NEVŞEHİR HACI BEKTAŞ VELİ ÜNİVERSİTESİ
Ara, bul.

Sosyal Bilimler Enstitüsü Dergisi
Cilt: 13 Sayı: 4 Aralık 2023
E-ISSN: 2149-3871

SYSTEMATIC REVIEW AND META-ANALYSIS WRITING PROCESS

SİSTEMATİK DERLEME VE META-ANALİZ YAZMA SÜRECİ

Hatice AKÇAKAYA
Muğla Sıtkı Koçman Üniversitesi
Eğitim Fakültesi
Özel Eğitim Bölümü
haticeakcakaya@mu.edu.tr
ORCID: 0000-0001-6508-9809

Geliş Tarihi:
10.01.2022

Kabul Tarihi:
12.12.2023

Yayın Tarihi:
25.12.2023

Anahtar Kelimeler
Derleme, Sistematik derleme, Meta-analizi

Keywords
Reviews, Systematic reviews, Meta-analysis

ÖZ

Sistematik derleme ve meta-analizi (SDM) çalışmaları, kanıt temelli uygulamalar hakkında klinisyenlere en kısa yoldan bilgi sağlamaları nedeniyle son derece önemlidir. Bu derlemenin amacı bir SDM çalışmasının nasıl yapıldığı ve hangi aşamalarla yürütüldüğünü açıklamaktır. Bir SDM'deki aşamalar; 1) Araştırma ekibinin oluşturulması 2) Araştırma sorusunun belirlenmesi 3) Ön araştırmanın yapılması 4) SDM'ye dahil edilecek ve dışlanacak çalışmaların özelliklerinin belirlenmesi 5) Literatür taraması stratejisinin belirlenmesi 6) SDM'nin bir protokole kaydedilmesi 7) İlgili veri tabanlarını taranması 8) Referans kayıt yönetimi 9) Araştırma havuzundaki çalışmaların başlık ve özetlerinin okunması ve SDM'ye dahil edilecek/dışlanacak çalışmaların seçilmesi 10) Kanıt kalitesinin belirlenmesi 11) Gözlemciler arası güvenirlik bakılması 12) İstatiksel analiz yöntemine karar verme ve verilerin analiz edilmesi 13) Ek analizlerin yapılması 14) Raporlamadır. Bu süreçte PRISMA bildirgesi kontrol listelerini takip etmek önemlidir. Sonuç olarak özellikle klinikte yapılan uygulamaları bilgilendiren SDM çalışmalarının elde edilecek sonuçların güvenirliği için titizlikle ve sistematik bir şekilde yürütülmesi gerekmektedir. Böylece sağlık bilimleri alanında klinikte uygulama yapan klinisyenler kısa yoldan güvenilir bir biçimde SDM sonuçlarını kullanabileceklerdir.

ABSTRACT

Because systematic review and meta-analysis (SRM) studies provide clinicians with the information quickly about evidence-based practices, they are crucial. It was aimed in this review to explain how an SRM study is conducted and with what stages it is carried out. The stages in an SRM are; 1) Creation of the research team 2) Determining the research question 3) Performing preliminary research 4) Describing the characteristics of studies to be included and excluded from the SRM 5) Determining of literature screening strategy 6) Recording the SRM into a protocol 7) Searching relevant databases 8) Reference record management 9) Reading the titles and abstracts of the studies in the research pool and choosing the studies to be included/excluded from the SRM 10) Determining the quality of evidence 11) Examining the inter-observer reliability 12) Deciding on the statistical analysis method and analyzing the data 13) Performing additional analyses 14) Reporting. In this process, following PRISMA statement checklists is important. Consequently, to ensure the accuracy of the results, SRM studies that guide clinical practices should be elaborately and systematically conducted. This will enable clinicians working in the health sciences (medicine, psychology, etc.) to quickly and accurately use the findings of SRM.

DOI: https://doi.org/10.30783/nevsosbilen.1232051
Atıf/Cite as: Akçakaya, H. (2023). Systematic review and meta-analysis writing process. *Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi, 13*(4), 1995-2007.

Introduction

With the emergence of evidence-based practices in the early 1990s, review articles are highly reliable sources for decision-making in clinics (Grant & Booth, 2009). Reviews are research syntheses that, regardless of the area of research, serve as a tool for evidence-based decision-making by reporting what has been written in the relevant scientific field up to that point and providing an interpretation of it (Sutton et al., 2019). According to the literature, there are five different types of reviews: meta-synthesis (Grant & Booth, 2009), narrative reviews, summative reviews, SRM (Dunst, 2018). A review that integrates and contrasts data from qualitative research is known as a meta-synthesis (Grant & Booth, 2009). The target of a narrative review is to provide a summary of a particular subject and to describe the types of findings that are attained in particular samples and the methodologies that are employed (Davies, 2000). Similar to narrative reviews, summative reviews contain numerical information about the type of treatment, settings, research designs, samples, and provide information about the latest situation on the subject (Dunst, 2018). Systematic reviews seek to systematically search, assess, and synthesize evidence from studies adhering to the methods outlined in various guidelines (Grant & Booth, 2009). Additionally, a meta-analysis is a sort of systematic review that makes use of effect sizes to ascertain whether there is a causal link between a treatment and an outcome in order to determine a treatment's effect (Dunst, 2018). According to these definitions, it may be argued that this review is a narrative review in the sense that it serves to summarize a certain subject.

Despite the fact that the terms SRM are often used synonymously in the literature, they are two different kinds of reviews. Studies that can provide evidence for a particular question are objectively identified, assessed, and synthesized in the systematic review. There is no need to produce a direct statistic such as meta-analysis. However, if there are more than two studies in a systematic review and it is possible to combine the data, meta-analysis can be performed (Hanratty, 2018). As a result, for a meta-analysis study, studies are systematically identified, assessed, synthesized, and statistical analysis is performed. Compared to randomized controlled trials, the findings of a meta-analysis offer a higher level of evidence (see Figure 1).

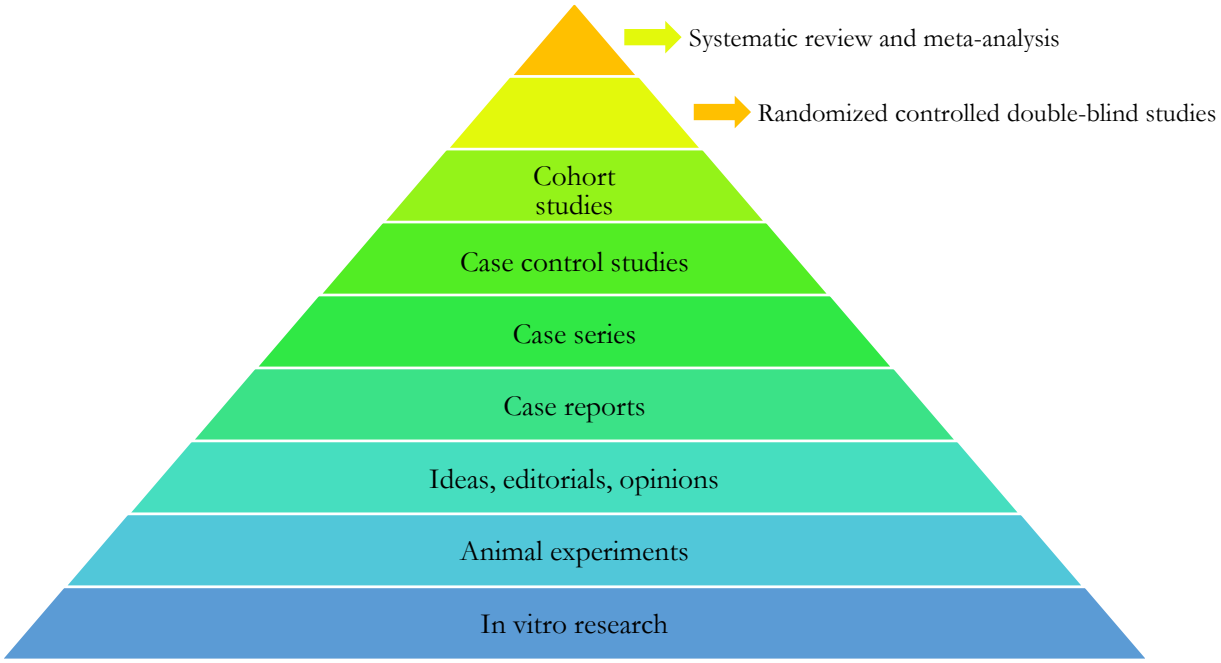


Figure 1. Levels of evidence (Ahn & Kang, 2018).

To write an SRM, there are specific steps that must be followed. These are:

1) Creation of the research team
2) Determining the research question
3) Performing preliminary research

4) Describing the characteristics of studies to be included and excluded from SRM
5) Determining of literature screening strategy
6) Recording the SRM into a protocol
7) Searching relevant databases
8) Reference Record Management
9) Reading the titles and abstracts of the studies in the research pool and selecting the studies to be included/excluded from the SRM
10) Determining the quality of evidence
11) Examining inter-observer reliability
12) Deciding on statistical analysis method and analyzing data
13) Performing additional analyses
14) Reporting (Crocetti, 2016; Moher et al., 2015; Şen, 2019; Tawfik et al., 2019).

Creation of the Research Team

Research team for SRM can consist of subject experts, and if possible, search experts such as librarians, SRM method experts or statisticians (Lê et al., 2022). According to Tawfik and his colleagues (2019), 2-3 review authors should work independently in all stages to ensure the accuracy and quality of data as well as determine whether the studies are eligible for review.

Determining the Research Question

Similar to individual research questions, an **SRM** starts with a research question or hypothesis These questions could be: Is there a statistically significant difference between those who received the treatment or therapy and those who did not from a variety of supplied treatments or therapies? Which form of therapy or treatment is more effective? Are the two determined variables significantly correlated? (Şen, 2019). When consisting of research questions, PICOS criteria [Participants (P) Intervention (I), Control (C), Outcomes (O), and Study designs (S)] can be used. According to PICOS criteria, the participants who will participate in the study, the intervention approaches to be used, the characteristics of the control group, the outcomes to be assessed, and the study design should be stated (Thomas et al., 2019). As an example, a review question might be: Do short-term memory and working memory outcomes differ statistically between children with learning disabilities who have taken working memory training and children with learning disabilities who have not? In this research question, children with learning disabilities who have taken working memory training are participants, working memory training is the intervention, and children with learning disabilities who have not taken working memory training are the control group, and short-term memory and working memory skills are outcomes. For this question, review authors may choose RCTs and other experimental designs as the study design. In addition, Cummings and his colleagues (2007) described the features of a good research question with abbreviations of FINER: Feasible (F), Interesting (I), Novel (N), Ethical (E), and Relevant (R). In this scope, feasible means an adequate number of participants, and includes manageable number of studies. Review authors should be aware of the probability of asking a question that might not be answerable using the existing findings. Interesting means getting the answer to interest the researchers and practitioners. A novel review addresses an authentic gap in knowledge, so review authors should be aware of previous reviews related to the topic. This reduces duplication of reviews. Authors should check for current reviews and meta-analyses in the published literature and also for ongoing reviews in the PROSPERO register of reviews before starting their own review. However, as new data published SRM can be updated (Liberati et al., 2009). Review questions should address problems that are important to review readers such as policymakers, and health professionals. Also, it should take into account the potential adverse effects of treatments or interventions in this stage (Thomas et al., 2019).

Performing Preliminary Research

Tawfik et al., (2019) suggest a preliminary search to determine relevant studies, provide the validity of the proposed idea, prevent duplication of previously addressed research questions, and ensure sufficient articles for conducting analysis. To do this, researchers can start by performing a simple search in Google Scholar or PubMed with search terms.

Describing the Characteristics of Studies to Included and Excluded from an SRM

Primary or secondary research could be included in the SRM. While data is collected data and analyzed first-hand in the primary research, collected data is re-analyzed in the secondary research (Glass, 1976). Studies that are included in an SRM as primary research may come from dissertations, papers published within the scope of a congress, articles, government reports, or project reports, whereas studies that are included as secondary research may come from old meta-analyses (Aslan, 2018). The parameters of the therapy supplied, or the methodology used should be well defined in order to appropriately respond to the research question. The effect of some confounding variables will also be eliminated in the case that the sample's characteristics are similar. According to PICOS criteria, after identifying the research question and characteristics of participants clarify which studies should be included and excluded in this stage. Listed below are some studies that will not be considered for the meta-analysis:

- Studies with interventions and sample unrelated to the research question
- Studies without comparison groups related to the research question
- Studies not written in a standard format
- Studies whose publication language is other than English (If the native language of the authors of synthesis is different or they are multilingual they can add publications in other languages)
- Studies outside the specified date ranges
- Studies that are accessed from different databases for the same publication (duplicate) (Basu, 2017).

Determining of Literature Screening Strategy

Following a description of the eligibility criteria for the SRM, identifying the eligibility criteria for research to be included facilitates the selection of relevant subject headings and words for the search strategy. Relevant published articles in databases may not be found if search strategies are ineffective or inappropriate. The search strategy should be based on the main concepts being inspected in a review. In general databases, a search strategy to determine studies generally has three groups of terms: (a) terms to use for the population; (b) terms to use for the intervention(s) evaluated; and (c) terms to use for the types of study design to be included (generally for randomized trials). However, for reviews of studies having different methodologies, it might be necessary to search only for the intervention or the population (Lefebvre et al., 2019). Also, outcomes of the intervention can be used as terms (Panitvisai et al., 2010). Moreover, to determine keywords, published studies or earlier reviews may be used. Sometimes, chosen keywords may not capture all necessary studies despite being appropriate for the review. In those situations, it may be useful to use just one term of intervention, use two or more terms, use a multicore or complicated approach that uses a number of searches, with different combinations of terms, or use citation searches on studies in addition to database searches (Lefebvre et al., 2019). The words that have been found to have the same or comparable meanings during the search are placed between the words with the conjunction 'or', while the words that do not have similar meanings are placed between the words with the conjunction 'and'. Additionally, 'near' or 'within 5 words' or 'adj5' screenings can be made. For instance, for a meta-analysis of salicylic acid and acute coronary syndrome, a search can be made ["salicylic acid" near "acute coronary syndrome"] or ["salicylic acid" within 5 words "acute coronary syndrome"] (Şen, 2019) Some databases such as PubMed/Medline archive articles under specific search terms (MeSH; Medical Subject Headings). Therefore, MeSH terms can be used in these databases (Basu, 2017).

Literature search can be conducted in databases such as Science Direct, PubMed, Scopus, EMBASE, PsycInfo, CINAHL Plus, ERIC and Cochrane Central Register of Controlled Trials (CENTRAL) in the field of health sciences (Ahn & Kang, 2018; Akçakaya et al., 2022). In addition, chosen keywords can be searched with ClinicalTrials.gov, Google Scholar, Open Gray. Also, it should be determined manual search strategy (for the detail see the title of Screening Relevant Databases).

Recording the SRM into a Protocol

The standardization and quality improvement of such studies are greatly aided by the registration of SRM on a recognized international platform (Ahn & Kang, 2018). Moreover, determining the subject of SRM, what kind of search strategy will be used in which databases with which keywords, inclusion/exclusion criteria, predetermining the statistics to be made and recording them in a certain system prevent waste of time and effort and publication bias for the same studies on the same subject (Moher et al., 2015). One such platform has been created by York University, Center for Reviews and Dissemination. For this purpose, the International

Prospective Register of Ongoing Systematic Reviews (PROSPERO; International Prospective Register of Ongoing Systematic Reviews, http://www.crd.york.ac.uk/prospero) was created in February 2011 (Moher et al., 2015). First of all, the first author's name, surname, address, e-mail, password, telephone number, institution, country name, professional category (academic, statistician, academic working actively in the clinic, researcher etc.), method of review (effect of health and social care interventions, side effects, diagnosis, economic evaluation, genetics, etc.) and related fields of health (cancer, paediatric health, education, ear, nose, throat, neurology etc.) must be filled and registered on the mentioned platform. Then, it is entered from the "Log in" section, "(Register your review now)" section is clicked, and 26 mandatory and 14 optional fields are filled. Meta-analysis studies referenced before the pandemic were first reviewed and then published. However, due to the pandemic, COVID-19 studies are prioritized, and studies pending more than 30 days are published as they are sent. It is suggested that extra care should be taken to ensure that the material is accurate because these studies are not reviewed by the PROSPERO team but go through an automatic control instead. Additionally, the registration must be confirmed within a month because the automatic registration confirmation is not notified by e-mail after 30 days. In this case, it would be appropriate for researchers to prepare their reviews in accordance with the preferred reporting items for SRM protocols (PRISMA-P) 2020 statement and to indicate this in the method section of their articles in order to indicate the systematicity of the study (see Page et al., 2021). In addition to PROSPERO, there is another platform to register: Cochrane (https://www.editorialmanager.com/cemd/default2.aspx). After signing up for this link the SRM proposal can be submitted.

Screening Relevant Databases

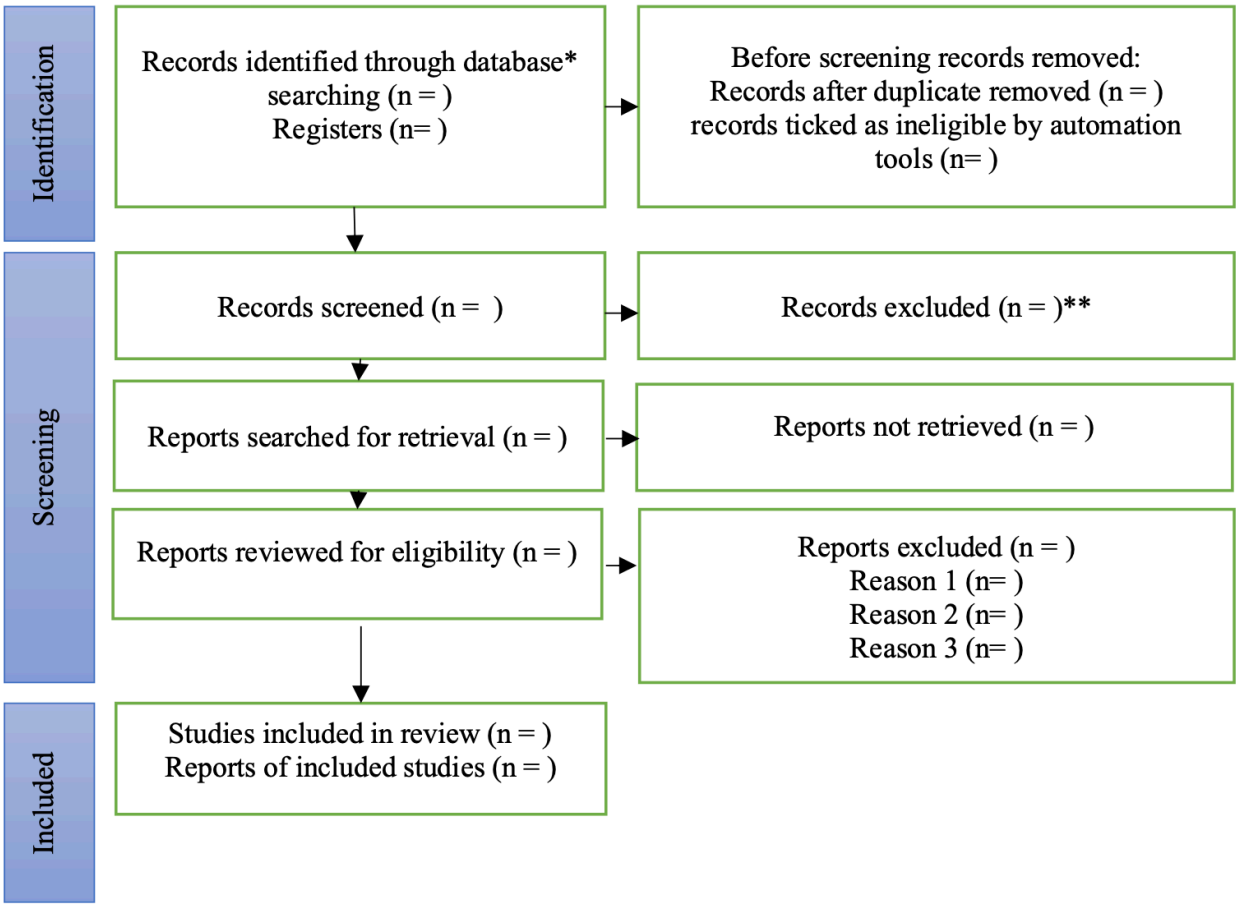
Title/abstract and full-text screening can be started in line with the protocol determined for SRM after PROSPERO registration. Searches are performed with specific keywords related to the subject of the review in the determined databases. In addition, manual search can reduce bias. For manual search, references from included reports can be searched, authors or experts can be contacted, and cited/related articles can be checked in Google Scholar and PubMed. Additionally, authors can track citations for each included study by tracking all the articles that cite that study. This may require electronic searching of databases. Furthermore, all 'similar' or 'related to' articles can be followed. These methods can be performed by 2-3 review authors to get all studies that need to be included in the review. If the full-texts are not accessed, some websites can be searched such as ResearchGate where we can direct full-text requests from authors. Also, full-texts not accessible can be requested from authors directly via e-mail (Tawfik et al, 2019).

Reference Record Management

Using software such as Endnote X8 (Thomson Reuters, 2016) and Covidence (www.covidence.org) while creating a research pool facilitates the process of identifying articles to include. For instance, after the necessary keywords are entered into the PubMed database, studies on the subject appear. Afterwards, these articles can be accessed by marking a date range or by ticking one of the full-text, abstract or free full-text access options. Additionally, the PubMed database can be scanned by bringing criteria related to the subject of interest in terms of article type, publication date, language, gender, and age. The articles that appear in the last case are transferred to the "citation manager" section from the four options that appear by clicking the "send to" button, resulting in a file that can be opened in the Endnote software. Then the file is opened in the Endnote software is exported as an XML file. Then it is registered with the Covidence software with an e-mail address. This software allows free use of up to 500 articles. In the Covidence program, articles are loaded from where it says "Import". This software finds different versions of the same article saved in different databases, preventing the same article from being scanned over and over again. Moreover, it is possible to access this software by installing the application on a mobile phone and this makes the job even easier. Then, all articles from other databases are uploaded to the Covidence program, and the software detects and deletes one of the articles that are the same (https://www.covidence.org). Other literature review softwares are DistillerSR, JBI SUMARI, and Rayyan (see Munn et al., 2019; Ouzzani et al., 2016).

Reading the Titles and Abstracts of the Studies in the Research Pool and Selecting the Studies to be Included/Excluded from the SRM

When all the screenings are finished, the two or three (three authors are more reliable) authors of the SRM independently (separately) read the titles and abstracts of all articles and the full text when they think it is necessary and vote on the articles with yes, no, or maybe according to the inclusion/exclusion criteria on Covidence or another literature review software (Tawfik et al., 2019). At this stage, the researcher is likely to encounter more than a thousand articles on the subject, and using such software makes the job easier. At the end of the process, the number of studies reached regarding the identification, screening, and number of included studies for SRM should be indicated in a figure (e.g., see Figure 2). It is recommended that the flowchart created for this process be added to SRM studies (Page et al, 2021).



*The number of records identified from each database or register searched (instead of the total number across all databases/registers) should be reported, if possible.

**How many records were excluded by a human and how many were excluded by automation tools should be indicated, if automation tools were used.

Figure 2. The process of determination of the articles' identification, screening, and inclusion

Determining the Quality of Evidence

The quality assessment includes external and internal validity. Methodological quality generally means internal validity. Internal validity can be affected by performance bias, selection bias, attrition bias, detection bias, reporting bias, and other biases during the research process. Hence, methodological quality assessment tools focus on these aspects, which is called risk of bias (Higgins et al., 2019) Examining the quality of the studies included in the SRM provides information about the certainty of what conclusions can be reached with the evidence obtained (Crowther et al., 2010). In order to rule out bias, it is important to ensure cross-checking with

at least two researchers to assess the quality of individual studies (Zeng et al., 2015). At this stage, a table including participant characteristics, number, variables of interest, quality score, kind of publication (e.g., thesis, article), name of the authors, and effect size, etc. is created. The reasons for excluding excluded studies should be on the table (Tawfik, 2019). Various tools are used in the literature to assess the quality of studies. Zeng et al. (2015) reviewed the methodological quality tools that can be used in SRM studies and suggested the use of a number of tools according to research types:

- The Newcastle-Ottawa Scale for cohort and case-control studies,
- The Methodological Index for Non-Randomized Studies (MINORS),
- Agency for Healthcare Research and Quality (ARHQ) methodology checklist for cross-sectional studies,
- For diagnostic accuracy testing studies, Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2),
- Systematic Review Centre for Laboratory Animal Experimentation; SYRCLE, for assessing animastudies,
- Assessment of Multiple Systematic Reviews (AMSTAR) for multiple systematic reviews/meta-analysis studies,
- An 18-item tool developed to assess case series studies and
- Appraisal of Guidelines, Research and Evaluation; AGREE-II tool for assessing clinical practice guidelines,

In addition to these tools, The Joanna Briggs Institute Critical Appraisal Tools have checklists for each research design such as case-control studies, analytical cross-sectional studies, case series, case reports, cohort studies, diagnostic test accuracy studies, prevalence studies, qualitative research, quasi-experimental studies, randomized controlled trials, and systematic reviews (see https://jbi.global/critical-appraisal-tools).

Inter-Observer Reliability

For inter-observer reliability, the researcher with a Covidence or other literature review software account invites the other researcher with his/her name and e-mail address. Thus, the researcher who accepts the invitation can vote on all files separately from the other researcher. Then, the extent to which the two authors agree on the studies to be included in the study can be determined using a formula. One of the formulas that can be used is Agreement÷(Agreement+Disagreement) X 100. According to this formula, 70% is necessary, 80% is sufficient, and 90% or more agreement is interpreted as good (House et al., 1981). Another statistical technique could be Cohen's Kappa. In Cohen's Kappa value, above 0.60 is considered acceptable, while 0.80 and above is interpreted as very good (Crocetti, 2016). Then, the authors make the final decision about the studies to be included in the SRM by discussing the studies that were voted "maybe" and the cases where one researcher voted yes and the other voted no, or vice versa. The authors should describe how decision rules were used to select data from multiple reports corresponding to a study, as well as how inconsistencies were resolved across reports (Page et al., 2020).

Deciding on Statistical Analysis Method and Analysing Data

All studies included in the systematic review before deciding on the statistical method may not be suitable for meta-analysis due to overlapping samples (Akçakaya et al., 2022). In such cases, even if only two studies are available from the last 20-30 years on the subject, a meta-analysis may be worthwhile. Although this kind of meta-analysis provides limited estimates (Ioannidis et al., 2008), it may ensure the quantitative accuracy of the meta-analysis by avoiding a narrative review (Akçakaya et al., 2022). The effect size gives information about the strength of the difference between the two groups or the strength/magnitude of the relationship between the two variables (Ellis, 2010). The effect size values used to perform meta-analysis differ from the effect size values obtained as a result of other statistical analysis methods. Effect size values that are frequently used for meta-analysis in the literature are risk ratio, correlation, and standardized mean difference values (Şen, 2019). Cohen's d of individual studies is calculated before calculating the standardized mean difference. Then, Hedges' g, Cohen's d or Glass' delta values are calculated for the standardized mean difference (Şen, 2019; Walker, 2005).

Hedges' g of these values is used to eliminate the bias that may arise due to the small number of samples (Hedges, 1981). Two different statistical models have been developed to determine the average effect size. These are the fixed effects model and the random effects model. While the fixed effects model makes calculations based on the assumption that the effect size parameters and the sample are homogeneous, the analysis is performed by assuming that the mean and variance are variable, and the sample is heterogeneous in the random effects model (Hedges & Vevea, 1998).

Analyses are performed using software such as ProMeta3 (https://idostatistics.com/prometa3/#prm_download) and Comprehensive Meta-Analysis (CMA), and forest graphs are obtained. As mentioned in the link provided above, ProMeta3 is currently free, while CMA has a 10-day free trial. As a result, if the effect size is 0.80 and above, it is interpreted as a high effect size, between 0.50 and 0.80 as medium-high, between 0.20 and 0.50 as small-medium effect size (Fritz et al., 2012).

Data Synthesis for Only Systematic Reviews

For research questions that concern issues other than the effectiveness of an intervention, current review methods are not always appropriate. This is because some research produces narrative rather than numerical data. This may lead to current systematic review methods being unable to synthesize the findings of some research and are consequently excluded from these reviews. This exclusion of non-RCT research has important inclusion for the psychology and healt care professionals, which has a considerable a number of different research methods (Evans, 2002). These research methods can be quantitative or qualitative research. For quantitative research, tabulation by grouping similar data, and using charts or graphics and narrative review can be used as the methodology (Moola, et al., 2015). As mentioned before, a review that integrates and contrasts data from qualitative research is known as a meta-synthesis (Grant & Booth, 2009). Narrative review, thematic synthesis, and critical interpretive synthesis can be used for analysis, in qualitative research (For the detail see Arai et al., 2007; Dixon-Woods et al., 2006; Thomas & Harden, 2008).

Performing Additional Analyses

After calculating the effect size, it is assessed whether there is heterogeneity between studies and its size, if any. For this, Q statistics and I^2 calculations are made (Crocetti, 2016). If the statistical value Q is greater than the Chi-square value at the same degrees of freedom, heterogeneity is assumed (Yıldırım & Şen, 2019). The other parameter assessing heterogeneity is I^2 value. The higher the I^2 value, the higher the heterogeneity. In other words, an I^2 the value between 0% and 40% is interpreted as low, between 41% and 60% as moderate, and above 61% as high heterogeneity (Moher et al., 2015). While the low level of heterogeneity indicates that the studies included in the meta-analysis are quite similar and consistent, the high level of heterogeneity may indicate that the number of studies is more than ten and that each study addresses the research question differently and is more often encountered in practice (Crocetti, 2016).

Reporting

In the reporting, the methods and results of the SRM should be described and presented. This phase consists of two steps: (a) describing the main elements in an **SRM** under a standard form (PRISMA 2020 report) and (b) preparing the manuscript comprehensively and succinctly according to the chosen academic journal (del Amo et al., 2018).

Conclusion

To sum up, writing an SRM is a quite systematic and planned process. The outcomes of the inferences are more valid and reliable when there is a systematic approach taken. This is crucial for the quality of the findings from SRM that guide practice.

References

Ahn, E., & Kang, H. (2018). Introduction to systematic review and meta-analysis. *Korean Journal of Anesthesiology, 71*(2), 103-112.

Akçakaya, H., Jayakody, D. M., & Doğan, M. (2022). Systematic review and meta-analysis of STM and WM in long-term CI users. *Contemporary School Psychology*, 1-20. https://doi.org/10.1007/s40688-022-00408-6

Arai, L., Britten, N., Popay, J., Roberts, H., Petticrew, M., Rodgers, M., & Sowden, A. (2007). Testing methodological developments in the conduct of narrative synthesis: a demonstration review of research on the implementation of smoke alarm interventions. *Evidence and Policy, 3*(3), 361-383.

Aslan, A. (2018). Sistematik derleme ve meta-analizi. *Acta Medica Alanya, 2*(2), 62-63. http://dx.doi.org/10.30565/medalanya.439541

Basu, A. (2017). How to conduct meta-analysis: A basic tutorial. https://ir.canterbury.ac.nz/bitstream/handle/10092/13709/peerj-preprints-2978%281%29.pdf?sequence=2 (accessed 18.06.2022).

Cummings S. R., Browner W. S., & Hulley S. B. (2007). Conceiving the research question and developing the study plan. In S. B. Hulley, S. R. Cummings, & W. S. Browner (Eds.), *Designing clinical research: An epidemiological approach* (4th ed., p. 14-22). Philadelphia (PA): Lippincott Williams & Wilkins.

Covidence systematic review software. Veritas Health Innovation, Melbourne, Australia. https://www.covidence.org. Accessed 4 Sep 2023

Crocetti, E. (2016). Systematic reviews with meta-analysis: Why, when, and how?. *Emerging Adulthood, 4*(1), 3-18.

Crowther, M., Lim, W., & Crowther, M. A. (2010). Systematic review and meta-analysis methodology. *Blood, The Journal of the American Society of Hematology, 116*(17), 3140-3146.

Davies, P. (2000). The relevance of systematic reviews to educational policy and practice. *Oxford Review of Education, 26*(3-4), 365-378.

del Amo, I. F., Erkoyuncu, J. A., Roy, R., Palmarini, R., & Onoufriou, D. (2018). A systematic review of Augmented Reality content-related techniques for knowledge transfer in maintenance applications. *Computers in Industry, 103*, 47-71.

Dixon-Woods, M., Cavers, D., Agarwal, S., Annandale, E., Arthur, A., Harvey, J., Hsu, R., Katbamna, S., Olsen, R., Smith, L., Riley, R., & Sutton, A. J. (2006). Conducting a critical interpretive synthesis of the literature on access to healthcare by vulnerable groups. *BMC Medical Research Methodology, 6*, 1-13.

Dunst, C. J. (2018). Kanıt-temelli erken çocukluk müdahale uygulamalarının belirlenmesinde araştırma sentezlerinin rolü [Role of research syntheses for identifying evidence-based early childhood intervention practices.]. (H. Akçakaya, Çev.). H. Bakkaloğlu & Ş. Demir (Eds.). Erken çocukluk özel eğitimi el kitabı içinde (ss. 539-562). Anı Yayıncılık.

Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results.* Cambridge, England: Cambridge University Press.

Evans, D. (2002). Systematic reviews of interpretive research: interpretive data synthesis of processed data. *The Australian Journal of Advanced Nursing, 20*(2), 22–26. https://search.informit.org/doi/10.3316/ielapa.405497388325103

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General, 141*(1), 2-18. http://dx.doi.org/10.1037/a0024338

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5*(10), 3-8.

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health İnformation and Libraries Journal, 26*(2), 91-108.

Hanratty, J. (2018). *What is the difference between a systematic review and a meta-analysis?* http://meta-evidence.co.uk/difference-systematic-review-meta-analysis/

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics, 6,* 107-128.

Hedges, L. V., & Vevea, J. L. (1998). Fixed-and random-effects models in meta-analysis. *Psychological Methods, 3*(4), 486-504.

Higgins, J. P. T., Savović, J., Page, M. J., Elbers, R. G., Sterne, J. A. C. (2019). Assessing risk of bias in a randomized trial. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page & V. A. & Welch VA (Eds.), *Cochrane handbook for systematic reviews of interventions* (2nd ed., pp. 205-228). Chichester (UK): John Wiley & Sons.

House, A. E., House, B. J., & Campbell, M. B. (1981). Measures of interobserver agreement: Calculation formulas and distribution effects. *Journal of Behavioral Assessment, 3*(1), 37-57.

Ioannidis, J. P., Patsopoulos, N. A., & Rothstein, H. R. (2008). Reasons or excuses for avoiding meta-analysis in forest plots. *British Medical Journal, 336*(7658), 1413-1415. https://doi.org/10.1136/bmj.a117

Lefebvre, C., Glanville, J., Briscoe, S., Littlewood, A., Marshall, C., Metzendorf, M. I., Noel-Storr, A., Rader, T., Shokraneh, F., Thomas, J., & Wieland, L. S. (2019). Searching for and selecting studies. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, V. A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (2nd Edition., pp.67-108). Chichester (UK): John Wiley & Sons.

Lê M., Loewen, H., Monnin, C., & Neilson, C. (March, 2022). *Part 1: Introduction to Systematic Reviews* [Video]. YouTube. https://www.youtube.com/watch?v=i2rmhf2bz_g

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Annals of Internal Medicine, 151*(4), W-65.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, L. A., & PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews, 4*(1), 1-25. https://doi.org/10.1186/2046-4053-4-1

Moola, S., Munn, Z., Sears, K., Sfetcu, R., Currie, M., Lisy, K., Tufanaru, C., Qureshi, R., Mattis, P., & Mu, P. (2015). Conducting systematic reviews of association (etiology): The Joanna Briggs Institute's approach. *JBI Evidence Implementation, 13*(3), 163-169.

Munn, Z., Aromataris, E., Tufanaru, C., Stern, C., Porritt, K., Farrow, J., Lockwood, C., Stephenson, M., Moola, S., Lizarondo, L., McArthur, A., Peters, M., Pearson, A., & Jordan, Z. (2019). The development of software to support multiple systematic review types: the Joanna Briggs Institute System for the Unified Management, Assessment and Review of Information (JBI SUMARI). *JBI Evidence Implementation, 17*(1), 36-43.

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews, 5*, 1-10. https://doi.org/10.1186/s13643-016-0384-4

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, r., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery, 88*, 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Panitvisai, P., Parunnit, P., Sathorn, C., & Messer, H. H. (2010). Impact of a retained instrument on treatment outcome: A systematic review and meta-analysis. *Journal of Endodontics, 36*(5), 775-780.

Sutton, A., Clowes, M., Preston, L., & Booth, A. (2019). Meeting the review family: Exploring review types and associated information retrieval requirements. *Health Information & Libraries Journal, 36*(3), 202-222. https://doi.org/10.1111/hir.12276

Şen, S. (2019). SPSS ile Meta-Analiz Nasıl Yapılır?. *Harran Maarif Dergisi, 4*(1), 21-49. http://dx.doi.org/10.22596/2019.0401.21.49

Tawfik, G. M., Dila, K. A. S., Mohamed, M. Y. F., Tam, D. N. H., Kien, N. D., Ahmed, A. M., & Huy, N. T. (2019). A step-by-step guide for conducting a systematic review and meta-analysis with simulation data. *Tropical Medicine and Health, 47*(1), 1-9. https://doi.org/10.1186/s41182-019-0165-6

Thomas, R. W. (2011). When student samples make sense in logistics research. *Journal of Business Logistics, 32*(3), 287-290. Retrieved from: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.2158-1592.2011.01023.x

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*(1), 1-10. https://doi.org/10.1186/1471-2288-8-45

Thomas, J., Kneale, D., McKenzie, J. E., Brennan, S. E., Bhaumik, S. (2019). Determining the scope of the review and the questions it will address. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, & T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions* (2nd ed., pp. 13-32). Chichester (UK): John Wiley & Sons.

Walker, D. A. (2005). A SPSS matrix for determining effect sizes from three categories: R and functions of r, differences between proportions, and standardized differences between means. *Journal of Modern Applied Statistical Methods, 4*(1), Article 30. https://doi.org/10.22237/jmasm/1114907400

Yıldırım, İ., & Şen, S. (2019). The effects of gamification on students' academic achievement: A meta-analysis study. *Interactive Learning Environments, 29*(8), 1-18. https://doi.org/10.1080/10494820.2019.1636089

Zeng, X., Zhang, Y., Kwong, J. S. W., Zhang, C., Li, S., Sun, F., Niu, Y., & Du, L. (2015). The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: A systematic review. *Journal of Evidence-Based Medicine, 8*(1), 2-10. https://doi.org/10.1111/jebm.12141

GENİŞLETİLMİŞ ÖZET

Sistematik derleme ve meta-analizi (SDM) sıklıkla birbirinin yerine kullanılsa da birbirinden farklıdır. Sistematik derlemede belirli bir soruya ilişkin kanıt oluşturabilecek çalışmalar nesnel bir şekilde belirlenir, değerlendirilir ve sentezlenir. Doğrudan bir istatistik yapılması şart değildir. Meta-analizi ise bir müdahalenin etkisini belirleyebilmek için müdahale ve sonuç arasında nedensel bir ilişki olup olmadığını belirlemek amacıyla etki büyüklüklerini kullanan sistematik bir derleme türüdür (Dunst, 2018). Sistematik bir derlemede ikiden fazla çalışma bulunuyorsa ve verileri birleştirmek mümkünse meta-analizi yapılabilmektedir (Hanratty, 2018). Bir SDM yazmak için izlenmesi gereken adımlar: (Crocetti, 2016; Moher vd., 2015; Şen, 2019).

Öncelikle araştırma ekibi oluşturulur. SDM ekibi konu uzmanlarından, SDM yöntemi/istatistik alanındaki uzmanlardan ve mümkünse kütüphanecilerden oluşabilir (Lê vd., 2022). Araştırma soruları oluşturulurken PICOS kriterlerine göre çalışmaya katılacak katılımcılar (**P**articipants), müdahale yaklaşımları (**I**ntervention), kontrol grubunun özellikleri (**C**ontrol), değerlendirilecek sonuçlar (**O**utcomes) ve çalışma tasarımı (**S**tudy design) belirtilmelidir. (Thomas vd., 2019).

Önerilen fikrin geçerliliğini sağlamak, daha önce ele alınan araştırma sorularının tekrarlanmasını önlemek ve yeterli makalenin olup olmadığını belirlemek amaçlarıyla bir ön araştırma yapılması önerilmektedir. Google Akademik veya PubMed'de arama terimleriyle basit bir arama yapılabilir (Tawfik vd., 2019).

Daha sonra sistematik derleme ya da meta-analizine dahil edilecek/dışlanacak çalışmaların özellikleri betimlenir. Birincil ya da ikincil araştırmalar bir SDM'ye dahil edilebilir (Aslan, 2018). Araştırma sorusunun en iyi şekilde yanıtlanabilmesi için verilen müdahalenin ya da uygulanan metodolojinin sınırlarının iyi çizilmesi gerekmektedir. Ayrıca müdahalenin uygulandığı örneklemin özelliklerinin benzer olması bazı karıştırıcı değişkenlerin etkisini bertaraf edecektir (Thomas, 2011). Ardından tarama stratejisi belirlenir. Bunun için konuyla ilgili tüm anahtar sözcükler belirlenir ve tarama stratejileri oluşturulur. Bazen seçilen anahtar kelimeler, derleme için uygun olmasına rağmen gerekli tüm çalışmaları kapsamayabilir. Böyle durumlarda, yalnızca bir müdahale terimi kullanmak, iki veya daha fazla terim kullanmak, farklı terim kombinasyonları kullanmak veya çalışmalarda veritabanı aramalarına ek olarak alıntı aramaları kullanmak yararlı olabilir. (Lefebvre vd., 2019).

Sağlık bilimleri alanında Science Direct, PubMed, Scopus, EMBASE, PsycInfo, CINAHL Plus, ERIC ve Cochrane Central Register of Controlled Trial gibi veri tabanlarında literatür taraması yapılabilir (Ahn & Kang, 2018; Akçakaya vd., 2022). Ayrıca ClinicalTrials.gov, Google Akademik, Open Grey ile meta-analizine dahil edilecek çalışmaların kaynakçası taranabilir.

SDM'lerin belirli bir sisteme kaydedilmesi aynı konu ile aynı çalışmalar için boşuna zaman ve emek harcanmasının ve yayın yanlılığının önüne geçmektedir (Moher vd., 2015). Bunun için PROSPERO (International Prospective Register of Ongoing Systematic Review) (Moher vd., 2015) veya Cochrane (https://www.editorialmanager.com/cemd/default2.aspx) kullanılabilir. Ayrıca yazarların çalışmalarını PRISMA-P 2020 beyanına uygun şekilde hazırlamaları ve bunu makalelerinin yöntem bölümünde belirtmeleri gerekmektedir (bk. Page vd., 2021). PROSPERO kaydı sonrası veri tabanlarında başlık/özet ve tam metin taramasına başlanabilir. Ayrıca manuel arama yanlılığı azaltabilir (Tawfik vd, 2019). Bu süreçte Endnote X8 (Thomson Reuters, 2016) ve Covidence (www.covidence.org) gibi bazı yazılımlar kullanılabilir. Ayrıca lüteratür taramak için DistillerSR, JBI SUMARI ve Rayyan kullanılabilir (bk. Munn vd., 2019; Ouzzani vd., 2016).

Tüm taramalar bittiğinde araştırma havuzundaki çalışmaların başlık ve özetleri okunur ve SDM'ye dahil edilecek/dışlanacak çalışmalar seçilir. SDM'nin iki yazarı birbirlerinden bağımsız bir şekilde tüm makalelerin başlık ve özetlerini ve gerekli olduğunu düşündüklerinde tam metnini okurlar ve dahil etme/dışlama ölçütleri doğrultusunda Covidence ya da başka bir literatür tarama programı kullanabilirler (Tawfik vd., 2019). Sürecin sonunda SDM için belirleme, tarama ve dahil edilen çalışmaların sayılarına ilişkin erişilen çalışma sayıları bir şekil içinde belirtilmelidir (bk. http://www.prisma-statement.org/PRISMAStatement/).

Daha sonra kanıt kalitesi belirlenmelidir. İç geçerlilik, araştırma süreci sırasındaki performans, seçim, yıpranma, tespit, raporlama yanlılığı ve diğer yanlılıklardan etkilenebilir. Dolayısıyla metodolojik kalite değerlendirme araçları, yanlılık riski olarak adlandırılan bu yönlere odaklanmaktadır (Higgins vd., 2019). Yanlılığı ekarte edebilmek için bireysel çalışmaların kalitelerini en az iki araştırmacı değerlendirmelidir. SDM çalışmalarında kullanılabilecek yöntemsel kalite araçlarını için bk. Zeng vd., 2015. Ardından gözlemciler arası güvenirlik incelenmelidir. Kullanılabilecek formüllerden biri; Görüş Birliği÷(Görüş Birliği+Görüş Ayrılığı) X100'dür. Bu formüle göre %70 gerekli, %80 yeterli ve %90 ve üzeri görüş birliği iyi olarak yorumlanmaktadır (House vd.,

1981). Bir diğer istatiksel teknik Cohen'in Kappası olabilir. Cohen'in Kappa değerinde ise 0,60 üzeri kabul edilebilir olarak kabul edilirken 0,80 ve üzeri çok iyi olarak yorumlanmaktadır (Crocetti, 2016). SDM'ye dahil olacak çalışmaları, "belki" oyu verilen çalışmalar ile bir araştırmacının evet, diğerinin hayır oyu verdiği ya da tam tersi durumlar birlikte tartışılarak dahil edilecek çalışmalar hakkında yazarlar nihai kararı verirler.
Ardından isttaiksel analiz aşamasına geçilir. SDM'ye dahil edilen tüm çalışmalar örtüşen örneklemler nedeniyle meta-analizi yapmak için uygun olmayabilir (Akçakaya vd., 2022). Bu gibi durumlarda konuyla ilgili son 20-30 yıla ait sadece iki çalışma mevcut olsa bile bir meta-analiz faydalı olabilir. Bu tür bir meta-analiz sınırlı tahminler sağlasa da (Ioannidis vd., 2008), öyküleyici bir derlemeden kaçınarak meta-analizin niceliksel doğruluğunu sağlayabilir (Akçakaya vd., 2022). Meta-analizi yapmak için sıklıkla kullanılan etki büyüklüğü değerleri risk oranı ve standartlaştırılmış ortalama farkı (SOF) korelasyondur (Şen, 2019). SOF'u hesaplamadan önce bireysel çalışmaların Cohen d'si hesaplanır. SOF için Hedges' g, Cohen's d veya Glass' delta değerleri hesaplanmaktadır (Şen, 2019; Walker, 2005). Bu değerlerden Hedges' g örneklem sayısının az olmasına bağlı olarak ortaya çıkabilecek yanlılığı gidermek amacıyla kullanılmaktadır (Hedges, 1981). Ortalama etki büyüklüğünü belirlemek üzere sabit etkiler ve rastgele etkiler modeli geliştirilmiştir. Sabit etkiler modeli, etki büyüklüğü parametreleri ve örneklemin homojen olduğu varsayımına dayanırken, rastgele etkiler modelinde ortalama ve varyansın değişken olduğu ve örneklemin heterojen olduğu varsayılmaktadır (Hedges & Vevea, 1998). Etki büyüklüğü 0,80 ve üzeri ise yüksek etki büyüklüğü, 0,50 ile 0,80 arası ise orta-yüksek, 0,20 ile 0,50 arası küçük-orta etki büyüklüğü şeklinde yorumlanmaktadır (Fritz vd., 2012).
Heterojenliğin değerlendirilmesi için Q istatistiği ve I^2 kullanılır (Crocetti, 2016). Aynı serbestlik derecelerinde Q istatistik değeri ki-kare değerinden büyükse, heterojenlik olduğu varsayılmaktadır (Yıldırım ve Şen, 2019). I^2 değeri arttıkça heterojenlik de artmaktadır. Yani I^2 değeri %0 ila %40 düşük, %41 ila %60 arası orta ve %61 üzerindeki değer ise yüksek düzeyde heterojenlik olarak yorumlanmaktadır (Moher vd., 2015). Düşük heterojenlik meta-analizine dahil edilen çalışmaların benzer ve tutarlı olduğunu gösterirken, yüksek heterojenlik çalışma sayısının ondan fazla olduğunu ve her bir çalışmanın araştırma sorusunu farklı şekilde ele alındığını gösterebilmektedir ve daha çok karşılaşılmaktadır (Crocetti, 2016). Raporlamada **ise** SDM'deki ana unsurlar PRISMA 2020 raporuna göre tanımlanmalı ve makale seçilen akademik dergiye göre kapsamlı ve kısa bir şekilde hazırlanmalıdır (del Amo vd., 2018). Özetle SDM yazma sürecindeki sistematiklik çıkarımların sonuçlarının daha geçerli ve güvenilir olmasını sağlamaktadır.