**Instructions: data extraction from epidemiology articles**

**Time of writing**

* Data extraction has not yet begun for any eligible article
* The first draft of the data extraction form has been piloted on one observational epidemiology study by MG, KD and RR (Dregan (2018) Arterial stiffness association with chronic inflammatory disorders in the UK Biobank study).

**Procedure**

The following procedure will be conducted for all articles classified as ‘observational epidemiology’

1. Two people (KD, RR, MG or BW) will independently identify what study designs (cohort, case-control or cross-sectional) all 178 articles contain by answering the first block of questions in the data extraction form. Any conflicts will be discussed by the two extractors and resolved by mutual consent if possible. If not resolved after discussion RR or MM will determine the study designs
2. If you articles contain multiple study designs they will confirm whether the article should be assessed STROBE guidelines at the article-level or study-level
3. One person (KD) will identify and exclude retracted articles by assessing the notices attached to the online version of the article.
4. One person (KD) will identify any articles with supplementary material which cannot be accessed via the University of Bristol subscriptions. We will attempt to retrieve any supplementary material we cannot access (e.g. by contacting the authors).
5. The items in Table 1 will be extracted manually by one person (KD) or automatically by a bibliographic manager.
6. The articles will be randomly selected for extraction and two people (MG, KD, BW or RR) will independently complete the data extraction form for each article. Any conflicts will be discussed by the two extractors and resolved by mutual consent if possible. If not resolved after discussion RR or MM will assign the category. Articles will continue to be selected at random until none remain, or until the authors resources are exhausted.

*Table 1. Methods for extracting article’s meta-data*

| Item | Definition | Method of extraction |
| --- | --- | --- |
| Article ID(s) of related articles | Article IDs of articles that contain similar analyses. | Examine articles with same first author, last author and UK Biobank application number. |
| Citation |  | Bibliographic |
| Publication date |  | Bibliographic |
| Authors |  |  |
| Journal |  |  |
| Volume |  |  |
| Issue |  |  |
| Open Access |  | (1) search for article on open access button (<https://openaccessbutton.org/>) and follow the link on non-university network  (2) if no on open access button check if there are links on google scholar |
| Countries of corresponding, first & last author |  | Extract country of author’s affiliations |
| UK Biobank Application number | Study’s approved UK Biobank data application number (if present) | (1) search article for all these terms:  ‘application’ , ‘approv’ , ‘number’ , ‘Resource’ (2) if not found examine any strings of numbers in article |
| Author keywords |  | (1)extract from page 1 (2) if not on page 1 search article for ‘keywords’ |
| Journal endorsement | Journal’s endorsement of STROBE before publication | (1) extract references to STROBE from the journal’s most recent ‘instructions/guide for authors’ before article’s publication date |
| Journal impact factor |  | Retrieve the impact factor of the journal for year of publication from Web of Science’s Journal Citation Reports (JCR) |
| Corresponding author address |  | Bibliographic (if missing (1)extract from page 1 (2) if not on page 1 search article for ‘corresponding’, ‘@’) |
| Conflicts of interest statement |  | (1) search article material for all these terms: ‘competing’, ‘interest’ ,‘conflict’ , ‘disclosure’ (2) if not found examine article for similar terms (3) if present extract conflicts of interest statements |
| Classification | Classification (MR, epi or other) as assigned by KD, BW and RR in Rayyan | Rayyan |

**Missing items**

We will not attempt to retrieve any missing items in the data extraction form, this is to give an accurate assessment of the version of record and to avoid exhausting the extractors resources.