Report on DIVERSE Task 1a4 (Screening and selection)

Study Reference: DIVERSE

Protocol title: DIVERSE project: protocol for the scoping review

Protocol version: 1.0

File: Seafile

Status: Final

Version: 1.11

Written by: Rosa Gini. Romin Pajouheshnia

Members of Task 1a4: Rosa Gini. Romin Pajouheshnia, Gillian Hall, Angelika Wientzek-Fleischmann, Dimitri Bennett, Mehmet Burcu, Caitlin Dodd, Edward_Lai, Helga Gardarsdottir, Giuseppe Roberto, Karine Marinier, Lin Li, Robert.Platt, Sigal Kaplan, Stephen Lanes, Soko Setoguchi Iwata, Xiaofeng.Zhou, Dana Teltsch, Carlos Duran Salinas, Kanaka Soman, Marloes Bazelier

Date: 15 Jul 2022

Contents

Aim

Before this task began, Task 1a1 had produced a draft selection tool and used it to select 24 'core' documents, and to label each of them with reasons for inclusion, and Task 1a3 extracted 687 journal articles and other documents based on a search strategy developed in Task 1a2. All such activities are reported separately.

The aims of Task 1a4 were to:

1) Refine the set of inclusion/exclusion criteria developed in Task 1a1, to be used to screen the 687 documents identified in 1a3;

2) Conduct screening and selection of the 687 documents for inclusion in the review and for data extraction.

Details of the overall background and aims of the project are given in the DIVERSE Seafile: DIVERSE_manuscript / Task folders / Objective1 / Task_1_0 / protocol.

Methods

The 24 core documents were automatically included in the study. The other 687 documents extracted in Task 1a3 underwent selection.

The screening tool produced by Task 1a1 was first piloted on 34 documents by 6 participants who worked in pairs (RG and MEB; RP and MAB; CD and SK). Each participant independently assessed a set of documents and then met with his/her pair to reconcile disagreements and produce written recommendations to improve the screening tool. Finally, all the participants to the pilot met and the screening tool was finalised.

Pilot using Titles and Abstracts, and full text selection

During the pilot the screening was performed on a sequence on title and abstract, and those passing this screening (15) were subsequently screened by the same assessors on the full text of the documents. During the pilot, agreement was sought among assessors.

After the pilot, the screening was divided in two phases

- Title and abstract (TIAB)
- Full text (FT)

Selection using Titles and Abstracts

TIAB selection was conducted in 3 phases

- During the first phase, the procedure to screen was piloted: 2 participants (RG and RP) assessed a first group of 100 documents and met to reconcile disagreements; this was then judged to be inefficient
- Second phase: 7 participants (CAD, CD, GH, LL, MAB, MEB, SK) assessed 509 of the documents, each document was assessed independently by two assessors and

whenever there was disagreement, reconciliation was performed by a third assessor (KM, RP, SL)

- Third phase: 2 assessors (GH and RG) assessed the 44 documents included from the last months of 2021, and whenever there was disagreement, reconciliation was performed by RP.

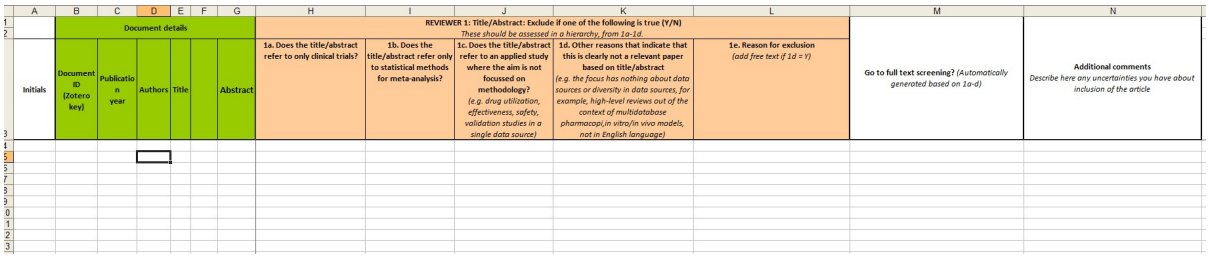
Selection using Full text

The papers filtered after TIAB were divided in 17 sets. Each set was screened by a pair of assessors, 17 assessors in total (AW CAD CD EL GH GR KM LL MAB MEB RG ROBP RP SK SL SS XZ). A group of 3 participants (GH, RG, RP) acted as third assessors for the papers where disagreement was observed.

Results

Final selection tool- TIAB

The final TIAB selection tool is shown in Figure 1 and is embedded in a spreadsheet.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1		Document details						REVIEWER 1: Title/Abstract: Exclude if one of the following is true (Y/N) These should be assessed in a hierarchy, from 1a-1d.						
2	Initials	Document ID (Zotero key)	Publication year	Authors	Title		Abstract	1a. Does the title/abstract refer to only clinical trials?	1b. Does the title/abstract refer only to statistical methods for meta-analysis?	1c. Does the title/abstract refer to an applied study where the aim is not focussed on methodology? (e.g. drug utilization, effectiveness, safety, validation studies in a single data source)	1d. Other reasons that indicate that this is clearly not a relevant paper based on title/abstract (e.g. the focus has nothing about data sources or diversity in data sources, for example, high-level reviews out of the context of multidatabase pharmacopi,in vitro/in vivo models, not in English language)	1e. Reason for exclusion (add free text if 1d = Y)	Go to full text screening? (Automatically generated based on 1a-d)	Additional comments Describe here any uncertainties you have about inclusion of the article
3														

Figure 1. Selection tool for title and abstract (TIAB) selection

The tool included four questions that were designed as exclusion criteria, based on the study protocol.

1a. Does the title/abstract refer to only clinical trials? *(Y/N)*

1b. Does the title/abstract refer only to statistical methods for meta-analysis? *(Y/N)*

1c. Does the title/abstract refer to an applied study where the aim is not focused on methodology? (e.g. drug utilization, effectiveness, safety, validation studies in a single data source) *(Y/N)*

1d. Other reasons that indicate that this is clearly not a relevant paper based on title/abstract

(e.g. the focus has nothing about data sources or diversity in data sources, for example, high-level reviews out of the context of multidatabase pharmacopi, in vitro in vivo models, not in English language)

If the reason for exclusion was based on the last exclusion criterion (1d), participants were requested to describe the reason for exclusion in a free text box.

If none of the exclusion criteria was selected, and no reason for exclusion was found, the document was proposed to proceed to full text selection, and a column for additional free text comments was available

Final selection tool- full text

The final selection tool is shown in Figure 2 and is embedded in a spreadsheet.

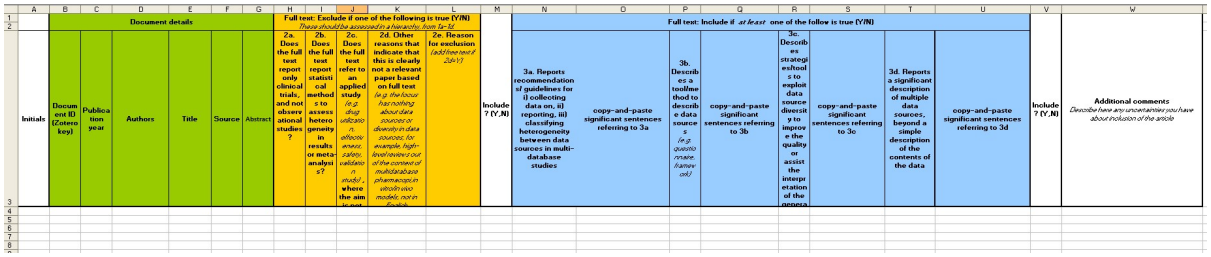


Figure 2. Selection tool for the full text selection

The tool included four questions that were designed as exclusion criteria, and four questions that were designed as inclusion criteria.

Exclusion criteria

2a. Does the full text report only clinical trials, and not observational studies? *(Y/N)*

2b. Does the full text report statistical methods to assess heterogeneity in results or meta-analysis? *(Y/N)*

2c. Does the full text refer to an applied study (e.g. drug utilization, effectiveness, safety, validation study), where the aim is not focused on methodology? *(Y/N)*

2d. Other reasons that indicate that this is clearly not a relevant paper based on full text (e.g. the focus has nothing about data sources or diversity in data sources, for example, high-level reviews out of the context of multidatabase pharmacopi, in vitro/in vivo models, not in English language). *(Y/N)*

2e. Reason for exclusion (add free text if 2d=Y)

Inclusion criteria (include text from the document or as a comment to support each criterion that is met)

3a. Reports recommendations/ guidelines for i) collecting data on, ii) reporting, iii) classifying heterogeneity between data sources in multi-database studies

3b. Describes a tool/method to describe data sources (e.g. questionnaire, framework)

3c. Describes strategies/tools to exploit data source diversity to improve the quality or assist the interpretation of the generated evidence

3d. Reports a significant description of multiple data sources, beyond a simple description of the contents of the data

Final comments

A free text field was available for additional comments.

Algorithm

Assessors were requested to

1. fill out the exclusion criteria hierarchically: if one criterion was met, the document was excluded and the assessment ended.
2. for those documents that were not excluded: fill out all the inclusion criteria; the document was included if at least one criterion was met, and excluded otherwise; for each criterion that was met, the assessor was invited to include supporting text from the document or their own comment, to support data extraction in the next phase.

Selection using Titles and Abstracts

The pilot was conducted on 34 documents, and 15 (44.1%) were included for full text screening (the final results of the pilot are reported in the next section 'Selection using full text').

The first phase included 100 documents, 30 (30.0%) were included for full text screening.

The second and third phases included, respectively, 509 and 44 documents. In those phases, reconciliation was not attempted by the two independent assessors, and agreement was as follows

```
. tab M_1 M_2 if  !mi(M_2) & origin > 2

                      |    Go to full text
     Go to full text  |       screening
           screening  |         N          Y |     Total
----------------------+----------------------+----------
                   N |       373         51 |       424
                   Y |        45         83 |       128
----------------------+----------------------+----------
               Total |       418        134 |       552

. kap M_1_num M_2_num if !mi(M_2)         & origin > 2

             Expected
Agreement    Agreement     Kappa   Std. Err.         Z      Prob>Z
-------------------------------------------------------------------
  82.61%       63.79%     0.5197     0.0425      12.21      0.0000
```

The kappa statistic was 52.0%

As a result, 96 documents (89, 17.5%, in the second phase; and 7, 15.9%, in the third phase) were re-assessed by a third assessor. Out of 96 documents, 57 (59.4%) were included for full text selection.

In summary, the inclusion for full text selection was as follows

Mcomplete	origin				Total
	pilot	first_gro	second_gr	third_gro	
N	19	70	378	34	501
	55.88	70.00	74.26	77.27	72.93
Y	15	30	131	10	186
	44.12	30.00	25.74	22.73	27.07
Total	34	100	509	44	687
	100.00	100.00	100.00	100.00	100.00

The reasons for exclusion are illustrated in the table below.

Reason for exclusion at TIAB (core papers not considered)	to_be_reassessed_complete		
	0	1	Total
clinical trial	26	0	26
	4.40	0.00	3.78
statistical methods	34	8	42
	5.75	8.33	6.11
study with no methodo	311	20	331
	52.62	20.83	48.18
other	91	11	102
	15.40	11.46	14.85
admitted to full text	129	57	186
	21.83	59.38	27.07
Total	591	96	687
	100.00	100.00	100.00

.

Out of 686 documents, 26 (3.8%) were excluded because they referred to clinical trials; 42 (6.1%) because they were focussed on statistical methods to address heterogeneity; 331 (48.2%) because they were studies with no methodological focus pertinent to the DIVERSE scoping review; and 102 (14.9%) for other reasons.

In summary, out of 687 documents, 186 (27.1%) were included for the full text selection.

Selection using Full text

Pilot and assessment were distributed as follows

```
. tab group_SFT if var_08_included_to_SFT == "Y" & var_07_core_paper == 0
```

group_SFT	Freq.	Percent	Cum.
Full text selection group 0	10	5.38	5.38
Full text selection group 1	11	5.91	11.29
Full text selection group 2	11	5.91	17.20
Full text selection group 3	11	5.91	23.12
Full text selection group 4	10	5.38	28.49
Full text selection group 5	9	4.84	33.33
Full text selection group 6	8	4.30	37.63
Full text selection group 7	9	4.84	42.47
Full text selection group 8	9	4.84	47.31
Full text selection group 9	10	5.38	52.69
Full text selection group 10	11	5.91	58.60
Full text selection group 11	11	5.91	64.52
Full text selection group 12	11	5.91	70.43
Full text selection group 13	10	5.38	75.81
Full text selection group 14	10	5.38	81.18
Full text selection group 15	10	5.38	86.56
Full text selection group 16	10	5.38	91.94
Pre-pilot	1	0.54	92.47
Pilot group 1	5	2.69	95.16
Pilot group 2	6	3.23	98.39
Pilot group 3	3	1.61	100.00
Total	186	100.00	

Out of the 15 papers that were assessed during the pilot phase, **4** were included in the scoping review (25.9%).

Across 171 papers that were evaluated by the selection groups, where agreement was not sought between the two assessors, agreement resulted a posteriori as follows

```
. tab SFTincl_sum_1 SFTincl_sum_2 if var_08_included_to_SFT == "Y" & group_SFT < 990
```

SFTincl_sum_1 \ SFTincl_sum_2	N	Y	Total
N	73	28	101
Y	25	45	70
Total	98	73	171

that is, **45** papers (26.3% of 171) were included in the scoping review by both assessors, 73 papers (41.5% of 171) were excluded, and 53 papers (31.0% of 171) were evaluated by a third assessor.

The kappa statistics was 36.3%;

```
. kap SFTincl_sum_1_num SFTincl_sum_2_num if !mi(SFTincl_sum_2_num) & var_08_included_to_SFT == "Y" & group_SFT < 990

             Expected
Agreement    Agreement      Kappa    Std. Err.         Z      Prob>Z
------------------------------------------------------------------
  69.01%       51.33%     0.3632       0.0764       4.75       0.0000
```

Out of the 53 papers evaluated by a third assessor, **18** (34.0%) were included in the scoping review.

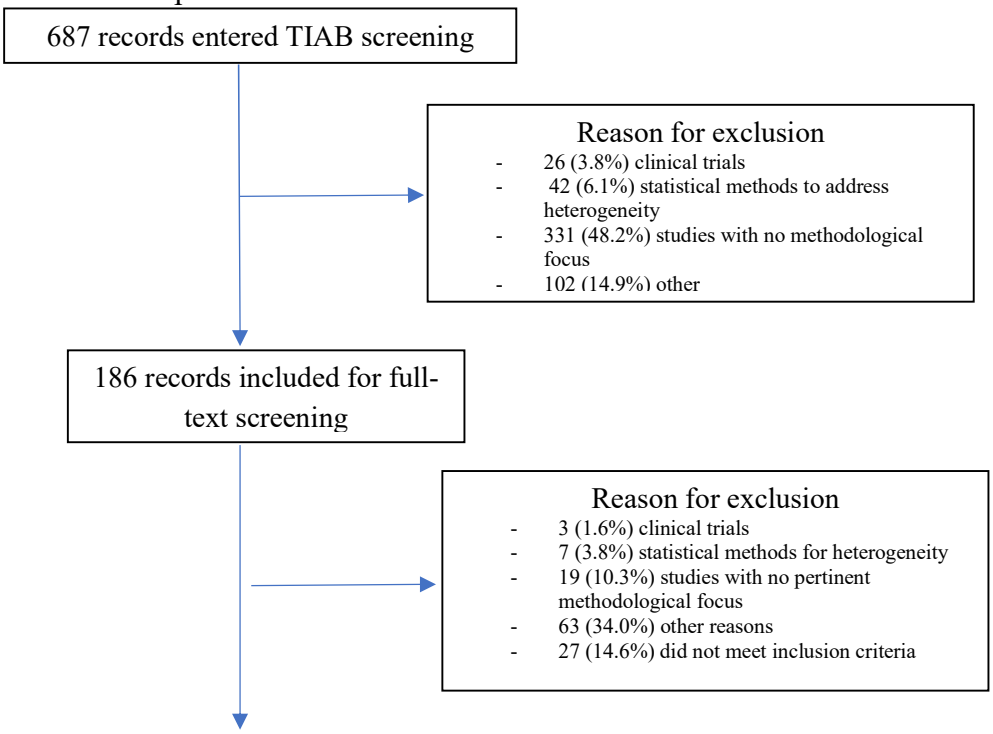
Reasons for exclusion across the 186 documents included to full text review were distributed as follows: 3 documents (1.6%) were clinical trials, 7 (3.18%) reported on statistical methods for addressing heterogeneity, 19 (10.2%) were studies with no pertinent methodology focus, 63 (33.9%) were excluded for other reasons, and 27 (14.5%) did not meet any inclusion criteria.

```
. tab SFTreason_exclusion
```

Reason for exclusion at full text selection(core papers and excluded after TIAB	Freq.	Percent	Cum.
clinical trial	3	1.61	1.61
statistical methods	7	3.76	5.38
study with no methodology focus	19	10.22	15.59
other reasons for exclusion	63	33.87	49.46
no inclusion criterion met	27	14.52	63.98
included in the scoping review	67	36.02	100.00
Total	186	100.00	

Overall, 67 papers were included after full text revision. Including the 24 core papers, we were left with a total of 91 documents included in the scoping review.

The flowchart below summarizes the process of inclusion.



67 included in the scoping review

(and additionally: 24 core papers) - total **91**

Reason for inclusion

The 91 documents included in the scoping review met one or more inclusion criteria described below:

3a. Reports recommendations/ guidelines for i) collecting data on, ii) reporting, iii) classifying heterogeneity between data sources in multi-database studies: 37 (40.7%);

3b. Describes a tool/method to describe data sources (e.g. questionnaire, framework): 53 (58.2.1%);

3c. Describes strategies/tools to exploit data source diversity to improve the quality or assist the interpretation of the generated evidence: 59 (64.8%);

3d. Reports a significant description of multiple data sources, beyond a simple description of the contents of the data: 51 (56.0%).

If we aggregate the reasons for inclusion 3a, 3b and 3d, which overall address diversity description, we find 75 documents (82.4%).

Summary

Out of 687 documents, 186 (27.1%) were included for full text screening and 67 (9.5%) were selected and appropriate for the scoping review. Agreement between assessors was good in the title and abstract phase, less so in the full text selection phase, which is not surprising, due to the exploratory nature of a scoping review. The judgement of a third assessor provided confidence that most relevant documents were included.

After adding the 24 core papers, 91 documents entered the scoping review.