



Quality Reporting of Systematic Review and Meta-Analysis According to PRISMA 2020 Guidelines: Results from Recently Published Papers in the *Korean Journal of Radiology*

Ho Young Park¹, Chong Hyun Suh¹, Sungmin Woo², Pyeong Hwa Kim¹, Kyung Won Kim¹

¹Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea;

²Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Objective: To evaluate the completeness of the reporting of systematic reviews and meta-analyses published in a general radiology journal using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines.

Materials and Methods: Twenty-four articles (systematic review and meta-analysis, n = 18; systematic review only, n = 6) published between August 2009 and September 2021 in the *Korean Journal of Radiology* were analyzed. Completeness of the reporting of main texts and abstracts were evaluated using the PRISMA 2020 statement. For each item in the statement, the proportion of studies that met the guidelines' recommendation was calculated and items that were satisfied by fewer than 80% of the studies were identified. The review process was conducted by two independent reviewers.

Results: Of the 42 items (including sub-items) in the PRISMA 2020 statement for main text, 24 were satisfied by fewer than 80% of the included articles. The 24 items were grouped into eight domains: 1) assessment of the eligibility of potential articles, 2) assessment of the risk of bias, 3) synthesis of results, 4) additional analysis of study heterogeneity, 5) assessment of non-reporting bias, 6) assessment of the certainty of evidence, 7) provision of limitations of the study, and 8) additional information, such as protocol registration. Of the 12 items in the abstract checklists, eight were incorporated in fewer than 80% of the included publications.

Conclusion: Several items included in the PRISMA 2020 checklist were overlooked in systematic review and meta-analysis articles published in the *Korean Journal of Radiology*. Based on these results, we suggest a double-check list for improving the quality of systematic reviews and meta-analyses. Authors and reviewers should familiarize themselves with the PRISMA 2020 statement and check whether the recommended items are fully satisfied prior to publication.

Keywords: Reporting quality; Systematic review; Meta-analysis; PRISMA 2020

INTRODUCTION

Systematic reviews and meta-analyses have several

Received: October 20, 2021 **Revised:** November 26, 2021

Accepted: December 21, 2021

Corresponding author: Chong Hyun Suh, MD, PhD, Department of Radiology and Research Institute of Radiology, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea.

• E-mail: chonghyunsuh@amc.seoul.kr

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

strengths over individual studies because they provide estimated outcomes with higher precision, address questions that cannot be asked in individual studies, and provide evidence-based guidance from conflicting results [1]. As a result, an increasing number of systematic reviews and meta-analyses are published every year in various medical fields [2]. Accordingly, the quality of reporting has been emphasized in systematic reviews and meta-analyses in order to provide clarity and transparency regarding study conduct procedures [3]. This is particularly important for systematic reviews and meta-analyses because the synthesized results are influenced by the results from individual studies and therefore can be misleading if the

individual results are biased [4].

In 2009, the first Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement was published with the aim to improve the quality of reporting [3]. Since then, methodological approaches, such as result synthesis and risk of bias assessment, have advanced, thereby necessitating update of the guidelines; thus, an updated version of the PRISMA statement was published in 2020 [5]. Despite the publication of the PRISMA statement, the quality of systematic reviews and meta-analyses still varies between individual articles and journals [6]. Moreover, there has been only a modest improvement in the quality of reporting in radiology articles since the publication of the PRISMA 2009 statement, suggesting that there is still room for further improvement [6].

To the best of our knowledge, the number of studies in the field of radiology that evaluated the quality of reporting in systematic reviews and meta-analyses using the PRISMA 2020 statement has been limited. Therefore, the goal of our study was to assess the reporting quality of recent publications in the *Korean Journal of Radiology* using the PRISMA 2020 statement. Based on the assessment, we aimed to provide suggestions for authors on how to improve the quality of their reports.

MATERIALS AND METHODS

Search Strategy and Study Selection

Using the MEDLINE database, we identified all potentially relevant systematic reviews, with or without meta-analysis, published in a single peer-reviewed journal, the *Korean Journal of Radiology*, between August 2009 and September 2021. Because the first PRISMA statement was published in July 2009 [3], we did not include studies published earlier than that date. The search terms were (“Korean Journal of Radiology”[Journal]) AND ((systematic review) OR (meta-analysis)). A total of 31 records (i.e., abstracts and titles) were retrieved from the MEDLINE database and two reviewers evaluated the eligibility of each article. Among them, two records [7,8] were removed before screening because they were published before 2009. Three records [9-11] were excluded while screening because they were guidelines for systematic reviews and meta-analyses. As a result, full texts from 26 publications were retrieved and assessed for eligibility, two of which [12,13] were excluded because they were not systematic reviews. Finally, 24 publications were included in our analysis (Fig. 1) [14-37].

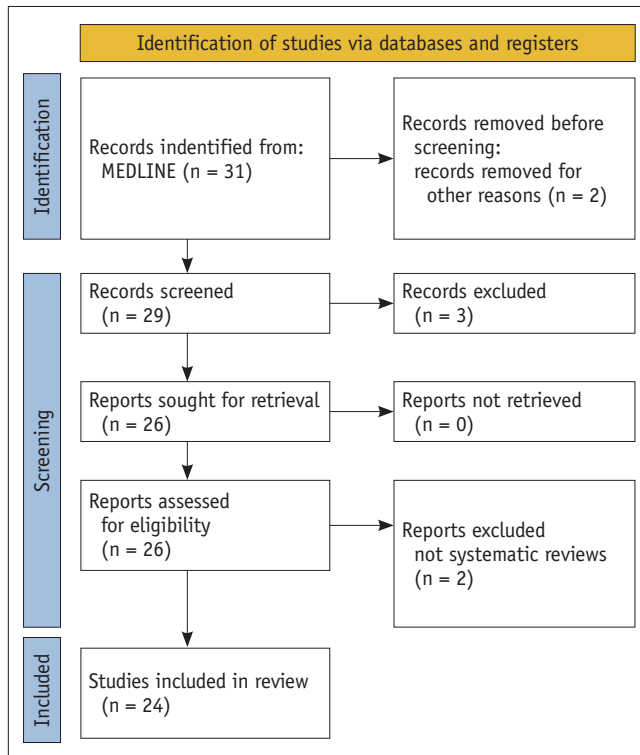


Fig. 1. Study selection process using the PRISMA 2020 flow diagram.

Data Extraction

Data extraction from the included studies was performed by two reviewers (with 2 and 8 years of experience in systematic review and meta-analysis studies, respectively), as shown in detail in the Supplement.

Major Changes in the PRISMA 2020 Statement

Several changes have been made in the PRISMA 2020 statement compared with the PRISMA 2009 statement. Although the number of main items in the checklists was unchanged (27 items), a large number of sub-items were added (42 in total, including sub-items) to provide more comprehensive guidelines. In addition, checklists for the abstracts (12 items) were included in the guidelines. Table 1 shows a brief summary of the major updates made in 2020 [5]. Two reviewers reviewed the checklists and agreed to a consensus for each item, as detailed in the Supplement.

Data Analysis

We extracted the PRISMA 2020 checklist items that were satisfied by fewer than 80% of the articles and grouped them into eight relevant domains. Suggestions for better quality of systematic reviews and meta-analyses were provided based on these domains. Evaluation of the

adherence of the included articles to the PRISMA 2020 statement is described in the Supplement.

RESULTS

Characteristics of the Included Studies

The characteristics of the 24 included studies are

summarized in Figure 2 and Supplementary Table 1. Briefly, 18 studies (14 univariate and four bivariate) were systematic reviews with meta-analyses [14,15,19-29,31-35] and six studies were systematic reviews without meta-analyses [16-18,30,36,37]. In terms of the type of data used for analyses, 13 studies used dichotomous data to measure the following outcomes: 1) efficacy or safety of an intervention

Table 1. Summary of the Major Updates in the PRISMA 2020 Statement

Major Updates
• Inclusion of checklists for abstract (item #2)
• Requiring full search strategies for all databases modified from full search strategy for at least one database (item #7)
• Emphasis on study selection process and data extraction, requiring how many reviewers evaluated for study eligibility and data extraction, and whether they worked independently (item #8). In addition, recommendation for authors to cite studies that seemed to meet inclusion criteria but excluded in the final stage and explain the reason for exclusion (item #16b)
• Detailed specification on result synthesis methods, providing subitems regarding data handling, visual data presentation (e.g., forest plot), statistical methods for pooling results and rationale for choosing one, methods to explore study heterogeneity, and sensitivity analysis used to evaluate robustness of the pooled results (items #13a-13f and #20a-20d)
• Addition of new items regarding the assessment of certainty of the evidence for an outcome (items #15 and #20)
• Emphasis on study registration and protocol (items #24a-c)

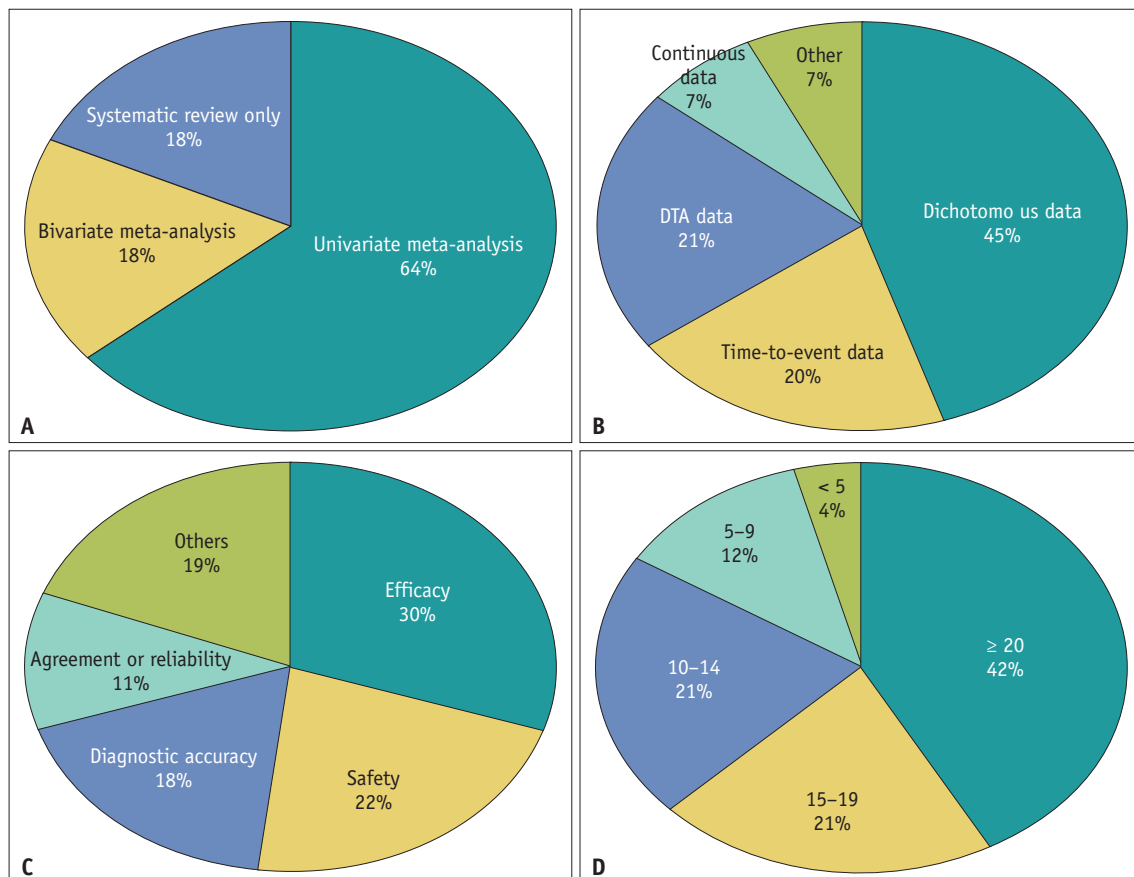


Fig. 2. Summary charts of the included studies according to (A) study type, (B) data type, (C) main outcome, and (D) number of included studies. “Other” in (B) data type was descriptive data regarding imaging protocols. “Others” in (C) main outcome were study or reporting quality, HRQoL score, imaging finding, imaging protocol, and diagnostic yield. DTA = diagnostic test accuracy, HRQoL = health-related quality of life

(proportion of tumor response, recurrence, or treatment-related complications), 2) efficacy of a diagnostic test (proportion of technical failure and unreliable measurement), 3) imaging features in a certain disease (proportion of specific imaging findings), 4) evaluation of study quality or reporting quality (proportion of studies that met the specific criteria), and 5) diagnostic yield [14,15,18,19,22-24,28,32-36]; six studies used time-to-event data to calculate the efficacy of a new intervention or the reliability between overall survival and imaging surrogate markers [15,22,31-34]; six studies used diagnostic test data to pool the diagnostic performance of index tests [16,25-27,29,37]; two studies used continuous data to evaluate the agreement and reliability of measurements between imaging methods [20,21]; one study used descriptive data from imaging protocols in randomized

controlled trials of acute ischemic stroke [30]; and one study used qualitative and quantitative data to assess the health-related quality-of-life in patients with hepatocellular carcinoma [17]. The number of included studies ranged from 4 to 516, with the majority (83%, 20 out of 24) of the articles including more than 10 studies. The statistical methods used in the included articles are summarized in Supplementary Table 2.

Assessment Using PRISMA 2020 Checklists

Overall Results

Each item in the PRISMA 2020 checklist and abstract checklist was evaluated for the included articles (Tables 2, 3). Of the 12 items in the abstract checklist, eight were reported in fewer than 80% of the articles (Fig. 3). To generate

Table 2. PRISMA 2020 Checklist for Abstract and the Number of Reported Articles in the Korean Journal of Radiology Since 2015

Section and Topic	Item #	Checklist Item	Number of Reported Articles (n/n, %)
TITLE			
Title	1	Identify the report as a systematic review	21/24 (88)
BACKGROUND			
Objectives	2	Provide an explicit statement of the main objective(s) or question(s) the review addresses	22/24 (92)
METHODS			
Eligibility criteria	3	Specify the inclusion and exclusion criteria for the review	0/24 (0)
Information sources	4	Specify the information sources (e.g., databases, registers) used to identify studies and the date when each was last searched	17/24 (71)
Risk of bias	5	Specify the methods used to assess the risk of bias in the included studies	3/23 (13)*
Synthesis of results	6	Specify the methods used to present and synthesize results	8/20 (40) [†]
RESULTS			
Included studies	7	Give the total number of included studies and participants and summarise relevant characteristics of studies	13/24 (54)
Synthesis of results	8	Present results for main outcomes, preferably indicating the number of included studies and participants for each If meta-analysis was done, report the summary estimate and confidence/credible interval. If comparing groups, indicate the direction of the effect (i.e., which group is favored)	22/24 (92)
DISCUSSION			
Limitations of evidence	9	Provide a brief summary of the limitations of the evidence included in the review (e.g., study risk of bias, inconsistency, and imprecision)	4/24 (17)
Interpretation	10	Provide a general interpretation of the results and important implications	24/24 (100)
OTHER			
Funding	11	Specify the primary source of funding for the review	12/24 (50)
Registration	12	Provide the register name and registration number	0/24 (0)

*One study was not applicable for item #5 [18], [†]Four studies were not applicable for item #6 [17,18,30,36].

Table 3. PRISMA 2020 Checklist and the Number of Reported Articles in the Korean Journal of Radiology Since 2015

Section and Topic	Item #	Checklist Item	Number of Reported Articles (n/n, %)
TITLE			
Title	1	Identify the report as a systematic review	22/24 (92)
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist	-
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge	24/24 (100)
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses	24/24 (100)
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses	23/24 (96)
Information sources	6	Specify all databases, registers, websites, organizations, reference lists, and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted	24/24 (100)
Search strategy	7	Present the full search strategies for all databases, registers, and websites, including any filters and limits used	20/24 (83)
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process	17/24 (71)
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process	21/24 (88)
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g., for all measures, time points, analyses), and if not, the methods used to decide which results to collect	23/24 (96)
	10b	List and define all other variables for which data were sought (e.g., participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information	23/24 (96)
Study risk of bias assessment	11	Specify the methods used to assess the risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process	17/22 (77)*
Effect measures	12	Specify for each outcome the effect measure(s) (e.g., risk ratio, mean difference) used in the synthesis or presentation of results	23/24 (96)

Table 3. PRISMA 2020 Checklist and the Number of Reported Articles in the Korean Journal of Radiology Since 2015 (Continued)

Section and Topic	Item #	Checklist Item	Number of Reported Articles (n/n, %)
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g., tabulating the study intervention characteristics and comparing against the planned groups for each synthesis [item #5])	17/22 (77)*
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions	10/20 (50)*†
	13c	Describe any methods used to tabulate or visually display the results of individual studies and syntheses	8/20 (40)*†
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s) If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used	7/20 (35)*†
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g., subgroup analysis, meta-regression)	13/19 (68)*†‡
	13f	Describe any sensitivity analyses conducted to assess the robustness of the synthesized results	5/18 (28)*†§
Reporting bias assessment	14	Describe any methods used to assess the risk of bias due to missing results in a synthesis (arising from reporting biases)	13/20 (65)*†
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome	2/22 (9)*
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram	23/24 (96)
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded	6/24 (25)
Study characteristics	17	Cite each included study and present its characteristics	21/22 (96)*
Risk of bias in studies	18	Present assessments of the risk of bias for each included study	7/22 (32)*
Results of individual studies	19	For all outcomes, present, for each study: 1) summary statistics for each group (where appropriate) and 2) an effect estimate and its precision (e.g., confidence/credible interval), ideally using structured tables or plots	22/22 (100)*
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies	0/20 (0)*†
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g., confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect	20/20 (100)*†
	20c	Present results of all investigations of possible causes of heterogeneity among study results	14/19 (74)*†‡
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results	5/18 (28)*†§
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed	13/20 (65)*†
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed	2/22 (9)*
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence	24/24 (100)
	23b	Discuss any limitations of the evidence included in the review	24/24 (100)
	23c	Discuss any limitations of the review processes used	14/24 (58)
	23d	Discuss implications of the results for practice, policy, and future research	23/24 (96)

Table 3. PRISMA 2020 Checklist and the Number of Reported Articles in the Korean Journal of Radiology Since 2015 (Continued)

Section and Topic	Item #	Checklist Item	Number of Reported Articles (n/n, %)
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered	0/24 (0)
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared	0/24 (0)
	24c	Describe and explain any amendments to the information provided at registration or in the protocol	0/24 (0)
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review	13/24 (54)
Competing interests	26	Declare any competing interests of review authors	15/24 (63)
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review	0/24 (0)

*Kim et al. [18] and Kang et al. [36] were not applicable for items #11, #13-15, #17-22 because these studies only calculated the proportions of the articles that met the certain criteria and no statistical method or modeling was used. †Kang et al. [17] and Suh et al. [30] were not applicable for items #13b-13f, #14, #20-21 because quantitative synthesis was not performed in these studies. ‡Wang et al. [33] was not applicable to items #13e and #20c because the study heterogeneity was minimal (supported by Cochran's Q-test and Higgins inconsistency index [I²] test), thereby not requiring an additional analysis. §Chung et al. [16] and Kim et al. [37] were not applicable for item #13f and #20d because meta-analysis was not performed in these studies and only narrative syntheses of results (i.e., providing ranges of values) were available.

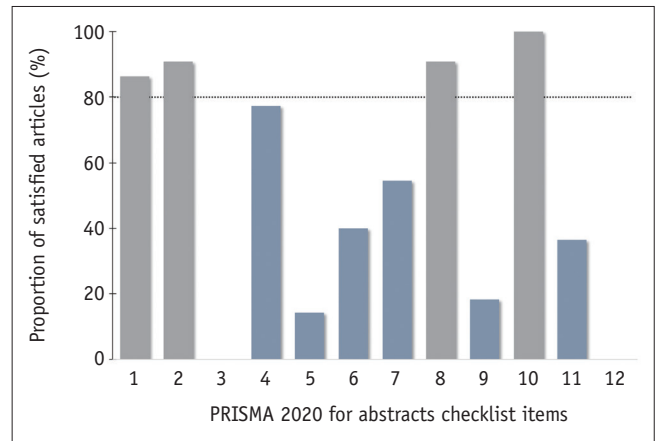


Fig. 3. Bar chart of the proportion of articles that satisfied each item of the PRISMA 2020 abstract checklist. Blue bars indicate the items that were satisfied by fewer than 80% of the studies. Gray bars indicate the items that were satisfied by greater than 80% of the studies.

abstracts with better quality, exclusion criteria, risk of bias assessment tools, statistical methods, and limitations of evidence should be included.

Of the 42 items (including sub-items) included in the guidelines for the main text, 24 were reported in fewer than 80% of the articles. While most studies satisfied the items in the Title, Introduction, and Discussion, incomplete reports were frequently observed in the Methods and the Results, especially in result synthesis (Fig. 4). The 24 items were grouped into eight domains for further exploration: 1) assessment of the eligibility of potential articles (items #8, #16b), 2) assessment of the risk of bias (items #11, #18), 3) synthesis of results (items #13a, #13b, #13c, #13d, #20a), 4) additional analysis (items #13e, #13f, #20c, #20d), 5) assessment of the non-reporting bias (items #14, #21), 6) assessment of the certainty of evidence (items #15, #22), 7) provision of limitations of the study (item #23c), and 8) additional information (items #24-#27).

Assessment of the Eligibility of Potential Articles (Items #8, #16b)

Seven articles [16,24,27,30-32,36] did not report how many reviewers participated in the evaluation of study eligibility or whether they worked independently (item #8). Eighteen articles [16-29,32,34,36,37] did not cite the studies that seemed to meet the inclusion criteria, but were excluded in the final stage or did not explain the reason for exclusion (item #16b).

The PRISMA 2020 guidelines emphasize transparency in the study selection process. In addition, the newly added

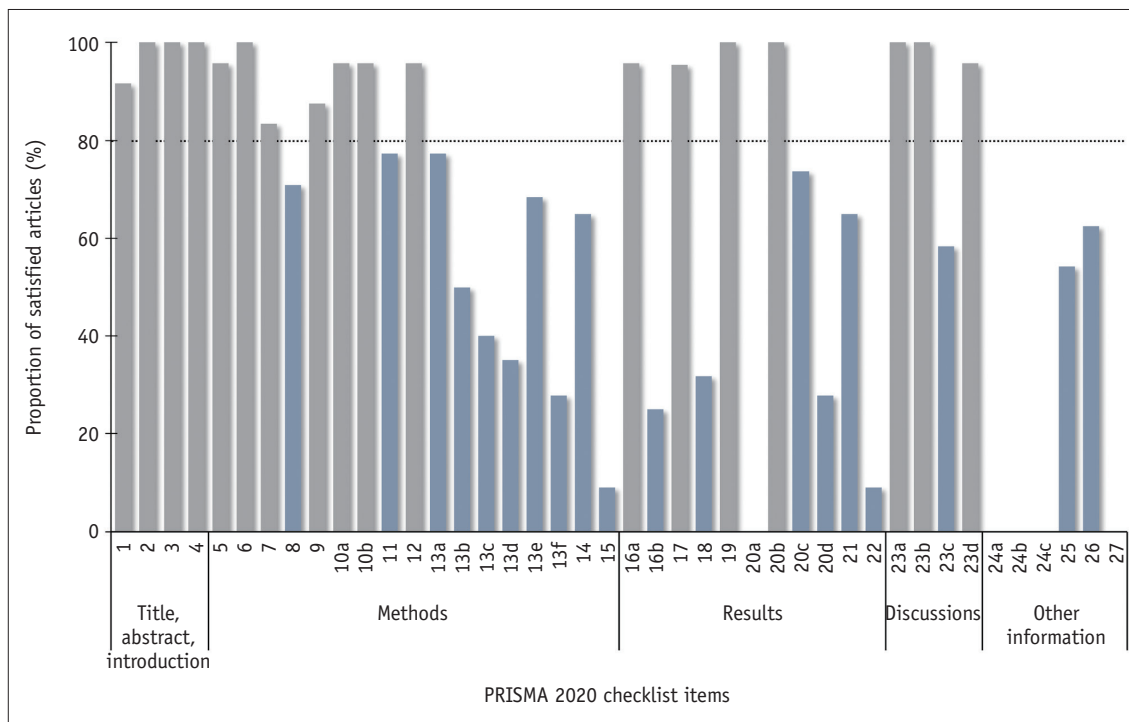


Fig. 4. Bar chart of the proportion of articles that satisfied each item of the PRISMA 2020 checklist. Blue bars indicate the items that were satisfied by fewer than 80% of the studies. Gray bars indicate the items that were satisfied by greater than 80% of the studies.

item #16b requires authors to provide the reasons for exclusion of potentially eligible studies [5].

Assessment of the Risk of Bias (Items #11, #18)

Four articles [24,29-31] did not evaluate the risk of bias in the studies and one article [33] did not report how many reviewers assessed the risk of bias. Various assessment tools were used in the remaining articles. For randomized controlled trials (RCTs), the Risk of Bias (RoB) tool or revised Jadad scale were implemented [32-34]. For non-RCTs, Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-2 [16,19-21,25,26,28], QUADAS [37], Risk of Bias Assessment tool for non-randomized studies [14,35], Risk of Bias in Non-randomized Studies of Interventions (ROBINS-I) [17], Newcastle Ottawa Scale [15,32,34], and the National Institute of Health (NIH) assessment tool were used [22,23]. Among the articles that evaluated the risk of bias, only seven provided the full assessment results of the individual studies [17,19,22,23,26,33,34].

Evaluation of the risk of bias in studies is essential for authors to understand result synthesis or to search for possible heterogeneity among the included studies, as well as for readers to evaluate the transparency of pooled results [5]. We advise authors to provide visual representation of assessment results for each study, rather than the overall

results of whole studies. Among the various options of assessment tools, the Cochrane guidelines [1] recommend the RoB, ROBINS-I, and QAUAS-2 as the preferred methods for assessing RCTs, non-RCTs on interventions, and diagnostic test accuracy (DTA) studies, respectively. Although there is no universally accepted tool for the evaluation of observational studies without interventions, Newcastle Ottawa Scale, the NIH assessment tool, and Joanna Briggs Institute critical appraisal checklists may be suitable options [38].

Synthesis of Results (Items #13a, #13b, #13c, #13d, #20a)

Among the articles that reported multiple pooled results, five [17,20-22,35] did not clearly report which studies were included for each outcome synthesis (item #13a). Ten articles [14,21,24-26,28,29,32,34,37] did not report how missing data were handled or how the data were converted for result synthesis (item #13b). Twelve articles [14,15,19,22-24,27,32-35,37] did not mention the methods used for the visual representation of the results from individual studies and syntheses, although forest plots were presented in the Results, except for one article [37] (item #13c). Thirteen articles with meta-analysis [14,15,19-23,25-28,31,35] did not report the rationale for choosing a specific statistical model (ex.

fixed- vs. random-effects model) (item #13d). Three studies [32-35] selected fixed- or random-effects models based on statistical values of study heterogeneity. Two studies [19,20] used the Dersimonian-Laird random-effects model for pooling rare events (e.g., complication rate). In the Results section, none of the studies reported a brief summary of the study characteristics and the risk of bias for each synthesis (item #20a).

The PRISMA 2020 checklist has elaborated the “synthesis of results” item to provide a more comprehensive evaluation regarding data preparation (item #13b), data visualization (item #13c), and statistical methods used for result synthesis (item #13d) [5]. When multiple results are pooled, authors are advised to cite the studies and report the number of the included studies for each outcome analysis (item #13a). When a meta-analysis is performed, authors should explain the rationale for choosing a statistical model. Choosing between fixed- and random-effects models should not be based on statistical methods for heterogeneity (i.e., Cochran’s Q-test or Higgins inconsistency index test) [1]; rather, it depends on the authors’ decision of whether effect sizes are truly identical between studies [1]. Therefore, the random-effects model is recommended when there is heterogeneity in study designs, which is very common when performing meta-analyses in the field of radiology.

Currently, the Cochrane guideline does not recommend a consensus method for result synthesis [1]. However, inverse-variance methods (including the DerSimonian and Laird method) should be avoided in meta-analyses of rare events [39]. Because these methods are based on the assumption of normal distribution of effect sizes, significant bias in pooled results may occur in meta-analyses of rare events [39,40]. In such cases, other methods such as the Peto method, Mantel-Haenszel method without zero-cell corrections, or generalized linear mixed models are preferred, although there is no generally accepted optimal method for dealing with rare events [39,41,42].

When reporting syntheses of multiple-effect sizes, authors should consider within-study covariance (i.e., correlation between outcomes). However, none of the articles included in this study considered within-study covariance despite the evident risk of correlation (e.g., pooling overall survival at multiple time points) [15,20-22,31,32,34]. The potential risks of correlation in these studies are summarized in Supplementary Table 3. When multiple-effect sizes are synthesized from data from the same participants, statistical dependency may occur and produce erroneous standard

errors in the pooled results [43]. Suggestions for managing within-study covariance are provided in the Supplement.

Additional Analysis (Items #13e, #13f, #20c, #20d)

Five articles [16,29,32,35,37] did not perform subgroup analysis or meta-regression. One article [34] did not mention the method used to explore study heterogeneity in the Materials and Methods section although subgroup analysis was provided in the Results section (item #13e, #20c). Thirteen articles [14,15,20-23,25,28,29,31,32,34,35] did not perform sensitivity analysis (item #13f, #20d).

The PRISMA 2020 guideline requires that authors perform subgroup analysis or meta-regression to evaluate the source of study heterogeneity, and sensitivity analysis to assess the robustness of the synthesized results.

Assessment of Non-Reporting Bias (Items #14, #21)

Seven articles [16,18,29,31,32,35,37] did not evaluate the non-reporting bias. Among the 13 articles that used funnel plots, four [15,21,24,27] did not further explore the source of bias, although asymmetry was observed. Two articles [33,34] performed a statistical test for funnel plot asymmetry, although fewer than 10 studies were included.

Non-reporting bias refers to the fact that reporting of the research findings is influenced by the *p* value and magnitude or direction of the results [44]. Although non-reporting is a broad term encompassing publication bias, time-lag bias, and selective non-reporting bias, publication bias has long been the focus of interest [1]. Funnel plots and statistical tests for asymmetry are frequently performed to evaluate non-reporting bias; however, the test for asymmetry has low statistical power and thus should not be used when fewer than 10 studies are included [45,46]. Moreover, it should be noted that asymmetry in funnel plots is not always due to non-reporting bias [46]. As a result, a contour-enhanced funnel plot may be preferred, because it may indicate whether the asymmetry is due to non-reporting bias or other factors [47]. Other potential sources of asymmetry include poor methodological quality in small-sized studies or true heterogeneity between studies [45]. For example, heterogeneity in the characteristics of study population or implementation of intervention between small vs. large-sized studies may cause asymmetry in the funnel plots [46]. When asymmetry is observed, authors should search for potential sources of asymmetry.

Assessment of the Certainty of Evidence (Items #15, #22)

Only two studies [24,33] used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach to evaluate the quality of evidence. The PRISMA 2020 guidelines included new items regarding the certainty of evidence for pooled results [5]. To evaluate the certainty of evidence, Cochrane has adopted the GRADE approach [1], which is composed of five domains: risk of bias, inconsistency, indirectness, imprecision, and publication bias [48]. By incorporating the evaluation for each domain, the final assessment for pooled results was classified into four categories: high, moderate, low, and very low quality. The certainty of evidence should be evaluated for each outcome, because the level of certainty often varies between outcomes [49]. Lastly, a “summary of findings” table should be presented by including the outcome of interest and its pooled result as well as the quality of evidence. Although the GRADE approach was first developed to evaluate studies on therapeutic intervention, it can be applied to DTA studies as well [50,51]. Supplementary Table 4 is an example of a “summary of findings” table, which may be produced using the GRADEpro GDT software (www.grade.org).

Provision of Limitations of the Study (Item #23c)

The PRISMA 2020 guidelines require that author provide limitations of not only the evidence but also the review process. Of all the included articles, 58% (14 out of 24) [16,17,19,20,22,24,26,27,30,32-34,36,37] described the limitations of the reviewing process in the Discussion section, which included: 1) limitation of search terms: “because we found HRQoL studies using the search term ‘quality of life,’ we might have missed studies using other terminology” [17], 2) limitation in the study selection process: “we included studies that were available only in the abstract form, and the reported data may not be as accurate and complete as those reported in the corresponding full text publication” [33], 3) limitations in data extraction: “there were limitations in extracting the exact survival data from the study regarding censored subjects and how these might affect the results” [22], and 4) inability to perform planned analysis due to lack of data: “because of the lack of sufficient data, we were unable to perform subgroup analyses to compare the effect of TACE plus RFA and surgical resection” [32].

Additional Information (Items #24–#27)

None of the studies reported the registration information (item #24) or which data in the review were publicly available to the readers (item #27). Eleven studies [14,16,17,23,24,28,29,32-35] did not report any financial or non-financial support (item #25), and nine studies [15,18,22,27,32-34,36,37] did not declare any competing interest for the authors (item #26).

The PRISMA 2020 guidelines require that authors provide registration information for the review (item #24a), a statement regarding accessibility of the registered protocol (item #24b), or any amendment made in the protocol (item #24c) [5]. PROSPERO is a database that authors can use to register their protocols [52]. Registering the protocol before conducting the systematic review enables the readers to evaluate whether the article properly followed the protocol and search for any differences between the pre-specified information and the finally reported information [5]. If the protocol was not registered, it should be stated so, and we suggest that authors discuss the potential limitations of not doing so. In addition, authors should report any financial or non-financial support received during the study, and their competing interests. Public sharing of the data used in the review is encouraged but is not widely performed in medical research [53]. Currently, there are several public data sharing platforms, such as Open Science Framework (<https://osf.io>) or Systematic Review Data Repository (<https://www.ahrq.gov/cpi/about/otherwebsites/srdr.ahrq.gov/index.html>).

DISCUSSION

Our study demonstrated that a substantial number of published systematic reviews with or without meta-analysis required further improvements to satisfy the PRISMA 2020 guidelines. These areas for improvement could be divided into eight domains for which thorough explanations and suggestions can be made: 1) assessment of the eligibility of potential articles, 2) assessment of the risk of bias, 3) synthesis of results, 4) additional analysis to explain study heterogeneity, 5) assessment of the non-reporting bias, 6) assessment of the certainty of evidence, 7) provision of limitations of the study, and 8) additional information such as protocol registration. In addition, for better quality abstracts, authors should report the exclusion criteria, the assessment tool for the risk of bias, the statistical methods, and limitations of the evidence.

Table 4. Checklist for Writing Good Quality Systematic Review and Meta-Analysis

Domain	PRISMA Items	Checklist	✓ or X
I. Assessing the eligibility of potential articles	#8, #16b	1. Report how many reviewers participated in article screening and whether they worked independently 2. Cite studies that seemed to meet inclusion criteria but excluded in the final stage and explain the reason for exclusion	
II. Assessing risk of bias in the studies	#11, #18	3. Report which method was used for evaluating the risk of bias*	
III. Synthesis of results	#13a, #13b, 13c, 13d, #20a	4. Present results regarding the assessment of the risk of bias for each included study 5. When performing syntheses of multiple outcomes, cite and report the number of studies included for each outcome synthesis 6. Consider within-study covariance when reporting syntheses of multiple outcomes (i.e., correlation between outcomes) [†] 7. Report how the missing data were handled or how the data were converted for synthesis 8. Report methods used for visual display of results of individual studies and syntheses (e.g., forest plot) 9. Report rationale of choosing one of the statistical models [‡] - Choosing between fixed vs. random-effects models should not be based on statistical values of study heterogeneity (i.e., Cochran's Q-test or Higgins inconsistency index test) - Check if reporting outcome is a rare event: in case of rare events, the DerSimonian-Laird method should be avoided 10. For each synthesis, present a brief summary of the characteristics and risk of bias regarding the contributing studies	
IV. Additional analysis	#13e, #13f, #20c, #20d	11. Present methods and results of additional analyses (i.e., subgroup analysis, meta-regression) to explain study heterogeneity 12. Present sensitivity analysis to assess robustness of the synthesized results	
V. Assessing non-reporting bias (i.e., publication bias)	#14, #21	13. Report any methods for evaluation of non-reporting bias (e.g., funnel plot) [§] - To interpret funnel plot, at least 10 studies needed	
VI. Assessing the certainty of evidence	#15, 22	14. Report certainty of evidence using the GRADE approach	
VII. Providing the limitation of the study	#23c	15. Provide any limitations of reviewing process as well as limitations of the evidence	
VIII. Additional information	#24–27	16. Provide registration information for the review - Register name (PROSPERO) and registration number - If unregistered, state that the review was not registered and indicate as a potential limitation 17. Report funding source and any competing interests	

*Based on the Cochrane guidelines, RoB, ROBINS-I, or QAUADAS-2 are preferred methods for assessing RCT, non-RCT of interventions, or DTA studies, respectively. †If one assumes a significant correlation between outcomes, a multivariate meta-analysis may be a preferred method over multiple independent univariate meta-analysis. There are a few other ways to deal with correlation between outcomes; however, these are beyond the scope of this paper. Refer to López-López et al. for more information [43]. ‡The Cochrane Handbook for Systematic Reviews of Interventions does not suggest universally preferred methods for result synthesis. However, the inverse-variance methods (including the DerSimonian and Laird method) should be avoided with rare events meta-analysis (e.g., complication rates, technical failure rates) [39]. Regarding the meta-analysis of DTA, the bivariate model and HSROC model are recommended methods [10]. §It should be noted that asymmetry in funnel plot is not always due to non-reporting bias. As a result, a contour-enhanced funnel plot may be preferred because it can suggest that whether the asymmetry is due to non-reporting bias or other factors. DTA = diagnostic test accuracy, HSROC = hierarchical summary receiver operating characteristic, QAUADAS = Quality Assessment of Diagnostic Accuracy Studies, RCT = randomized controlled trial, RoB = Risk of Bias in randomized trials, ROBINS-I = Risk of Bias in Non-randomized Studies of Interventions

Table 5. Recommended Statistical Methods for Meta-Analysis

	Meta-Analysis of Usual Proportion	Meta-Analysis of Diagnostic Accuracy Test
Result synthesis	Fixed-effects model	Inverse variance method Mantel-Haenszel method Peto method Not recommended
	Random-effects model	Dersimonian Laird method REML method Paule-Mandel method Bivariate model HSROC model
Non-reporting/publication bias assessment tool	Funnel plot Begg's test or Egger's test	Deeks' funnel plot Deeks' asymmetry test
Risk of bias assessment tool*	RCT: RoB 2 tool Non-RCT: ROBINS-I tool	QUADAS-2 tool
Evaluation of study heterogeneity	Chi-squared test (Cochrane Q statistics) Higgins I ² statistic	Chi-squared test (Cochrane Q statistics) Higgins I ² statistic Analysis of threshold effect - Visual evaluation of coupled forest plot - Spearman correlation analysis between sensitivity and specificity
Additional analysis for study heterogeneity	Subgroup analysis or meta-regression Sensitivity analysis	
Certainty of evidence evaluation	GRADE approach	

*The Cochrane Handbook for Systematic Reviews of Interventions recommends the RoB-2 tool and ROBINS-I tool as bias assessment tools in RCT and non-RCT, respectively. GRADE = Grading of Recommendations, Assessment, Development and Evaluations, HSROC = hierarchical summary receiver operating characteristic, QUADAS = Quality Assessment of Diagnostic Accuracy Studies, REML = restricted maximum likelihood, RoB = Risk of Bias in randomized trials, ROBINS-I = Risk of Bias in Non-randomized Studies of Interventions

Based on our results, we developed a double-check list consisting of the items in PRISMA 2020 guidelines that had been frequently missed in published articles (Table 4). In the checklist, we made specific suggestions for each domain and provided further comments regarding the errors in statistical analyses identified in some published articles (e.g., determining fixed- vs. random-effects model based on statistical values of study heterogeneity). To help authors properly utilize the statistical models and assessment tools, we summarized the recommended methods in Table 5. These recommended methods are mainly based on the Cochrane Handbook for Systematic Reviews of Interventions and previous guideline articles for DTA studies [1,9,11,54-56].

A substantial proportion of meta-analysis in the field of radiology is DTA research. In 2018, an extension of PRISMA 2009 statement has been developed for systematic reviews of DTA studies (PRISMA-DTA statement) [57]. When compared to PRISMA 2020 statement, PRISMA-DTA statement requires specific information regarding index test, including the clinical role of the index test and 2 x 2 data (true positive, false positive, false negative, and true negative) for each study. However, the PRISMA 2020 statement provides more comprehensive checklists in the remaining fields such

as data extraction, data handling, statistical analysis, and result presentation. Thus, authors who conduct systematic reviews of DTA studies should follow the PRISMA 2020 statement in general and refer to the PRISMA-DTA statement for DTA specific requirements [5].

There are several limitations to our study. First, the articles used in our study were from a single journal, which may impair extrapolation of the results. However, the *Korean Journal of Radiology* may serve as a proper representative sample given its reputation in the field of radiology, nuclear medicine, and imaging (rank: 36 out of 452 journals in Scopus) and wide coverage of topics as a general journal of radiology. Second, detailed statistical background for each method in meta-analyses was not provided. In addition, we did not cover advanced techniques in meta-analysis, such as individual participant data meta-analysis or network meta-analysis [58,59]. Third, while our study focused on the reporting qualities of systematic reviews and meta-analysis, the reporting quality does not necessarily indicate the quality of the study itself. Proper research questions based on the population, intervention, comparison, outcome (PICO) framework and the purpose of conducting the systematic review or meta-analysis should be well

established beforehand [1]. Despite these limitations, our study clearly identified which items should be improved for high-quality systematic review articles. Authors and reviewers who are interested in systematic reviews or meta-analyses should be familiar with the PRISMA 2020 statement. Our checklists may help authors to identify which items of the PRISMA 2020 statement should be reinforced prior to submission.

Supplement

The Supplement is available with this article at <https://doi.org/10.3348/kjr.2021.0808>.

Availability of Data and Material

The datasets generated or analyzed during the study are available from the corresponding author on reasonable request.

Conflicts of Interest

Chong Hyun Suh who is on the editorial board of the *Korean Journal of Radiology* was not involved in the editorial evaluation or decision to publish this article. All remaining authors have declared no conflicts of interest.

Author Contributions

Conceptualization: Chong Hyun Suh, Kyung Won Kim. Data curation: Ho Young Park. Formal analysis: Ho Young Park, Chong Hyun Suh. Funding acquisition: Chong Hyun Suh. Methodology: Ho Young Park, Pyeong Hwa Kim, Chong Hyun Suh. Project administration: Sungmin Woo, Chong Hyun Suh. Supervision: Kyung Won Kim. Writing—original draft: Ho Young Park. Writing—review & editing: Sungmin Woo, Pyeong Hwa Kim, Chong Hyun Suh, Kyung Won Kim.

ORCID iDs

Ho Young Park

<https://orcid.org/0000-0003-2318-9806>

Chong Hyun Suh

<https://orcid.org/0000-0002-4737-0530>

Sungmin Woo

<https://orcid.org/0000-0001-8459-8369>

Pyeong Hwa Kim

<https://orcid.org/0000-0003-4276-8803>

Kyung Won Kim

<https://orcid.org/0000-0002-1532-5970>

Funding Statement

This research was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI18C2383).

REFERENCES

- Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. Cochrane handbook for systematic reviews of interventions version 6.2 (updated February 2021). Training. cochrane.org Web site. <https://training.cochrane.org/handbook>. Accessed August 8, 2021
- Booth A, Clarke M, Ghera D, Moher D, Petticrew M, Stewart L. An international registry of systematic-review protocols. *Lancet* 2010;377:108-109
- Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med* 2009;6:e1000097
- Nawijn F, Ham WHW, Houwert RM, Groenwold RHH, Hietbrink F, Smeeing DPJ. Quality of reporting of systematic reviews and meta-analyses in emergency medicine based on the PRISMA statement. *BMC Emerg Med* 2019;19:19
- Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021;372:n160
- Tunis AS, McInnes MD, Hanna R, Esmail K. Association of study quality with completeness of reporting: have completeness of reporting and quality of systematic reviews and meta-analyses in major radiology journals changed since publication of the PRISMA statement? *Radiology* 2013;269:413-426
- Wang P, Guo YM, Liu M, Qiang YQ, Guo XJ, Zhang YL, et al. A meta-analysis of the accuracy of prostate cancer studies which use magnetic resonance spectroscopy as a diagnostic tool. *Korean J Radiol* 2008;9:432-438
- Yu SH, Kim CB, Park JW, Kim MS, Radosevich DM. Ultrasonography in the diagnosis of appendicitis: evaluation by meta-analysis. *Korean J Radiol* 2005;6:267-277
- Kim KW, Lee J, Choi SH, Huh J, Park SH. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-part I. general guidance and tips. *Korean J Radiol* 2015;16:1175-1187
- Lee J, Kim KW, Choi SH, Huh J, Park SH. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers-part II. statistical methods of meta-analysis. *Korean J Radiol* 2015;16:1188-1196
- Suh CH, Park SH. Successful publication of systematic review and meta-analysis of studies evaluating diagnostic test

- accuracy. *Korean J Radiol* 2016;17:5-6
12. Kim SY, Chung HW, Oh TS, Lee JS. Practical guidelines for ultrasound-guided core needle biopsy of soft-tissue lesions: transformation from beginner to specialist. *Korean J Radiol* 2017;18:361-369
 13. Oh SW, Cheon GJ. Prostate-specific membrane antigen PET imaging in prostate cancer: opportunities and challenges. *Korean J Radiol* 2018;19:819-831
 14. Cho SJ, Kim HS, Suh CH, Park JE. Radiological recurrence patterns after bevacizumab treatment of recurrent high-grade glioma: a systematic review and meta-analysis. *Korean J Radiol* 2020;21:908-918
 15. Choi SH, Kim JW, Kim JH, Kim KW. Efficacy and safety of microwave ablation for malignant renal tumors: an updated systematic review and meta-analysis of the literature since 2012. *Korean J Radiol* 2018;19:938-949
 16. Chung SR, Choi YJ, Suh CH, Lee JH, Baek JH. Diffusion-weighted magnetic resonance imaging for predicting response to chemoradiation therapy for head and neck squamous cell carcinoma: a systematic review. *Korean J Radiol* 2019;20:649-661
 17. Kang D, Shim S, Cho J, Lim HK. Systematic review of studies assessing the health-related quality of life of hepatocellular carcinoma patients from 2009 to 2018. *Korean J Radiol* 2020;21:633-646
 18. Kim DW, Jang HY, Kim KW, Shin Y, Park SH. Design characteristics of studies reporting the performance of artificial intelligence algorithms for diagnostic analysis of medical images: results from recently published papers. *Korean J Radiol* 2019;20:405-410
 19. Kim DW, Suh CH, Kim KW, Pyo J, Park C, Jung SC. Technical performance of two-dimensional shear wave elastography for measuring liver stiffness: a systematic review and meta-analysis. *Korean J Radiol* 2019;20:880-893
 20. Kim JY, Suh YJ, Han K, Choi BW. Reliability of coronary artery calcium severity assessment on non-electrocardiogram-gated CT: a meta-analysis. *Korean J Radiol* 2021;22:1034-1043
 21. Kim JY, Suh YJ, Han K, Kim YJ, Choi BW. Cardiac CT for measurement of right ventricular volume and function in comparison with cardiac MRI: a meta-analysis. *Korean J Radiol* 2020;21:450-461
 22. Kim PH, Choi SH, Kim JH, Park SH. Comparison of radioembolization and sorafenib for the treatment of hepatocellular carcinoma with portal vein tumor thrombosis: a systematic review and meta-analysis of safety and efficacy. *Korean J Radiol* 2019;20:385-398
 23. Kim PH, Kim M, Suh CH, Chung SR, Park JE, Kim SC, et al. Neuroimaging findings in patients with COVID-19: a systematic review and meta-analysis. *Korean J Radiol* 2021;22:1875-1885
 24. Kim PH, Suh CH, Kim HS, Kim KW, Kim DY, Lee EQ, et al. Immune checkpoint inhibitor with or without radiotherapy in melanoma patients with brain metastases: a systematic review and meta-analysis. *Korean J Radiol* 2021;22:584-595
 25. Kim TH, Woo S, Han S, Suh CH, Ghafoor S, Hricak H, et al. The diagnostic performance of the length of tumor capsular contact on MRI for detecting prostate cancer extraprostatic extension: a systematic review and meta-analysis. *Korean J Radiol* 2020;21:684-694
 26. Ko MJ, Park DA, Kim SH, Ko ES, Shin KH, Lim W, et al. Accuracy of digital breast tomosynthesis for detecting breast cancer in the diagnostic setting: a systematic review and meta-analysis. *Korean J Radiol* 2021;22:1240-1252
 27. Liao XL, Wei JB, Li YQ, Zhong JH, Liao CC, Wei CY. Functional magnetic resonance imaging in the diagnosis of locally recurrent prostate cancer: are all pulse sequences helpful? *Korean J Radiol* 2018;19:1110-1118
 28. Lim SJ, Kim M, Suh CH, Kim SY, Shim WH, Kim SJ. Diagnostic yield of diffusion-weighted brain magnetic resonance imaging in patients with transient global amnesia: a systematic review and meta-analysis. *Korean J Radiol* 2021;22:1680-1689
 29. Park SH, Cho SH, Choi SH, Jang JK, Kim MJ, Kim SH, et al. MRI assessment of complete response to preoperative chemoradiation therapy for rectal cancer: 2020 guide for practice from the Korean Society of Abdominal Radiology. *Korean J Radiol* 2020;21:812-828
 30. Suh CH, Jung SC, Kim B, Cho SJ, Woo DC, Oh WY, et al. Neuroimaging in randomized, multi-center clinical trials of endovascular treatment for acute ischemic stroke: a systematic review. *Korean J Radiol* 2020;21:42-57
 31. Suh CH, Kim HS, Jung SC, Choi CG, Kim SJ, Kim KW. Optimized image-based surrogate endpoints in targeted therapies for glioblastoma: a systematic review and meta-analysis of phase III randomized controlled trials. *Korean J Radiol* 2020;21:471-482
 32. Wang WD, Zhang LH, Ni JY, Jiang XY, Chen D, Chen YT, et al. Radiofrequency ablation combined with transcatheter arterial chemoembolization therapy versus surgical resection for hepatocellular carcinoma within the Milan criteria: a meta-analysis. *Korean J Radiol* 2018;19:613-622
 33. Wang X, Hu Y, Ren M, Lu X, Lu G, He S. Efficacy and safety of radiofrequency ablation combined with transcatheter arterial chemoembolization for hepatocellular carcinomas compared with radiofrequency ablation alone: a time-to-event meta-analysis. *Korean J Radiol* 2016;17:93-102
 34. Zhu ZX, Liao MH, Wang XX, Huang JW. Transcatheter arterial chemoembolization plus ¹³¹I-labelled metuximab versus transcatheter arterial chemoembolization alone in intermediate/advanced stage hepatocellular carcinoma: a systematic review and meta-analysis. *Korean J Radiol* 2016;17:882-892
 35. Kim HJ, Cho SJ, Baek JH. Comparison of thermal ablation and surgery for low-risk papillary thyroid microcarcinoma: a systematic review and meta-analysis. *Korean J Radiol* 2021;22:1730-1741
 36. Kang TW, Rhim H, Lee MW, Kim YS, Choi D, Lim HK. Terminology and reporting criteria for radiofrequency ablation of tumors in the scientific literature: systematic review

- of compliance with reporting standards. *Korean J Radiol* 2014;15:95-107
37. Kim G, Cho YZ, Baik SK, Kim MY, Hong WK, Kwon SO. The accuracy of ultrasonography for the evaluation of portal hypertension in patients with cirrhosis: a systematic review. *Korean J Radiol* 2015;16:314-324
 38. Ma LL, Wang YY, Yang ZH, Huang D, Weng H, Zeng XT. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? *Mil Med Res* 2020;7:7
 39. Efthimiou O. Practical guide to the meta-analysis of rare events. *Evid Based Ment Health* 2018;21:72-76
 40. Thorlund K, Wetterslev J, Awad T, Thabane L, Gluud C. Comparison of statistical inferences from the DerSimonian-Laird and alternative random-effects model meta-analyses - an empirical assessment of 920 Cochrane primary outcome meta-analyses. *Res Synth Methods* 2011;2:238-253
 41. Bradburn MJ, Deeks JJ, Berlin JA, Russell Localio A. Much ado about nothing: a comparison of the performance of meta-analytical methods with rare events. *Stat Med* 2007;26:53-77
 42. Stijnen T, Hamza TH, Ozdemir P. Random effects meta-analysis of event outcome in the framework of the generalized linear mixed model with applications in sparse data. *Stat Med* 2010;29:3046-3067
 43. López-López JA, Page MJ, Lipsey MW, Higgins JPT. Dealing with effect size multiplicity in systematic reviews and meta-analyses. *Res Synth Methods* 2018;9:336-351
 44. McGauran N, Wieseler B, Kreis J, Schüler YB, Kölsch H, Kaiser T. Reporting bias in medical research - a narrative review. *Trials* 2010;11:37
 45. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629-634
 46. Sterne JA, Sutton AJ, Ioannidis JP, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ* 2011;343:d4002
 47. Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Contour-enhanced meta-analysis funnel plots help distinguish publication bias from other causes of asymmetry. *J Clin Epidemiol* 2008;61:991-996
 48. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;336:924-926
 49. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. *J Clin Epidemiol* 2011;64:401-406
 50. Gopalakrishna G, Mustafa RA, Davenport C, Scholten RJ, Hyde C, Brozek J, et al. Applying grading of recommendations assessment, development and evaluation (GRADE) to diagnostic tests was challenging but doable. *J Clin Epidemiol* 2014;67:760-768
 51. Schünemann HJ, Mustafa R, Brozek J, Santesso N, Alonso-Coello P, Guyatt G, et al. GRADE guidelines: 16. GRADE evidence to decision frameworks for tests in clinical practice and public health. *J Clin Epidemiol* 2016;76:89-98
 52. Booth A, Clarke M, Dooley G, Ghersi D, Moher D, Petticrew M, et al. The nuts and bolts of PROSPERO: an international prospective register of systematic reviews. *Syst Rev* 2012;1:2
 53. Naudet F, Sakarovitch C, Janiaud P, Cristea I, Fanelli D, Moher D, et al. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in the BMJ and PLOS Medicine. *BMJ* 2018;360:k400
 54. McGrath TA, Alabousi M, Skidmore B, Korevaar DA, Bossuyt PMM, Moher D, et al. Recommendations for reporting of systematic reviews and meta-analyses of diagnostic test accuracy: a systematic review. *Syst Rev* 2017;6:194
 55. McGrath TA, McInnes MD, Korevaar DA, Bossuyt PM. Meta-analyses of diagnostic accuracy in imaging journals: analysis of pooling techniques and their effect on summary estimates of diagnostic accuracy. *Radiology* 2016;281:78-85
 56. Takwoingi Y, Leeftang MM, Deeks JJ. Empirical evidence of the importance of comparative studies of diagnostic test accuracy. *Ann Intern Med* 2013;158:544-554
 57. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA* 2018;319:388-396
 58. Owen RK, Cooper NJ, Quinn TJ, Lees R, Sutton AJ. Network meta-analysis of diagnostic test accuracy studies identifies and ranks the optimal diagnostic tests and thresholds for health care policy and decision-making. *J Clin Epidemiol* 2018;99:64-74
 59. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010;340:c221