This is a provisional PDF only. Copyedited and fully formatted version will be made available soon.

How to write an Umbrella Review? A step-by-step tutorial with tips and tricks

Authors: Michał Bonczar, Patryk Ostrowski, Anthony V. D'Antoni, Richard Shane Tubbs, Joe Iwanaga, Sanjib Kumar Ghosh, Ilona Klejbor, Marcin Kuniewicz, Jerzy Walocha, Janusz Moryś, Mateusz Koziej

How to write an Umbrella Review? A step-by-step tutorial with tips and tricks

Michał Bonczar et al., How to write an Umbrella Review? A step-by-step tutorial with tips and tricks

Michał Bonczar[1]*, Patryk Ostrowski[1]*, Anthony V. D'Antoni[2], Richard Shane Tubbs[3, 4], Joe Iwanaga[3, 4], Sanjib Kumar Ghosh[5], Ilona Klejbor[6], Marcin Kuniewicz[1], Jerzy Walocha[1], Janusz Moryś[7], Mateusz Koziej[1]

[1]Department of Anatomy, Jagiellonian University Medical College, Krakow, Poland

[2]Division of Anatomy, Department of Radiology, Weill Cornell Medicine, New York, New York, United States

[3]Department of Neurosurgery, Tulane Center for Clinical Neurosciences, Tulane University School of Medicine, New Orleans, Louisiana, United States

[4]Department of Neurology, Tulane Center for Clinical Neurosciences, Tulane University School of Medicine, New Orleans, LA, United States

[5]Department of Anatomy, All India Institute of Medical Sciences, Phulwarisharif, Patna, India

[6]Department of Anatomy, Collegium Medicum, Jan Kochanowski University of Kielce, Poland

[7]Department of Normal Anatomy, Pomeranian Medical University in Szczecin, Poland

Address for correspondence: Mateusz Koziej, Department of Anatomy, Jagiellonian University Medical College, ul. Mikołaja Kopernika 12, 33-332 Kraków, Poland, e-mail: mateusz.koziej@gmail.com

*Equal Contribution

Abstract

The number of meta-analyses (MA) and systematic reviews (SR) on various medical issues has increased during the last two decades. The MA and SR results may differ from one another due to a number of factors such as inaccurate or diverse searches through the

databases, discrepancies in the extraction process or in statistical analysis, among others. Some results may even contradict one another, resulting in confusion among readers. Umbrella reviews (UR) have allowed the collection of all available data on a medical issue into one concise study, making it the source of evidence-based medical knowledge to the highest degree. Furthermore, UR can resolve those problems by collecting all data and taking into account both MA and SR, making it the superior tool for physicians. Although the pros of UR are clear and the overall popularity of these types of study has increased tremendously, there is no available step-by-step guide on how to conduct one. Therefore, the objective of the present study was to provide researchers with a detailed tutorial on how to conduct an UR. UR represent the next major step in the advancement of evidence-based medicine, with great practical potential for physicians looking for the most up-to-date data on their topic of interest. We hope that our step-by-step guide may be a useful tool for researchers conducting UR in the future.

Key words: umbrella review, evidence based medicine, EBM, meta-analysis, systematic review

INTRODUCTION

The number of meta-analyses (MA) and systematic reviews (SR) on various medical issues has increased during the last two decades. These types of studies were the best sources of up-to-date and evidence-based information on certain topics for medical professionals. The MA allowed us to make sense of numerous and often contrasting findings about different medical issues, pushing our knowledge forward. However, with the increased number of MA and SR, medical professionals can feel overwhelmed when looking for evidence-based knowledge on their topic of interest. Gathering more precise and significant statistical conclusions is the basis of scientific development.

Umbrella reviews (UR) have allowed the collection of all available data on a medical issue into one concise study, making it the source of evidence-based medical knowledge to the highest degree (Figure 1). They are reviews based on previous MA and SR, presenting all available data on a specific topic in the literature. Therefore, the popularity of UR has increased significantly in recent years. The difference in methodology between MA and SR can present different conclusions about the same topic, which may give an incorrect clinical picture of the issue at hand. The MA and SR results may differ from one another due to a number of factors such as inaccurate or diverse searches through the databases, discrepancies

in the extraction process or in statistical analysis, among others. Some results may even contradict one another, resulting in confusion among readers. UR can resolve those problems by collecting all data and taking into account both MA and SR, making it the superior tool for physicians. Although the pros of UR are clear and the overall popularity of these types of study has increased tremendously, there is no available step-by-step guide on how to conduct one. Therefore, the objective of the present study was to provide researchers with a detailed tutorial on how to conduct an UR. Lastly, we will also discuss the potential limitations associated with UR.

HOW TO PREPARE AN UMBRELLA REVIEW?
Planning and establishing goals

The initial step of creating an UR, and probably the most crucial one, is a precise and accurate establishment of the aim of the study. Additionally, it is essential to understand the sense, importance, and overall assumptions of the UR. The goal of UR is to gather and summarize the available data in the literature and/or to standardize inaccuracies. The MA and SR results may differ from one another due to a number of factors such as inaccurate or diverse searches through the databases, discrepancies in the extraction process or in statistical analysis, among others. Additionally, some MA and SR may be related to the same topics from different points of view. Therefore, UR are created in order to gather, summarize, and/or idealistically standardize the results and dispel the doubts in the literature. Additionally, UR minimize the time that is needed to be spent on the recognition in the literature by the physician, which is purely beneficial for the clinical practice. Therefore, by creating the UR it is aimed to gather all MA and SR regarding the studied topic. At the beginning, we highly recommend doing an initial recognition in the literature, especially with respect to the overall number and availability of SR and MA for the proposed topic. There are no contradictions for the use of a search filter, especially the one regarding the type of study, as it is only the first, initial search, only to recognize whether the topic is suitable for further investigation. It is critical to evaluate whether the found SR and MA concern the same or at least related topics. Additionally, cohesion of the methodologies among the said studies and mutually exclusive results increase the relevance and need for such UR. This step is essential, as the creation of an UR is only possible if there are enough appropriate studies and data in the literature on a specific research issue. Therefore, initial estimates of the number of SR and MA on the topic studied allow evaluation of the risk of incomplete, unsuccessful, or impossible research.

Systematic search

After the initial evaluation of the topic, it is time to perform a systematic search through the databases. Therefore, the search phrases have to be established, preferably with the usage of the Boolean technique. The Boolean technique is based on mathematical logic. Between each of the phrases we can add one of three conjunctions: 'AND', 'OR' and 'NOR'. Using the conjunction 'AND' will result in the appearance of only studies that refer to all those phrases simultaneously. Using the conjunction 'OR' will result in the appearance of studies that are related to at least one of the phrases. Using the 'NOR' conjunction will result in the appearance of all studies related to the initial phrase (the one used before the 'NOR' conjunction) but with the omission of all studies related to the other phrase (the one used after the 'NOR' conjunction). It is possible to use brackets to create multilevel, complicated search phrases (example: (((Phrase1 OR Phrase2 OR Phrase3 OR Phrase 4) AND (Phrase5 OR Phrase6)) NOR (Phrase6 OR Phrase7)). For the UR, we recommend for the search phrases to be as wide as possible in order to minimize the chances of losing any relevant study. We recommend applying as many synonyms for each phrase when using the 'OR' conjunction. In addition, it is beneficial to add as many abbreviations as possible commonly used in clinical practice or the literature on the studied topic. It is also beneficial to use all historical eponyms in the researched issue. The given topic might have been studied for a long time before modern nomenclature. We strongly advise against using the 'NOR' conjunction, as it greatly increases the chances of an unsuccessful search or missing some of the potentially relevant studies. We advise against applying phrases that refer to 'meta-analysis' or 'systematic review' (example: (Phrase1 OR Phrase2) AND (Meta OR Meta-analysis OR Systematic OR Systematic Review)). Additionally, it is possible to use quotation marks (example: 'Phrase1' OR Phrase2) in order to search only for the particular word used as a phrase, without its varieties; however, we strongly advise against using this function in systematic searches. Graphical illustration of the Boolean conjunctions and their influence on the searched phrases is presented on Figure 2.

After the establishment of the search phrases, the search can be conducted through the databases. We recommend searching through as many databases as possible, especially the main ones such as PubMed, Embase, Web of Science, Scopus, Google Scholar and Cochrane Library, as it maximizes the chances of gathering all relevant studies. In some of the databases, the search phrases will need to be adjusted for the specific search coding systems used on those sites. We strongly advise against the usage of filters at this stage of the process. The aim of this part of the study is to collect all MA and SR regarding the topic researched. It

is acceptable to create UR regarding only MA, with omission of the SR; however, we recommend gathering both types of studies. During the search, determining the relevance and reference to the topic of the study should be done in the first step by evaluating the title and abstract and by reading the article in the second step. Studies in foreign languages should be evaluated only by reviewers who speak fluently the said language. Clearly, irrelevant research should not be screened for further evaluation. It is also recommended for the search to be done twice by two independent reviewers. Afterwards at the final step of the search, we recommend performing a manual search through the references of the submitted studies to find more research which may not have been found by the database's algorithms.

Furthermore, all submitted studies should be evaluated. For this purpose, inclusion and exclusion criteria should be established. This should be the final point of clarification of the detailed aim and topic of the study, as the submitted studies must have results in at least one identical category so they can be compared with each other. We recommend excluding studies that are burdened with strong bias. There are not many tools created for the evaluation of bias and the evaluation of quality in MA and SR. However, there are many detailed and specific tools for the evaluation and assessment of primary studies that are used in the process of conducting MA and SR. We recommend the MeaSurement Tool to Assess Systematic Reviews 2 (AMSTAR 2) [6], the ROBIS tool [7] for the quality assessment. It is also beneficial to use more dedicated tools, such as AQUA or CATAM, for anatomical MA [1, 2]. We recommend using these tools, as there is a high similarity between MA, SR, and UR in methodology. Irrelevant, biased, low-quality, and duplicated research should be deleted.

Double extraction and statistical dilemma

Later, after the systematic search has been completed, the extraction process must be performed. During this part, it is aimed to extract and gather all possible data, both quantitative and qualitative, from the qualified studies. It is crucial to perform this part with great care and precision. We strongly suggest performing the extraction twice by two independent researchers. During this part of the process of creating MA and SR, the authors gather all extractable data from the primary studies. The UR is a relatively new type of review study, and its statistical results are based on the results of the MA, which are based on the results of the primary studies. It comes down to the statistical dilemma of whether it is acceptable to establish new results taking into consideration only the results of the MA. During our experience, we agreed that statistical analysis should be done based on the results of the primary studies and only compared to the results of the MA. The results of MA results

may differ from one another due to different searches within the different sets of databases. However, some of the studies may have been included in more than one MA. Therefore, we strongly suggest performing a double extraction to prevent multiplicated use of the results from the primary studies in the new results. All discrepancies need to be discussed and resolved to create a one-consistent version. The first part of the double extraction should be performed on the MA and SR. All qualitative and quantifiable data should be collected. Subsequently, the data for statistical analysis should be extracted from the primary studies. In order to perform this part, the authors must evaluate the studies on which each of the MA performed its statistics. We recommend searching through the reference lists of the MA as all primary studies should have been reported there. It is essential to extract data from primary studies to prevent multiplication of the results. Data from MA and UR should be gathered in separate tables in order to avoid unnecessary confusion. The aim should be to collect all relevant data from qualified studies. Qualitative results, for example, the year when the study was conducted, the continent and country in which the study has been performed, studied groups, the methodology of the study (for example: cadaveric study, intraoperative observations, study based on the computed tomography results, comparison of the surgical techniques, comparison of the groups using different methods, etc.) and more. Quantitative results are all numeric data obtained during the study by the researchers. These are strongly varied depending on the purpose of the said study. However, all quantitative results should be presented using mathematical tools like means, medians, risk ratios, odds ratios, pooled prevalences, percentages, standard deviations, etc. It is essential for those data to be properly interpreted and understood by the researcher so they can be grouped and qualified as analogical ones. Incorrect extraction leads to incorrect statistical analysis, resulting in false conclusions.

Statistical analysis

Statistical analysis of the UR should be performed based on the results of the primary studies instead of the results of the MA. It is crucial to properly collect and group the extracted data and establish an accurate method of analysis. For UR regarding the results from SR in which MA were not performed, we recommend using the Summary of Finding table to present the data in clear format [3, 5].

At the beginning of the statistical analysis of the UR, we recommend evaluating the types of outcomes in the submitted MA. From a mathematical point of view, we can output data in numerous ways. However, we recommend presenting the newly established UR

results in the same way that the authors of the MA presented the analogous results in their study. This will allow us to compare the new results with those from MA and potentially resolve the contradictions occurring in the literature. Additionally, it will minimize the potential confusion associated with numerous mathematical and statistical methods in which the same topic is discussed. We strongly recommend providing possibly the most accurate results but, simultaneously, presenting them in a simple way. This will allow clinicians to understand the mathematics behind the conclusions. Establishing the way in which it is preferred to present the results allows us to properly choose statistical tools.

Authors creating UR will most commonly find three types of outcomes in MA: (1) comparisons of risks between cohorts, presented as risk ratios, odds ratios, hazard ratios, etc., (2) proportions or prevalences, and (3) means. Independently, heterogeneity among the gathered data has to be assessed, preferentially using both the Chi-squared test and the I-squared statistic. For the Chi-squared test, the p-value of Cochran's Q should be reported, with a value of < 0.10 considered to indicate statistically significant heterogeneity between studies [3, 4]. The I-squared statistic, an overall measure of heterogeneity, should be reported with its 95% CI. The I-squared statistic should be interpreted as follows: 0% to 40% might not be important; 30% to 60% may represent moderate heterogeneity; 50% to 90% may represent substantial heterogeneity; and 75% to 100% may represent considerable heterogeneity, according to the guidelines in the current version of the Cochrane Handbook of Systematic Reviews of Interventions [3, 4]. This will allow authors to choose either fixed- or random-effects models. Mostly, the data will be heterogeneous, therefore the random-effects model will be used much more commonly rather than the fixed-effects model. Unlike a fixed-effects model which assumes that the differences between the results of studies are due solely to chance, a random-effects model assumes that the effects being estimated in the different studies are not identical [3, 4].

To establish new risks (1) we recommend using the Comprehensive Meta-analysis 3.0 (Biostat Inc., Englewood, NJ, USA) tool. It is a highly efficient, accurate and user-friendly tool with numerous tutorials and examples implemented in it. Additionally, it allows one to choose an outcome type after the insertion of group sizes and number of positive outcomes in each group (or more if needed for specific type of outcome). Therefore, we can decide whether to present new outcomes such as risk ratios, odds ratios, hazard ratios, or others. Additionally, both fixed- and random-effects models can be assessed depending on heterogeneity of data. For proportions and prevalences (2) we recommend using MetaXL (EpiGear International Pty Ltd, Wilston, Queensland, Australia), which is a highly efficient

plug-in for Microsoft Excel. To establish pooled prevalences, one will need to implement the quantity of the study group and the number of individuals who present such variation. Evaluation of means (3) is only possible when, in addition to the mean itself, the authors extracted data such as standard deviations, interquartile range, standard error, or any other results that mathematically reflect the diversity of data. For this purpose, we recommend using Comprehensive Meta-analysis V3 tool (Biostat Inc., Englewood, NJ, USA) tool, analogously to the establishment of risks (1). We recommend presenting the data gathered in tables and figures (graphs, forest plots, etc.), rather than in a descriptive form, as it is much more reader-friendly.

Writing the manuscript

When writing the manuscript, we would like to recommend following a structure similar to that of MA, that is, separate sections for the introduction, materials and methods, results, and discussion. The manuscript should be concise, but comprehensive, covering the wholesome of the studied topic. Using simple language and avoidance of sublime phrases or complicated statistical descriptions (if not needed) could also be beneficial, as it would be more understandable for a wider group of readers.

The introduction should contain general information about the research at hand, why it is important, and how it impacts what we know now. It should also cover topics and definitions that have to be introduced in order for the reader to understand the full content of the study. Finally, the introduction should explain how the paper at hand will fill a missing piece of information in the literature.

Materials and methods should present the search strategy, that is, how the data were searched from the literature. Later, one has to explain the process of assessing the eligibility of the searched studies and the extraction techniques used. Subsequently, the methodological quality of all MA submitted must be assessed. Finally, the authors need to describe how the statistical analysis was performed. If any other steps were taken, in order to prepare the UR, the authors must state them in this chapter. There are no standardized, dedicated checklists or guidelines on what to include in the manuscript of an UR. However, we recommend using the PRISMA and MOOSE guidelines, which are implicitly designed for MA and SR.

The results should be concisely presented in the 'Results' section. For readability, all results should be categorized into sub-paragraphs and briefly presented. We recommend presenting the results as simply as possible, without paying attention, for example, to repeated words. This section should be the essence of the results, deprived of any stylistic

additives which may cause potential confusion or misunderstanding. More in-depth data should be presented in additional tables and/or figures.

The 'Discussion' section should present the clinical significance of the results obtained from the UR. Comparing the results with data from other MA or SR may be beneficial, as it may show the possible discrepancies and contradictory results on the topic at hand. The discussion should also address the practical implications of the gathered data and potentially suggest the direction for further research. Lastly, it is crucial to acknowledge the limitations of the study, as UR are only as good as the submitted studies. A simplified step-by-step diagram of how to prepare an umbrella review can be found in Figure 3.

CONCLUSIONS

UR represent the next major step in the advancement of evidence-based medicine, with great practical potential for physicians looking for the most up-to-date data on their topic of interest. We hope that our step-by-step guide may be a useful tool for researchers conducting UR in the future.

Funding

Conflict of interest: None declared

REFERENCES

1. D'Antoni A v., Tubbs RS, Patti AC, et al. The Critical Appraisal Tool for Anatomical Meta-analysis: A framework for critically appraising anatomical meta-analyses. Clinical Anatomy. 2022; 35(3): 323–331, doi: 10.1002/ca.23833.
2. Henry BM, Tomaszewski KA, Ramakrishnan PK, et al. Development of the Anatomical Quality Assessment (AQUA) Tool for the quality assessment of anatomical studies included in meta-analyses and systematic reviews. Clinical Anatomy. 2017; 30(1): 6–13, doi: 10.1002/ca.22799.

3.   Henry BM, Tomaszewski KA, Walocha JA. Methods of Evidence-Based Anatomy: a guide to conducting systematic reviews and meta-analysis of anatomical studies. Annals of Anatomy - Anatomischer Anzeiger. 2016; 205: 16–21, doi: 10.1016/j.aanat.2015.12.002.

4.   Higgins JPT, Thomas J, Chandler J, et al. (eds). Cochrane Handbook for Systematic Reviews of Interventions. Wiley 2019, doi: 10.1002/9781119536604.

5.   Langendam MW, Akl EA, Dahm P, et al. Assessing and presenting summaries of evidence in Cochrane Reviews. Syst Rev. 2013; 2(1): 81, doi: 10.1186/2046-4053-2-81.

6.   Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017; j4008, doi: 10.1136/bmj.j4008.

7.   Whiting P, Savović J, Higgins JPT, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol. 2016; 69: 225–234, doi: 10.1016/j.jclinepi.2015.06.005.
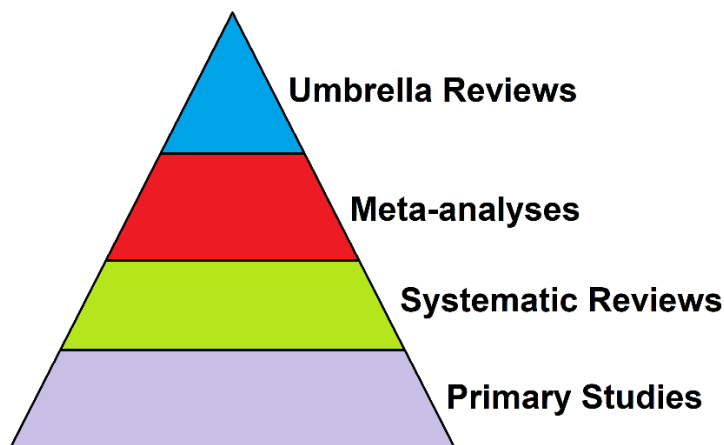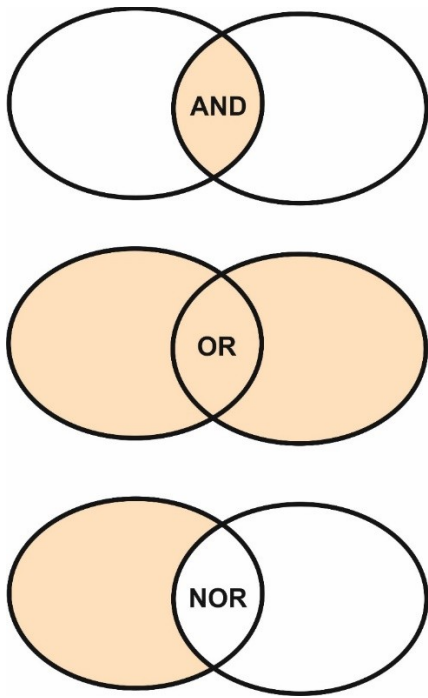
Figure 1. The evidence-based medicine pyramid.

Figure 2. Scheme, illustrating the Boolean conjunctions and their influence on the searched phrases.
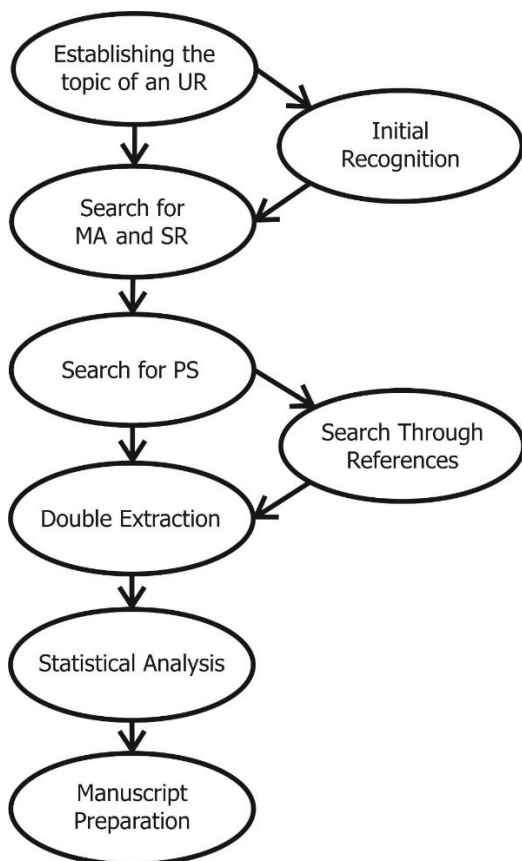

Figure 3. A simplified step-by-step diagram of how to prepare an umbrella review (UR); MA — meta-analyses; SR — systematic reviews; PS — primary studies.