete concordance between the primary outcome(s) specified on the registry and the primary outcome(s) identified in the paper?

Explanation:
Complete concordance is defined as:

(1) All primary outcomes prespecified on the registry are presented as primary outcomes in the results section of the paper
    (a) Consistent measurement tool(s)/measure(s) and time-point(s)

(b) In the results section of the paper, a result for the primary outcome is reported (e.g. as textual description of result, p-value, reporting of effect size, etc)

(2) All primary outcomes presented in the results section of the paper are prespecified on the registry
   (a) Consistent measurement tool/measure and time-point
   (b) Pre-specified under Primary Outcome heading on the registry

- Labeling of an outcome as “primary”: the primary outcome(s) must be clearly labeled as “primary,” in the intro, methods, or results section of the paper. If there is a discrepancy, the label in the results section will be used for comparisons with the registry. Alternative terms for “primary” such as “main,” “principal,” “central” or other unambiguous descriptions will be allowed. The descriptor “key” will not be considered an unambiguous identifier of an outcome given that it is often used to define important secondary endpoints.
- Defining the measurement tools/measures and time-points of the primary outcome(s): the primary outcome(s) must have clearly identified measurement tool(s)/measure(s) and timepoint(s), in the intro, methods, or results section of the paper. If there is a discrepancy, the measurement tool(s)/measure(s) and timepoint(s) used for the primary outcome(s) in the results section will be used for comparisons with the registry.
- A primary outcome must match in measurement tool/measure and time point between registry and paper, <u>unless this difference is explained or discussed in the paper.</u> The only difference between the registry and the paper in the measurement tool/measure that will be scored as concordant is if a general version of a measurement tool (i.e. HAM-D) is registered on the registry and a more specific version of the tool (i.e. HAM-D-17) that is in widespread use and measures the same overall construct is reported in the paper. Reporting in the paper of a subscale of the outcome measure pre-specified on the registry will be scored as not concordant.
- If the measurement tool/measure or time point is not specified in the paper or the registry, this counts as not concordant
- If time point is on the registry without any additional information (e.g. “6 weeks”), it will be scored as concordant in time point if the paper presents one of the following: (1) change from baseline to that time point, (2) condition-by-time interaction effect bound by the baseline and the specified time point, or (3) the absolute values at that time point.

- Specific additional questions that we hope to answer include:
  - Outcomes on the registry ending up in paper. Within our sample:
    - What is the proportion of all primary outcomes from the registry that are reported across all papers?
    - Out of all primary outcomes on the registry that are NOT reported (with full concordance in measurement tool/measure, time-point, and labeling) in the paper:

- What is the proportion in which the measurement tool/measure on the registry is not reported anywhere in the paper?
- What is the proportion in which the measurement tool/measure on the registry is reported in the paper, but the time-point is different?
- What is the proportion in which the primary outcome from the registry is reported in the paper with concordant measurement tool/measure and time-point, but it has been labeled as a secondary outcome in the paper?

- Outcomes in paper that were/were not on the registry. Within our sample:
  - What is the mean number of primary outcomes per paper that were not prespecified on the registry?
  - Out of all primary outcomes in the paper that were NOT prespecified (with full concordance in measurement tool/measure, time-point, and labeling) on the registry:
    - What is the proportion in which the measurement tool/measure in the paper is not prespecified anywhere on the registry?
    - What is the proportion in which the measurement tool/measure in the paper is prespecified on the registry, but the time-point is different?
    - What is the proportion in which the primary outcome was listed with concordant measurement tool/measure and time-point, but labeled as a secondary outcome on the registry?

SO2: In what percentage of abstracts is there complete concordance between the primary outcome(s) specified on the registry and the primary outcome(s) specified in the abstract?

Explanation:
Complete concordance is defined as:

(1) All primary outcomes prespecified on the registry are presented as primary outcomes in the abstract
   (a) Consistent measurement tool(s)/measure(s) and time-point(s)
   (b) Any reporting of result in abstract (e.g. as textual description of result, p-value, reporting of effect size, etc)
(2) All primary outcomes presented in the abstract are prespecified on the registry
   (a) Consistent measurement tool/measure and time-point
   (b) Pre-specified under Primary Outcome heading on the registry

- Labeling of an outcome as “primary”: the primary outcome(s) must be clearly labeled as “primary” within the abstract. Alternative terms for “primary” such as “main,” “principal,” “central” or other unambiguous descriptions will be allowed. The descriptor “key” will not be considered an unambiguous identifier of an outcome given that it is often used to define important secondary endpoints.

- Defining the measurement tool(s)/measure(s) and time-points of the primary outcome(s): the primary outcome(s) must have clearly identified measurement tool(s)/measure(s) and time-point(s).
- A primary outcome must match in measurement tool/measure and time point between abstract and registry, unless this difference is explained or discussed in the paper. The only difference between the registry and the abstract in the measurement tool/measure that will be scored as concordant is if a general version of a measurement tool (e.g. HAM-D) is registered on the registry and a more specific version of the tool (e.g. HAM-D-17) that is in widespread use and measures the same overall construct is reported in the abstract. Reporting in the abstract of a subscale of the outcome measure pre-specified on the registry will be scored as not concordant.
- If the measurement tool/measure or time point is not specified in the abstract or the registry, this counts as not concordant
- If time point is on the registry without any additional information (e.g. “6 weeks”), it will be scored as concordant in time point if the paper presents one of the following: (1) change from baseline to that time point, (2) condition-by-time interaction effect bound by the baseline and the specified time point, or (3) the absolute values at that time point.

- Specific additional questions that we hope to answer include:
    - Outcomes on registry ending up in abstract. Within our sample:
        - What is the proportion of all primary outcomes from registries that are included across all abstracts?
    - Outcomes in abstract that were/were not on the registry
        - What is the mean number of primary outcomes per abstract that were not prespecified on the registry?

SO3: In what percentage of papers is there complete concordance between the primary outcome(s) in the paper and the primary outcome(s) in the abstract?

Explanation:
Complete concordance is defined as:

(1) All primary outcomes in paper are presented as primary outcomes in the abstract
    (a) Consistent measurement tool/measure and time-point
    (b) Any reporting of result (e.g. as textual description of result, p-value, reporting of effect size, etc)
(2) All primary outcomes presented in the abstract are also primary outcomes in the results section of the paper
    (a) Consistent measurement tool/measure and time-point

- Labeling of an outcome as “primary”: the primary outcome(s) must be clearly labeled as “primary,” in the intro, methods, or results section of the paper. If there is a discrepancy, the label in the results section will be used for comparisons. It also must be clearly labeled as “primary” anywhere in the abstract. Alternative terms for “primary” such as

"main," "principal," "central" or other unambiguous descriptions will be allowed. The descriptor "key" will not be considered an unambiguous identifier of an outcome given that it is often used to define important secondary endpoints.
- Defining the measurement tool/measure and time-points of the primary outcome(s): the primary outcome(s) must have clearly identified measurement tool(s)/measure(s) and timepoint(s), in the intro, methods, or results section of the paper. If there is a discrepancy, the measurement tool(s)/measure(s) and timepoint(s) used for the primary outcome(s) in the results section will be used for comparisons with the abstract.
- A primary outcome must match in measurement tool/measure and time point between abstract and paper.
- If the measurement tool/measure or time point is not specified in the paper or the abstract, this counts as not concordant. If the measurement tool is not specified in either the paper or the abstract, this counts as not concordant.

- Specific additional questions that we hope to answer include:
    - Within our sample, out of all primary outcomes in papers, what proportion are included in abstracts?
    - What is the percentage of primary outcomes in abstracts that do have concordant measurement tools/measures and time-points in the papers, but that are identified as secondary outcomes in the papers?

We will also include descriptive measures (counts, means, proportions) summarizing the ways in which discrepancies occur in primary outcomes between the paper, abstract and the registry, including information about journal, publication year, affiliation with industry, diagnostic category (by DSM-IV chapter), and intervention type.

Other questions we plan to answer

- What is the total percentage of statistically significant primary outcomes in papers that are also included in abstracts?

- What is the total percentage of not-statistically significant primary outcomes in papers that are also included in abstracts?

  Explanation: Each outcome will be scored as statistically significant or not statistically significant based on the statistics presented in the paper.

- Out of all the primary outcomes in papers that were not prespecified on the registry (either because they were upgraded from a secondary outcome, not listed as primary, or not concordant with any outcome on the registry), what percentage are statistically significant?

- Out of all the primary outcomes on the registry that were changed to secondary outcomes in papers (i.e. demoted), what percentage are statistically significant?

Explanation: Secondary outcomes in the paper that have the same measurement tool/measure and time point as primary outcomes identified on the registry will be identified and classified as statistically significant or not statistically significant.

- What are the percentages of complete concordance (among SO1, SO2, SO3) among industry-affiliated and non-industry-affiliated trials?

  Explanation: Industry-affiliated studies will be determined based on the sponsor and collaborator listed on the registry. Any listed industry sponsor or collaborator will classify the trial as industry-affiliated.