Observational Study

Medicine® OPEN

Characteristics, trend, and methodological quality of systematic reviews and meta-analyses in nuclear medicine

A bibliometric analysis of studies published between 2005 and 2016

Jung Ui Hong, MD[a], Jun Ho Kim, MD[a,*], Kyung Hee Lee, MD[a], Minkyung Lee, MD[b], In Young Hyun, MD[b], Soon Gu Cho, MD[a], Yeo Ju Kim, MD[a], Ha Young Lee, MD[a], Ga Ram Kim, MD[a]

Abstract

To evaluate the characteristics, trend, and quality of systematic reviews and meta-analyses in nuclear medicine.

We performed a PubMed search to identify systematic reviews and meta-analyses published between 2005 and 2016 in the field of nuclear medicine. The following data were extracted: journal name, impact factor, type of study, topics with cancer type, imaging modalities, authors (number, country, affiliation, presence of nuclear medicine specialists and statisticians, discordance between the first and corresponding authors), funding, methodological quality, methods used for quality assessment, and statistical methods.

We included 185 nuclear medicine articles. Meta-analyses (n=164; 88.6%) were published about 7 times more frequently than systematic reviews. Oncology was the most commonly studied topic (n=125, 67.6%). The first authors were most frequently located in China (n=73; 39.5%). PET was the most commonly used modality (n=150; 81.1%). Both the number of authors and the ratio of discordance between the first and corresponding authors tended to progressively increase over time.

The mean AMSTAR score increased over time (5.77 in 2005–2008, 6.71 in 2009–2012, and 7.44 in 2013–2016). The proportion of articles with quality assessment increased significantly (20/26 in 2005–2008, 54/65 in 2009–2012, and 79/94 in 2013–2016). The most commonly used assessment tool was quality assessment of diagnostic accuracy studies (n=85; 54.9%).

The number and quality of systematic reviews and meta-analyses in nuclear medicine have significantly increased over the review period; however, the quality of these articles varies. Efforts to overcome specific weaknesses of the methodologies can provide opportunities for quality improvement.

Abbreviations: AMSTAR = a measurement tool to assess systematic reviews, FDG = 2-fluoro-2-deoxy-D-glucose, HSROC = hierarchical summary ROC curve, IF = impact factor, PET = positron emission tomography, PRISMA = the preferred reporting items for systematic reviews and meta-analyses, QUADAS = quality assessment of diagnostic accuracy studies, SCI = science citation index, SCIE = science citation index expanded, SPECT = single-photon emission computed tomography, SROC = summary receiver operating characteristic curve.

Keywords: bibliometrics, meta-analysis, nuclear medicine, quality assessment, systematic review

Editor: Antonio Palazón-Bru.

This work was supported by Inha University Research Grant (2019).

The study sponsor had no involvement in the conduct of the study or writing of the article.

The authors have no conflicts of interest to declare.

Supplemental Digital Content is available for this article.

[a] Department of Radiology, [b] Department of Nuclear Medicine, Inha University Hospital, Inha University School of Medicine, Jung-gu, Incheon, Korea.

[*] Correspondence: Jun Ho Kim, Department of Radiology, Inha University Hospital, Inha University School of Medicine, Inhang-ro 27, Jung-gu, Incheon 22332, South Korea (e-mail: jhkim7227@gmail.com).

Medicine (2019) 98:21(e15785)

Received: 19 November 2018 / Received in final form: 24 April 2019 / Accepted: 27 April 2019

http://dx.doi.org/10.1097/MD.0000000000015785

1. Introduction

A systematic review is a strong scientific methodology that can provide evidence-based summary when results of individual relevant studies regarding a particular research topic are inconclusive. A meta-analysis is part of the systematic review, and the former applies statistical methods to synthesize the data from multiple relevant studies to provide a summary estimate.[1,2] Recently, the need for systematic reviews and meta-analyses in all fields of research, including the field of diagnostic testing, has increased in proportion to the significant increase in the number of published primary scientific research articles and the importance of evidence-based medicine.[1–3] Appropriately conducted systematic reviews and meta-analyses are essential to ensure transparency, complete reporting of data, and minimization of bias. There are some assessment tools that can assess the quality of systematic reviews and meta-analyses, such as a measurement tool to assess systematic reviews (AMSTAR).[2–4]

Bibliometric analysis provides a general picture of the trend and characteristics of the articles within a particular research field. In addition, the analysis is important in guiding the direction of the specific research field.[3–5]

Recently, technological advancements that include hybrid imaging and new radiopharmaceuticals have allowed the growth of application of nuclear medicine procedures.[6] Previous bibliometric analyses have shown that the number of publications in the field of nuclear medicine has grown rapidly over last decade in proportion to continued growth of the clinical importance of nuclear medicine.[7,8] However, no bibliometric analysis has been performed to investigate the quality and the trend of systematic reviews and meta-analyses in the field of nuclear medicine.

Therefore, the purpose of this study was to assess the quality and evaluate the trend and characteristics of systematic reviews and meta-analyses in the field of nuclear medicine over the last decade or more (2005–2016).

2. Methods

Our study was a retrospective bibliometric analysis that did not involve human subjects, and approval was not required from our institutional review board.

2.1. Search strategy

We searched the National Library of Medicine's PubMed database for articles that were published between January 2005 and December 2016 to identify systematic reviews and meta-analyses in the field of nuclear medicine. We used the following previously reported search filters: ("diagnostic test accuracy" or dta[tiab] or sensitivity) and (specificity[mh] or specificit*[tw] or "false negative"[tw] or accuracy[tw]).[9]

We searched the journals that were included in the 2016 ranking by Clarivate Analytics. The articles published in journals that were not indexed in the science citation index (SCI)/science citation index expanded (SCIE) were excluded from the analysis. The search was restricted to articles in radiology, nuclear medicine, and medical imaging as defined by Clarivate Analytics.

We excluded articles that were not related to nuclear medicine based on their titles and abstracts. Full-text articles were independently reviewed by 2 reviewers (JHK and JUH). The disagreement between the 2 reviewers was resolved through a discussion with a third reviewer (KHL). After review of full texts, we also excluded original research, review articles, and guidelines.

2.2. Data extraction

The aforementioned 2 reviewers extracted the following information from all articles: journal of publication, impact factor (IF), type of study (systematic review or meta-analysis), year of publication, topics (oncology, such as various types of primary, metastatic cancers, including primary unknown cancers [breast, central nervous system, gynecologic, head and neck, hematologic, liver, lung, pancreatic, prostate, stomach, colorectal/anal/rectal, and others], cardiovascular diseases, inflammation/infection, such as various infections and inflammatory responses, autoimmune diseases, and fever of unknown origin, endocrinology, neurology, and radiopharmaceuticals [tracers]), imaging modalities (positron emission tomography [PET], including PET alone, PET-CT, and PET-MR, single-photon emission computed tomography [SPECT], and scintigraphy), and tracers (2-fluoro-2-deoxy-D-glucose [FDG] and others [including the use of multiple tracers]). In addition, 2 reviewers recorded the number of authors (<4, 4–7, or >7), Affiliation Department of the first and corresponding authors (nuclear medicine, radiology, medicine or surgery, statistics, or others), presence of a nuclear medicine specialist and a statistician among the authors, discordance between the first and corresponding authors, source of funding, country of origin of the first author, methodological quality, methods of quality assessment, methods of statistical analysis (univariate, bivariate, summary receiver operating characteristic [SROC] curve, or hierarchical summary ROC [HSROC] curve analysis or a combination of these methods), meta-analysis outcome, and the reasons that a meta-analysis was not performed in a systematic review article.

If an epidemiologist was included among the authors, he or she was reported as if from department of statistics. If the first author was affiliated with more than 1 country, the study's country of origin was determined by checking the corresponding author's country of origin.

The major guidelines for the systematic reviews and meta-analyses were from the preferred reporting items for systematic reviews and meta-analyses (PRISMA) statement, which was published in 2009.[10] The time periods before and after 2009 were called the first and second periods. To evaluate the recent trends, the second period was divided into 2 even periods: 2009 to 2012 and 2013 to 2016. Finally, the 3 periods of review were 2005 to 2008, 2009 to 2012, and 2013 to 2016.

2.3. Quality assessment of methodology

The 2 investigators (JHK and JUH) used the AMSTAR checklist to evaluate the quality of the systematic reviews and meta-analyses

All articles that were included were assessed, and a response to each question was recorded as "yes," "no," "cannot answer," or "not applicable." If the answer to a given question was "yes" or "not applicable," a point was assigned to that response; the total score for all 11 questions was 11 points. The response to items 9 and 10 of the AMSTAR checklist for the systematic review was "not applicable" and was assigned a point. In addition, if the response for item 7 was "no," a response of "yes" was not given for item 8. The 2 reviewers discussed the first 10 papers to identify the differences in their scoring methods before all the articles were reviewed. Any disagreements were discussed in a consensus meeting until a consensus was reached. In addition, the score for item 4 has been observed to be very low in previous studies.[2–4] If an article was not assigned a point on item 4, we investigated for the reason. We divided the reasons into 5 categories: language restriction, exclusion of small-sized studies, exclusion of conference papers, exclusion of low-quality articles based on the results of quality assessment, and other reasons. An article could have more than 1 reason.

2.4. Data analysis

The studies were divided into 3 groups based on the period of review: 2005 to 2008, 2009 to 2012, and 2013 to 2016. Continuous variables were compared among these groups using ANOVA. With respect to the categorical variables, the Cochran-Armitage test for trend in proportions was used; this was combined with the linear by linear association test where appropriate. Statistical analyses were performed using either

SPSS for Windows (SPSS version 19.0; Chicago, IL) or dBSTAT for Windows (dBSTAT version 5.0; Seoul, South Korea). *P*-values <.05 were considered statistically significant.

Agreement between the 2 investigators in the AMSTAR assessment of the reviews of the first 10 papers was assessed using the κ static. A κ value of 0 to 0.20 indicated slight agreement; 0.21 to 0.40 showed fair agreement, 0.41 to 0.60 demonstrated moderate agreement; 0.61 to 0.80 indicated substantial agreement; and 0.81 to 1.00 showed almost perfect agreement.

3. Results

3.1. Results of search

A total of 10,710 articles that were meta-analyses or systematic reviews in the field of diagnostic test accuracy that were published between 2005 and 2016 were retrieved from the PubMed database. Among these articles, 1071 articles were published in radiology, nuclear medicine, and medical imaging journals that were indexed by Clarivate Analytics and had a 2016 ranking. After a review of the titles and abstracts, it was determined that only 260 articles were related to nuclear medicine. Seventy-five articles were excluded after a full-text review: 52 review articles, 15 original research articles, 1 guideline, 1 case report, and 6 editorials. Finally, 185 articles were included and assessed in the present study (Fig. 1).

Assessment of the included 185 articles showed that the *Nuclear Medicine Communications* (IF = 1.472) was the most commonly used journal (29/185; 15.68%), followed by the *European Journal of Nuclear Medicine and Molecular Imaging* (20/185; 10.81%; IF = 7.277), the *Journal of Nuclear Medicine* (18/185; 9.73%; IF = 6.646), and others (Table 1).

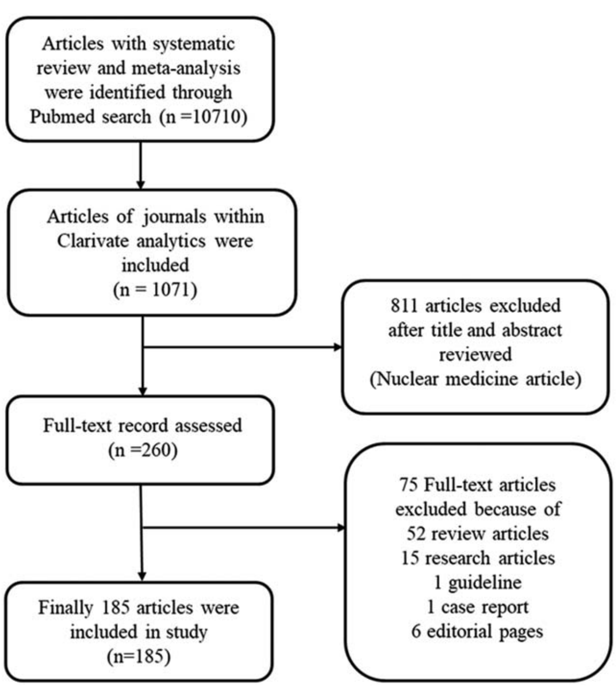


Figure 1. Results of the search: In our PubMed search, we identified 10,710 meta-analysis and systematic review articles. Finally, 185 articles were included and were assessed in terms of their characteristics, trend, and quality.

Table 1

Number of articles and the impact factor of the journals that published meta-analyses and systematic reviews between 2005 and 2016.

Journal name	Impact factor (2016)	5-yr impact factor	Number of articles
Nucl Med Commun.	1.472	1.489	29 (16)
Eur J Nucl Med Mol Imaging.	7.277	6.218	21 (11)
J Nucl Med.	6.646	6.459	18 (10)
Eur J Radiol.	2.462	2.507	15 (8)
Clin Nucl Med.	4.563	3.725	12 (6)
Radiology.	7.296	7.648	11 (6)
Acta Radiol.	2.011	1.751	10 (5)
Acad Radiol.	2.128	2.057	6 (3)
Ann Nucl Med	1.396	1.619	6 (3)
Eur Radiol.	3.967	4.277	5 (3)
J Nucl Cardiol.	3.93	2.996	5 (3)
Am J Roentgenol.	2.778	3.105	4 (2)
Rev Esp Med Nucl.	0.951	0.942	4 (2)
Am J Neuroradiol	3.55	3.888	3 (2)
Mol Imaging Biol.	3.466	2.726	3 (2)
Br J Radiol.	2.05	2.235	3 (2)
Skeletal Radiol.	1.737	1.913	3 (2)
J Med Imaging Radiat Oncol.	1.189	1.36	3 (2)
Circ Cardiovasc Imaging.	6.803	7.069	2 (1)
Radiother Oncol.	4.328	4.687	2 (1)
Q J Nucl Med Mol Imaging.	2.481	2.16	2 (1)
Nucl Med Biol.	2.426	2.318	2 (1)
Cancer Imaging.	2.404	2.282	2 (1)
J Neuroimaging.	1.664	1.637	2 (1)
Pediatr Radiol.	1.465	1.545	2 (1)
Nuklearmedizin.	1.087	1.153	2 (1)
Semin Nucl Med.	3.63	3.5	1 (1)
J Cardiovasc Comput Tomogr.	3.185	3.239	1 (1)
J Am Coll Radiol.	2.993	2.813	1 (1)
Abdom Imaging.	1.842	1.93	1 (1)
Radiol Oncol.	1.681	1.723	1 (1)
Dentomaxillofac Radiol.	1.594	1.698	1 (1)
Hell J Nucl Med.	1.048	0.991	1 (1)
Clin Imaging.	1.015	1.036	1 (1)
Total			185 (100)

Data are number of articles, with percentage in parentheses.

3.2. Characteristics and trend of articles

The characteristics of the included papers are presented in Table 2. During the review period, meta-analyses (n = 164; 88.6%) were published about 8 times more frequently than systematic reviews were (n = 21; 11.7%).

Most studies were on the oncology topic (n = 126; 68.1%). Among them, the number of studies on prostate cancers increased significantly during each period (*P* = .29).

PET imaging (PET alone, PET-CT, or PET-MR) was the most commonly used imaging modality in research (n = 150; 81.1%), followed at a distance by scintigraphy (n = 37; 17.6%), and SPECT (n = 30; 16.2%). The most commonly used radiotracer for PET was FDG (n = 134; 89.3%).

The number of authors tended to increase over the review period; however, this change was not statistically significant. On the other hand, the proportion of discordance between the first and corresponding authors increased significantly (*P* = .028). Most of both the first (n = 96, 51.9%) and the corresponding (n = 96; 51.9%) authors were affiliated with the department of nuclear medicine, followed by the department of medicine or

Table 2

Characteristics of the meta-analyses and systematic reviews published between 2005 and 2016 in Nuclear Medicine Journals.

	Period			
Characteristic	**2005–2008 (26 articles)**	**2009–2012 (65 articles)**	**2013–2016 (94 articles)**	**Total (185 articles)**
Type of research				
Meta-analysis	24 (92.3)	60 (92.3)	80 (85.1)	164 (88.6)
Systematic review	2 (7.7)	5 (7.7)	14 (14.9)	21 (11.4)
Topics				
Oncology	17 (65.4)	48 (73.8)	61 (64.9)	126 (68.1)
Cardiovascular disease	5 (19.2)	6 (9.2)	9 (9.6)	20 (10.8)
Inflammation/infection	4 (15.4)	5 (7.7)	10 (10.6)	19 (10.3)
Endocrinology	0 (0)	2 (3.1)	5 (5.3)	7 (3.8)
Neurology	0 (0)	3 (4.6)	3 (3.2)	6 (3.2)
Radiophamaceutical (tracer)	0 (0)	0 (0)	1 (1.1)	1 (0.5)
Others	0 (0)	1 (1.5)	5 (5.6)	6 (3.2)
Cancer type[†]				
Breast	2 (11.8)	4 (8.3)	5 (8.2)	11 (8.7)
Central nervous system	1 (5.9)	1 (2.1)	4 (6.6)	6 (4.8)
Gynecologic[‡]	2 (11.8)	4 (8.3)	2 (3.3)	8 (6.3)
Head and neck	1 (5.9)	4 (8.3)	9 (14.8)	14 (11.1)
Hematologic	2 (11.8)	4 (8.3)	2 (3.3)	8 (6.3)
Liver	0 (0.0)	2 (4.2)	0 (0.0)	2 (1.6)
Lung	2 (11.8)	5 (10.4)	10 (16.4)	17 (13.5)
Pancreatic	0 (0.0)	1 (2.1)	0 (0.0)	1 (0.8)
Prostate[*]	0 (0.0)	1 (2.1)	7 (11.5)	8 (6.3)
Stomach	1 (5.9)	4 (8.3)	3 (4.9)	8 (6.3)
Colorectal/anal/rectum	0 (0.0)	2 (4.2)	3 (4.9)	5 (4.0)
Other[*]	6 (35.3)	16 (33.3)	16 (26.2)	38 (30.2)
Imaging modalities				
PET	18 (69.2)	57 (87.7)	75 (79.8)	150 (81.1)
SPECT	6 (23.1)	7 (10.8)	17 (18.1)	30 (16.2)
Scintigraphy	6 (23.1)	13 (20.0)	18 (19.1)	37 (20.0)
PET tracer				
FDG	18 (100)	53 (92.9)	63 (84.0)	134 (89.3)
Other	0 (0)	4 (7.0)	19 (25.3)	23 (15.3)
No. of authors				
<4	5 (19.2)	12 (18.5)	10 (10.6)	27 (14.6)
4–7	18 (69.2)	42 (64.6)	67 (71.3)	127 (68.6)
>7	3 (11.5)	11 (16.9)	17 (18.1)	31 (16.8)
Affiliation department of the first author				
Nuclear medicine	8 (30.8)	38 (58.5)	50.0 (53.2)	96.0 (51.9)
Medicine or surgery	(23.1)	10 (15.4)	24.0 (25.5)	40.0 (21.6)
Radiology	7 (26.9)	11 (16.9)	15.0 (16.0)	33.0 (17.8)
Statistics[‡,§]	5 (19.2)	2 (3.1)	2.0 (2.1)	9.0 (4.9)
Others	0 (0)	4 (6.2)	3.0 (3.2)	7.0 (3.8)
Affiliation department of the corresponding author				
Nuclear medicine	8 (30.8)	35 (53.8)	52.0 (55.3)	96.0 (51.9)
Medicine or surgery	5 (19.2)	13 (20.0)	25.0 (26.6)	43.0 (23.2)
Radiology[‡]	8 (30.8)	13 (20.0)	13.0 (13.8)	34.0 (18.4)
Statistics[‡]	4 (15.4)	2 (3.1)	2.0 (2.1)	8.0 (4.3)
Others	0 (0)	2 (3.1)	2.0 (2.1)	4.0 (2.2)
Nuclear medicine specialist				
Yes	21 (80.8)	48 (73.8)	62 (66.0)	131 (70.8)
No	5 (19.2)	17 (26.2)	32 (34.0)	54 (29.2)
Statistician				
Yes[¶]	13 (50.0)	22 (33.8)	25 (26.6)	60 (32.4)
No	13 (50.0)	43 (66.2)	69 (73.4)	125 (67.6)
Discordance between the first and the corresponding author				
Yes[*]	11 (42.3)	38 (58.5)	66 (70.2)	115 (62.2)
No[¶]	15 (57.7)	27 (41.5)	28 (29.8)	70 (37.8)
Funding				
Yes	9 (34.6)	24 (36.9)	37 (39.4)	70 (37.8)
No	17 (65.4)	41 (63.1)	57 (60.6)	115 (62.2)

Data are number of articles, with percentage in parentheses.
[*] Statistically significant increase.
[†] Oncology category only.
[‡] Statistically significant decrease.
[§] Epidemiologists were included among statisticians.
[¶] Includes acknowledgment.

Table 3

Countries of origin of the first authors in the meta-analyses and systematic reviews published between 2005 and 2016 in Nuclear Medicine Journals.

	Period			
Country	2005–2008 (26 articles)	2009–2012 (65 articles)	2013–2016 (94 articles)	Total (185 articles)
China*	2 (7.6)	28 (43.0)	43 (45.7)	73 (39.5)
Netherlands	7 (26.9)	10 (15.3)	8 (8.5)	25 (13.5)
United States	7 (26.9)	6 (9.2)	9 (9.6)	22 (11.9)
Italy	0	10 (15.3)	10 (10.6)	20 (10.8)
Japan	2 (7.6)	1 (1.5)	2 (2.1)	5 (2.7)
Spain	3 (11.5)	2 (3.0)	0	5 (2.7)
United Kingdom	1 (3.8)	1 (1.5)	3 (3.2)	5 (2.7)
Denmark	0	0	3 (3.2)	3 (1.6)
Greece	1 (3.8)	1 (1.5)	1 (1.1)	3 (1.6)
Iran	0	1 (1.5)	2 (2.1)	3 (1.6)
South Korea	0	1 (1.5)	2 (2.1)	3 (1.6)
Switzerland	0	1 (1.5)	2 (2.1)	3 (1.6)
Australia	0	0	2 (2.1)	2 (1.1)
France	0	1 (1.5)	1 (1.1)	2 (1.1)
Canada	1 (3.8)	1 (1.5)	0	2 (1.1)
France	0	1 (1.5)	1 (1.1)	2 (1.1)
Norway	0	0	2 (2.1)	2 (1.1)
Bangladesh	1 (3.8)	0	0	1 (0.5)
Belgium	1 (3.8)	0	0	1 (0.5)
Chile	0	0	1 (1.1)	1 (0.5)
Germany	0	0	1 (1.1)	1 (0.5)
India	0	1 (1.5)	0	1 (0.5)
Singapore	0	0	1 (1.1)	1 (0.5)
Sweden	0	0	1 (1.1)	1 (0.5)
Total	26 (100)	65 (100)	94 (100)	185 (100)

Data are number of articles with percentages in the parentheses.
* Statistically significant increase (includes articles published in Taiwan).

surgery. The proportion of first and corresponding authors affiliated with the department of nuclear medicine and medicine or surgery increased over the review period, but the change was not statistically significant. The proportion of articles that had a nuclear medicine specialist among the authors was more than half (n=131; 70.8%).

On the other hand, the proportion of articles that had a statistician among the authors decreased significantly over the review period ($P=.003$).

The country where most first authors lived was China, followed by the Netherlands, the United States, and others. Furthermore, the proportion of articles coming from China increased significantly over the review period ($P=.003$); the proportion of articles coming to the remaining 23 countries did not significantly change during the review period (Table 3).

3.3. Quality assessment

The agreement between the 2 investigators on the assessment of the AMSTAR score was moderate based on the kappa score ($\kappa=0.670$). The mean ± standard deviation score of the AMSTAR assessment was 5.77 ± 1.75 in 2005 to 2008, 6.71 ± 1.44 in 2009 to 2012, and 7.44 ± 1.61 in 2013 to 2016; and the score was increased significantly at each period ($P<.001$) (Table 4). And we present the list of all 185 articles and the AMSTAR score of each article in Appendix A, http://links.lww.com/MD/C1000.

The mean AMSTAR score of each journal that had more than 4 publications is shown in Figure 2. The highest mean AMSTAR score was 7.6 for the *European Radiology*, followed by 7.5 for the *Journal of Nuclear Medicine*, and 7.4 for the Journal of Nuclear Cardiology; the lowest score was 4.25 for the *American Journal of Roentgenology*. Appendix B, http://links.lww.com/MD/C1000 presents the AMSTAR score of each journal between 2005 and 2016.

Item 4 (Was the status of publication [eg, grey literature] used as an inclusion criterion?) and item 5 (Was a list of studies [included and excluded] provided?) had the lowest score (the percentage of "yes" or "not applicable" answers). Although the scores for item 4 tended to decrease over the review period, the scores for the other 10 items tended to increase gradually (Fig. 3).

Among the 185 articles, 155 (83.8%) did not assign a point on item 4:19 (73.1%] in 2005 to 2008, 53 (81.5%) in 2009 to 2012, and 83 (88.3%) in 2013 to 2016. Exclusion of small-sized studies was most common (61.3%) for this observation, followed by language restriction (56.1%), exclusion of conference papers (29.7%), other reasons (4.5%), and exclusion of low-quality articles (3.2%). Analysis of items 6 and 7 in the AMSTAR showed that various methodologies were used to evaluate the quality of meta-analyses and systematic reviews (Table 4). Articles in which quality assessment was performed increased in over the review period, but the change was not statistically significant. Quality assessment of diagnostic accuracy studies (QUADAS) and QUADAS-2 were the most commonly used methods to perform quality assessment among the 153 articles that were evaluated in the present study: 84 (54.9%) and 32 (20.9%) articles, respectively, used these methodologies for quality evaluation. The ratio of articles using QUADAS-2 significantly increased throughout the review period ($P<.001$).

Table 4

Results of AMSTAR assessment, methods for quality assessment, and statistical analysis of the meta-analyses and systematic reviews published between 2005 and 2016 in Nuclear Medicine Journals.

Characteristic	Period			
	2005–2008 (26 articles)	2009–2012 (65 articles)	2013–2016 (94 articles)	Total (185 articles)
AMSTAR assessment				
Mean score*	5.77 ± 1.75†	6.71 ± 1.44	7.44 ± 1.61	6.95 ± 1.67
Proportion of quality assessment				
Meta-analysis	18/24 (75.0)	53/60 (88.3)	72/80 (90.0)	143/164 (87.1)
Systematic review	2/2 (100.0)	1/5 (20.0)	7/14 (50.0)	10/21 (47.6)
Total	20/26 (76.9)	54/65 (83.0)	79/94 (84.0)	153/185 (82.7)
Methods of quality assessment‡				
QUADAS	8 (40.0)	36 (66.7)	40 (50.6)	84 (54.9)
QUADAS-2*	0 (0.0)	1 (1.9)	31 (39.2)	32 (20.9)
Recommendations of the Cochrane methods working group on systematic review of screening and diagnostic tests	3 (15.0)	8 (14.8)	0 (0.0)	11 (7.2)
OCEBM level of evidence: diagnostic domain	1 (5.0)	3 (5.6)	3 (3.8)	7 (4.6)
Own criteria§	8 (40.0)	6 (11.1)	5 (6.3)	19 (12.4)
Statistical analysis (meta-analysis)¶				
Univariate only	5 (20.8)	6 (10.0)	8 (10.0)	19 (11.6)
Univariate and SROC	15 (62.5)	42 (70.0)	50 (62.5)	107 (65.2)
Subtotal	20 (83.3)	48 (80.0)	58 (72.5)	126 (76.8)
Bivariate only	2 (8.3)	2 (3.3)	2 (2.5)	6 (3.7)
Bivariate and HSROC	2 (8.3)	10 (16.7)	20 (25.0)	32 (19.5)
Subtotal	4 (16.7)	12 (20.0)	22 (27.5)	38 (23.2)
Meta-analysis outcome¶				
DTA (sensitivity, specificity . . .)	23.0 (95.8)	54.0 (90.0)	74.0 (92.5)	151.0 (92.1)
Odds ratio	0.0 (0.0)	0.0 (0.0)	4.0 (5.0)	4.0 (2.4)
Hazard ratio	0.0 (0.0)	2.0 (3.3)	1.0 (1.3)	3.0 (1.8)
Risk ratio	1.0 (4.2)	2.0 (3.3)	0.0 (0.0)	3.0 (1.8)
Prevalence	0.0 (0.0)	1.0 (1.7)	1.0 (1.3)	2.0 (1.2)
Survival outcome	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
Others	0.0 (0.0)	1.0 (1.7)	0.0 (0.0)	1.0 (0.6)
Reason meta-analysis was not performed‖				
Heterogeneity	0 (0.0)	0 (0.0)	2 (14.3)	2 (9.5)
Insufficient studies	0 (0.0)	0 (0.0)	2 (14.3)	2 (9.5)
Study quality	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Not applicable	1 (50.0)	0 (0.0)	6 (42.9)	7 (33.3)
No specified	1 (50.0)	5 (100.0)	4 (28.6)	10 (47.6)

Data are number of articles with percentages in the parentheses.

AMSTAR = a measurement tool to assess systematic reviews, HSROC = hierarchical summary receiver operating characteristic, OCEBM = Oxford Centre for evidence-based medicine level of evidence: diagnostic domain, QUADAS = quality assessment of diagnostic accuracy studies, SROC = summary receiver operating characteristic.

* Statistically significant increase.

† Mean ± standard deviation.

‡ Includes only 153 articles in which quality assessment was performed.

§ Statistically significant decrease.

¶ Includes only 164 articles in which meta-analysis was performed.

‖ Includes only 21 articles in which a systematic review was performed.

The ratio and number of articles using the authors' own criteria significantly decreased throughout the review period.

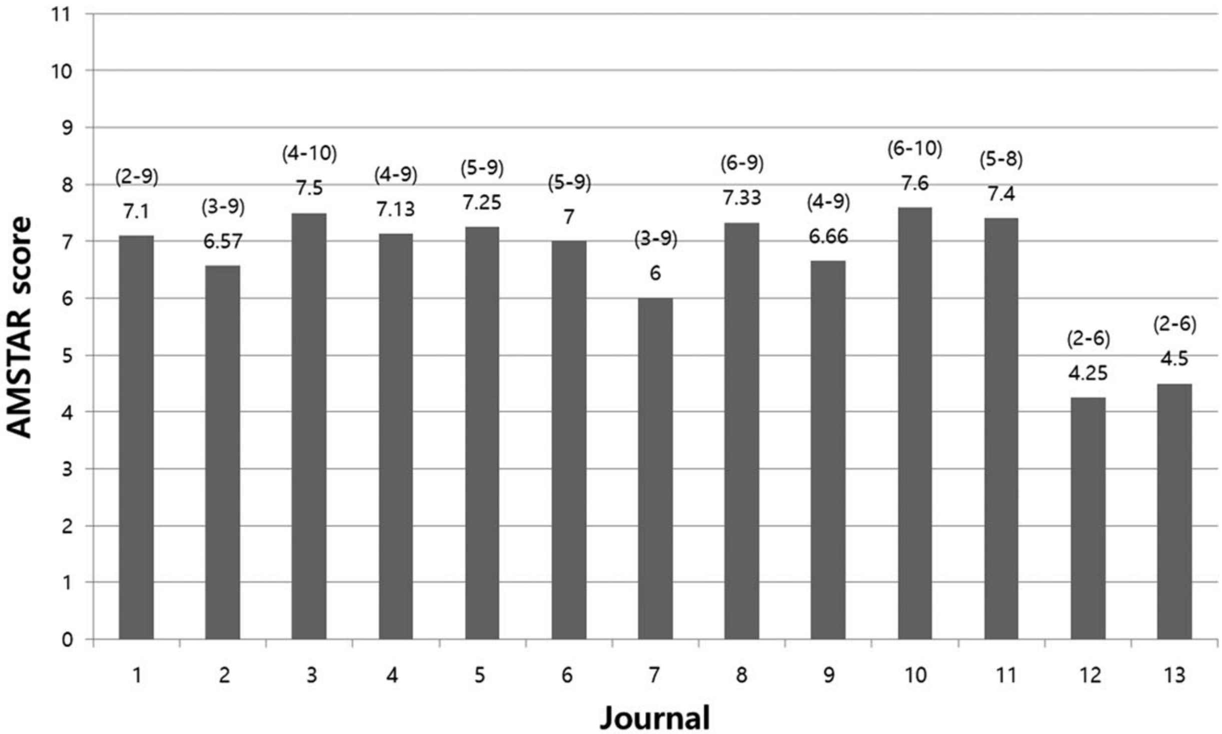
Various statistical methods were used in the meta-analyses to evaluate item 9 of the AMSTAR (were the methods used to combine the findings of studies appropriate?). Among the 164 meta-analyses that were evaluated in the present study, 126 used univariate analysis only or univariate and SROC curve analysis (Table 4). Bivariate analysis, HSROC analysis, or both methods were used in the remaining 38 articles. During each period, the proportion of the articles in which a bivariate model was used increased slightly, but the change was not statistically significant.

The most commonly using meta-analysis outcome was DTA (n = 151; 92.1%); other outcomes that were used were odds ratio, hazard ratio, risk ratio, and prevalence.

There were 21 systematic review articles in which data pooling was not performed in the present study: 11 articles mentioned the reason, but 10 (47.6%) articles did not mention the reason.

4. Discussion

Systematic reviews that include meta-analyses are considered to be the highest level of evidence because of the summarization of the included primary studies; these reviews are increasingly used for not only evidence-based decision making in therapeutic/interventional studies but also generalization of diagnostic accuracy tests in the field of imaging research. As the importance of evidence-based medicine has increased in recent years, the need for systematic reviews in the field of nuclear medicine has also increased. Previous studies had reported that there was a



Note. Data is mean score, with minimum and maximum value in parentheses.
1=Nuclear Medicine Communications, 2=European Journal of Nuclear Medicine and Molecular Imaging, 3= Journal of Nuclear Medicine, 4=European Journal of Radiology, 5=Clinical Nuclear Medicine, 6=Radiology, 7=Acta Radiologica, 8=Academic Radiology, 9=Annals of Nuclear Medicine, 10=European Radiology, 11=Journal of Nuclear Cardiology, 12= American Journal of Roentgenology, 13=Revista Española de Medicina Nuclear e Imagen Molecular

Figure 2. The total AMSTAR score in the included nuclear medicine journals that published more than 4 articles*: the European Radiology* had the highest mean AMSTAR score (7.6), whereas the *American Journal of Roentgenology* had the lowest (4.25). AMSTAR = a measurement tool to assess systematic reviews.

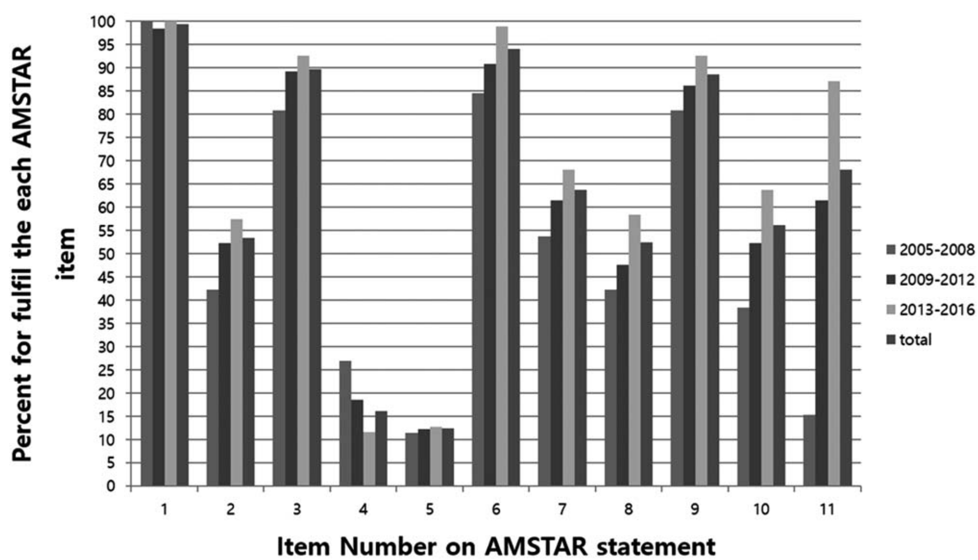


Figure 3. Percentage of fulfillment of each AMSTAR item at each stage: Item 4 (Was the status of publication [ie, grey literature] used as an inclusion criterion?) and item 5 (Was a list of studies [included and excluded] provide?) had the lowest proportion of "yes" or "not applicable" answers. All items, except item 4, tended to improve over time in terms of their percentage of "yes" or "not applicable" answers. AMSTAR = a measurement tool to assess systematic reviews.

considerable growth in the number of published systematic reviews and meta-analyses regarding diagnostic test accuracy,[3] and similar observations have been made in the field of nuclear medicine, especially with regard to PET and PET/CT imaging in oncology.[11,12]

Bibliometric analysis is a method that evaluates published articles in a specific field over a certain period of time; thus, this methodology enables one to understand the trend and characteristics of the published articles in a specific field.[3,5] We performed a bibliometric analysis of systematic reviews and meta-analyses published in general radiology journals and found that a significant number of systematic reviews in the field of nuclear medicine has already been published.[3] However, to the best of our knowledge, there was no bibliometric research focusing on the field of nuclear medicine on the same topic.

The absolute number of systematic reviews and meta-analyses in the field of nuclear medicine has increased in the past 12 years. The reason for the increase in the number of these studies in this field is thought to be the increase in the importance of the evidence-based medicine. Many studies have reported that the importance of evidence-based medicine has increased, and the number of meta-analyses and systematic reviews have also increased not only in the field of clinical medicine[13–15] but also in radiology.[3] We found that there is a similar tendency in the field of nuclear medicine.

Oncology was found to be the most productive field in the present study (68.1%; 126/185) although this observation was not statistically significant. Nuclear medicine procedures can demonstrate various metabolic and biochemical processes; therefore, these procedures play incremental roles in the diagnosis of a number of tumors.[16] The most common imaging modality in this study was PET (n = 150, 81.1%). The introduction of PET-CT scanners that use various tracers has contributed to better detection of malignant tumors and better acquisition of additional information about the biochemical behavior of various tumors.[16] A previous bibliometric study on PET demonstrated that the number of published research articles had increased rapidly.[8] The results reflect the increased clinical application of PET and the increased author's interest in nuclear imaging. The most commonly used tracer in PET imaging in the present study was FDG (n = 134; 89.3%). The use of FDG and other radiotracers increased during review period, but this change was not statistically significant.

Within the oncology topic, the proportion of studies focusing on prostate cancer increased significantly. The role of nuclear medicine methods in the evaluation of prostate cancer has increased due to the introduction of new hybrid devices, such as PET/CT and PET/MR and the development of new PET tracers, such as ^{11}C-choline, ^{18}F-choline, and ^{68}Ga-prostate specific membrane antigen.[6,17,18]

The present study demonstrated that most of the articles had had 4 to 7 authors (68.6%; 127/185). The absolute number and proportions of articles with more than 4 authors increased during the review period; however, the difference was not statistically significant. In addition, we discovered that the discordance between first and corresponding author increased significantly during the review period. These results were similar to the results in previous studies on the number of authors.[3,5,19,20] The sustained increase in the number of authors for scientific article in the field of diagnostic test accuracy may reflect the complexity of the study designs and the need for collaborative efforts in present research.

The proportion of first and corresponding authors with affiliation to nuclear medicine departments was 51.9%; conversely, 48.1% of first and corresponding authors were not affiliated with nuclear medicine departments in the present study. Various researchers in non-nuclear medicine fields have performed systematic reviews and meta-analyses in nuclear medicine-related research. In addition, more than a quarter of articles (29.2%; 54/185) did not include nuclear medicine specialists among the authors. These observations may reflect the growing interest in nuclear medicine and an increased trend of multidisciplinary research.[3,19] The great interest of clinicians in nuclear medicine research may be attributed to the significant impact of nuclear medicine on clinical management that is a result of the advancements in nuclear medicine modalities and various radiopharmaceuticals

The proportion of statisticians as both first and corresponding authors significantly decreased over the review period. In addition, the proportion and absolute number of articles that did not include a statistician among the authors increased although the increase was not statistically significant. These results may reflect the fact that many clinicians have become familiar with systematic reviews and meta-analyses with the trend of recent promotion of evidence-based medicine.

In our study, China was the country with the greatest number of authors who published systematic reviews and meta-analyses in the field of nuclear medicine (39.5%; 73 of 185). China was the only country where the proportion of main authors of the article increased over the review period. Previous studies have also reported that the annual number of meta-analyses that were published from China increased 40- to 200-fold over about a decade.[21,22] Park et al reported a similar trend in publications in general radiology journals during a similar period.[3] Previous researchers gave an explanation of this geometric increase:

(1) meta-analyses are relatively cost-effective and less time-consuming compared with clinical research because they do not involve designing and execution of the actual research projects and
(2) there is academic pressure for professional promotion.[21,22]

Systematic reviews and meta-analyses are excellent forms of studies that provide a great level of evidence provided that strict adherence to appropriate methodology was achieved.[1,2,4] In addition, a significant increase in the number of published systematic reviews does not necessarily guarantee the improvement in the methodological quality.[23] There are many challenges to the validity of systematic reviews and meta-analyses, such as systematic and transparent conduct and reporting, poor methodological quality of the included studies, risk of random errors, unrecognized and unaccounted statistical and clinical heterogeneity, data dredging in non-predefined statistical analyses, and lack of assessment of the overall quality of evidence.[23] Inconsistent or poor methodological quality of systematic reviews and meta-analyses that have been published in several disciplines of the medical field has been noted.[24–29] Thus, several tools have been developed for quality assessment of systematic reviews and meta-analyses, including AMSTAR, meta-analysis of observational studies in epidemiology, and the PRISMA statement, although they have some limitations.[3,23,30] We chose AMSTAR for evaluation of methodological quality in this study. The other 2 tools were for appraising systematic reviews and meta-analyses rather than for quality assessments.[3,30] However, AMSTAR has a flaw in that it depends on the reporting quality rather than on the methodological quality.[31] Although new tools, such as the revised AMSTAR and AMSTAR-2, have been developed to overcome the drawback of the original tool,[32,33] they have not yet been sufficiently validated.[34,35] AMSTAR has been previously validated as an assessment tool and still remains the major means for the assessment of systematic reviews because of its simplicity.[30,34,35] In addition, AMSTAR has high inter-rater reliability.[3,4,30,36] In the present study, the interobserver agreement between the 2 investigators was moderate ($\kappa = 0.670$) using AMSTAR.

In our study, the mean AMSTAR score was significantly increased during each period. These results mean there was modest improvement in the study quality. As researchers become accustomed to the performance of systematic reviews and meta-analyses and the use of guidelines and assessment tools, they may become more organized, concise, and reliable in research design. Several previous studies have shown results similar to ours in the field of diagnostic test accuracy in radiology.[3,30] Scores for all items, except item 4, gradually increased during the review period; especially item 1, 3, 6, 9 had more than 90% compliance in last study period 2013 to 2016. However, item 4 and 5 had very low scores, with less than 20% compliance, similar to previous studies in the field of radiology.[3,30] Searching for grey literature on item 4 can be challenging in systematic reviews on imaging that include both nuclear imaging and radiological imaging, and the lack of literature can cause an important deficit in obtained representative sample for performance of an appropriate study. Our study showed that the most common factors related to this deficit were the exclusion of small-sized studies and language restriction. Inclusion or exclusion of small-sized studies was related to publication bias and a small-study effect that can often distort the results. Several guidelines, including one on the use of AMSTAR, have suggested that the small-study effect should be always assessed to reduce this bias.[37] In addition, it has been recommended that authors should consider the impact of language restriction, especially if the studies on the specific topic are likely to have been performed using various language although the effect of language restriction has been reported to be small. However, authors who want to restrict language should clearly cite the excluded studies.[2,38] Although systematic reviews start with a search of the bibliographic database using many peer-reviewed scientific studies, internet searches have increased as an additional source of grey literature, which produce more comprehensive and applicable outcomes of scientific research but result in controversies regarding scientific reproducibility.[39,40] Our results seem to reveal the trend that the internet is not yet widely used for accessing grey literature in the field of nuclear medicine. Item 5 was related to the provision of a list of included and excluded studies. In the present study, the score for this item was a result of a problem that most researchers did not strictly report the excluded studies. In another words, the existence of items 4 and 5 with low scores meant that there is potential for quality improvement if researchers would follow a guideline.

Assessment of risk of bias is an essential part of systematic reviews and meta-analyses to reduce the probability of systematic errors.[2] There are several methods for quality assessment.[2,3] In our study, the most commonly used quality assessment tool was QUADAS and QUADAS-2. For studies of diagnostic test accuracy, QUADAS-2 was developed from the original form, QUADAS; QUADAS-2 is known as a tool of a choice.[41] The ratio of articles using QUADAS-2 significantly increased, especially in the third period, because QUADAS-2 method was published in 2011. The ratio and numbers of articles using own criteria significantly decreased throughout the review period; this observed tendency was because QUADAS and QUADAS-2 methods have become standard tools for quality assessment. In addition, the proportion of studies without quality appraisal decreased although this change was not statistically significant: 76.9% in 2005 to 2008 and 84.0% in 2013 to 2016.

In the process of data summarization, there are traditional statistical methods, such as SROC curve, univariate analysis, and newly developed bivariate methods that include the HSROC curve method and the bivariate analysis.[9,42] Newly developed methods are recommended for researchers because traditional methods can give a false impression of the accuracy of diagnostic tests due to ignorance of threshold effects; in addition, a correlation between sensitivity and specificity in the univariate model can also overestimate diagnostic accuracy and produce a narrower confidence interval.[2,3,42] In a previous study performed in a radiology field, it was observed that the use of HSROC or bivariate methods had gradually increased, although univariate and SROC analyses were still used in more than half of the meta-analysis articles.[3] The present study also showed that more than 3 quarters of the included studies used traditional statistical methods. Despite the superiority of newly developed methods, SROC model and univariate methods are still the mainstay of statistical methods that are used in the meta-analyses. However, the proportion of articles in which bivariate and HSROC were used, increased over the review period although the increase was not statistically significant.

In this study, the diagnostic test accuracy was most commonly reported outcome in the meta-analyses (92.1%). We used the search filters that focused on diagnostic test accuracy for identifying relevant studies.[9] Therefore, this result was expected. In addition, the present study included meta-analyses that used other outcomes, such as odds ratio, hazard ratio, and prevalence; this is in contrast to what was performed in a previous study using the same filter in the field of general radiology whereby the only outcome for all included studies was diagnostic test accuracy.[3] This result may reflect the fact that nuclear medicine research plays an important role in improving diagnostic test accuracy, and it is also used in clinical practice, including in the performance of therapeutic procedures.

Our study had several limitations. First, we searched only the PubMed database to identify relevant articles. Moreover, only journals within radiology, nuclear medicine, and medical imaging as defined by Clarivate Analytics were evaluated. Non-imaging journals were excluded. Although article screening was performed in a standard way, which was used in previous studies,[3,9] there is potential for a bias. The included journals may have represented only a fraction of the total world literature on nuclear medicine. In addition, the inclusion of only SCI- or SCI-E-indexed journals may have led to the overestimation of the methodological quality. However, the purpose of this study was not to retrieve all nuclear medicine systematic reviews but to explore a large representative sample for a broad audience in the field of nuclear medicine. In addition, by limiting the number of studies that qualified for this study, we think that we obtained credible results by excluding poorly conducted and biased studies. Second, there was a possibility that there was subjectivity in the assessment by using AMSTAR because each question could not be equally assessed in all included studies that had indeterminate and unclear results. Thus, we performed a duplicate quality assessment to minimize disagreements, and the results show that there was a consensus in the assessment. Finally, it can be argued that the overall score from separate AMSTAR items was used in this study. However, previous studies have frequently used the total AMSTAR score to evaluate and compare the methodological quality of systematic reviews due to the ease of using AMSTAR and its simplicity, despite its flaws.[3,4,26,30] In addition, we showed the trend of each item of the AMSTAR in Figure 3. Furthermore, if recently developed tools – such as the revised AMSTAR and AMSTAR 2 – are further validated, future investigation using these tools can enable assessment of the methodological quality of systematic reviews.

5. Conclusions

The present study revealed that there was a modest increment in the number and improvement in the quality of systematic reviews and meta-analyses in the field of nuclear medicine over the 12-year review period based on the bibliometric analysis. In addition, we confirmed that there is a possibility of improvement in quality in specific areas of weakness with respect to study methodologies.

Author contributions

Conceptualization: Jun Ho Kim, Soon Gu Cho.
Data curation: Jung Ui Hong, Jun Ho Kim.
Formal analysis: Jung Ui Hong, Jun Ho Kim.
Investigation: Jung Ui Hong, Jun Ho Kim, Minkyung Lee, Yeo Ju Kim.
Methodology: Jun Ho Kim.
Supervision: In Young Hyun, Soon Gu Cho.
Validation: Kyung Hee Lee, Yeo Ju Kim, Ga Ram Kim.
Writing – original draft: Jung Ui Hong, Jun Ho Kim.
Writing – review and editing: Kyung Hee Lee, Minkyung Lee, Ha Young Lee.
Jun Ho Kim orcid: 0000-0002-7913-7506.

References

[1] Kim KW, Lee J, Choi SH, et al. 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;16:1175–87.
[2] McInnes MDF, Bossuyt PMM. Pitfalls of systematic reviews and meta-analyses in imaging research. Radiology 2015;277:13–21.
[3] Park JY, Lee KH, Ku YJ, et al. Characteristics, trends, and quality of systematic review and meta-analysis in general radiology between 2007 and 2015. Acad Radiol 2017;24:1013–22.
[4] Hasan H, Muhammed T, Yu J, et al. Assessing the methodological quality of systematic reviews in radiation oncology: a systematic review. Cancer Epidemiol 2017;50:141–9.
[5] Hong SJ, Yoon DY, Cho YK, et al. Characteristics and quality of radiologic randomized controlled trials: a bibliometric analysis between 1995 and 2014. Am J Roentgenol 2016;206:917–23.
[6] Delbeke D, Segall GM. Status of and trends in nuclear medicine in the United States. J Nucl Med 2011;52:24S–8S.
[7] Pajares Vinardell M, Freire Macías JM. Twenty-five years of the Spanish Journal of Nuclear Medicine. Bibliometric study. Rev Esp Med Nucl 2007;26:345–53.
[8] Baek S, Yoon DY, Min KJ, et al. Characteristics and trends of research on positron emission tomography: a bibliometric analysis, 2002-2012. Ann Nucl Med 2014;28:455–62.
[9] McGrath TA, McInnes MDF, Korevaar DA, et al. 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.
[10] Moher D, Altman DG, Liberati A, et al. PRISMA statement. Epidemiol 2011;22:128.
[11] Sadeghi R, Zakavi R, Kakhki VR. How to apply the evidence-based medicine concept to nuclear medicine diagnostic studies – a review. Nucl Med Rev 2009;12:59–64.
[12] Treglia G, Sadeghi R. Meta-analyses and systematic reviews on PET and PET/CT in oncology: the state of the art. Clin Transl Imaging 2013;1:73–5.
[13] Hung BT, Long NP, Hung LP, et al. Research trends in evidence-based medicine: a joinpoint regression analysis of more than 50 years of publication data. PloS One 2015;10:e0121054.
[14] Guilera G, Barrios M, Gómez-Benito J. Meta-analysis in psychology: a bibliometric study. Scientometrics 2013;94:943–54.
[15] D'Agostino RBJr. Editorial commentary: meta-analyses are on the rise, but are we learning more? Arthroscopy 2016;32:538–9.
[16] Lind P, Gallowitsch H-J. Nuclear medicine and oncology. Wien Med Wochenschr 2012;162:405–6.
[17] Giovacchini G, Giovannini E, Leoncini R, et al. PET and PET/CT with radiolabeled choline in prostate cancer: a critical reappraisal of 20 years of clinical studies. Eur J Nucl Med Mol Imaging 2017;44:1751–76.
[18] Zuchora Z, Browne M, Colgan N. Role of nuclear medicine methods in prostate cancer management. Phys Med 2016;32:956–7.
[19] Lim KJ, Yoon DY, Yun EJ, et al. Characteristics and trends of radiology research: a survey of original articles published in AJR and radiology between 2001 and 2010. Radiology 2012;264:796–802.
[20] Baek S, Yoon DY, Cho YK, et al. Trend toward an increase in authorship for leading radiology journals. Am J Roentgenol 2015;205:924–8.
[21] Ioannidis JPA, Chang CQ, Lam TK, et al. The geometric increase in meta-analyses from China in the genomic era. PLoS One 2013;8:e65602.
[22] Ye XF, Yu DH, He J. The rise in meta-analyses from China. Epidemiology 2013;24:336–7.
[23] Møller MH, Ioannidis JPA, Darmon M. Are systematic reviews and meta-analyses still useful research? We are not sure. Intensive Care Med 2018;44:518–20.
[24] Delaney A, Bagshaw SM, Ferland A, et al. The quality of reports of critical care meta-analyses in the Cochrane Database of Systematic Reviews: an independent appraisal. Crit Care Med 2007;35:589–94.
[25] Gagnier JJ, Kellam PJ. Reporting and methodological quality of systematic reviews in the orthopaedic literature. J Bone Joint Surg Am 2013;95:e771–7.
[26] Braga LH, Pemberton J, DeMaria J, et al. Methodological concerns and quality appraisal of contemporary systematic reviews and meta-analyses in pediatric urology J Urol 2011;186:266–72.
[27] Elangovan S, Avila-Ortiz G, Johnson GK, et al. Quality assessment of systematic reviews on periodontal regeneration in humans. J Periodontol 2013;84:176–85.
[28] Han JL, Gandhi S, Bockoven CG, et al. The landscape of systematic reviews in urology (1998 to 2015): an assessment of methodological quality. BJU Int 2017;119:638–49.
[29] Samargandi OA, Hasan H, Thoma A. Methodologic quality of systematic reviews published in the plastic and reconstructive surgery literature: a systematic review. Plast Reconstr Surg 2016;137:225e–36e.
[30] Tunis AS, McInnes MDF, Hanna R, et al. 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;269:413–26.
[31] Wegewitz U, Weikert B, Fishta A, et al. Resuming the discussion of AMSTAR: what can (should) be made better? BMC Med Res Methodol 2016;16:111.
[32] Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ 2017;358:J4008.
[33] Kung J, Chiappelli F, Cajulis OO, et al. From systematic reviews to clinical recommendations for evidence-based health care: validation of revised assessment of multiple systematic reviews (R-AMSTAR) for grading of clinical relevance. Open Dent J 2010;4:84–91.
[34] Dosenovic S, Jelicic Kadic A, Vucic K, et al. Comparison of methodological quality rating of systematic reviews on neuropathic pain using AMSTAR and R-AMSTAR. BMC Med Res Methodol 2018;18:37.
[35] Pieper D, Buechter RB, Li L, et al. Systematic review found AMSTAR, but not R(evised)-AMSTAR, to have good measurement properties. J Clin Epidemiol 2015;68:574–83.
[36] Kang D, Wu Y, Hu D, et al. Reliability and external validity of AMSTAR in assessing quality of TCM systematic reviews. Evid Based Complement Alternat Med 2012;2012:732195.
[37] Schwarzer G, Carpenter JR, Rücker G. Small-Study Effects in Meta-Analysis. Meta-Anal with R Cham. 2015;Springer, 107–141.

[38] Bown MJ, Sutton AJ. Quality control in systematic reviews and meta-analyses. Eur J Vasc Endovasc Surg 2010;40:669–77.
[39] Ćurković M, Košec A. Bubble effect: including internet search engines in systematic reviews introduces selection bias and impedes scientific reproducibility. BMC Med Res Methodol 2018;18:130.
[40] Curkovic M. The implications of using internet search engines in structured scientific reviews. Sci Eng Ethics 2019;25:645–6.
[41] Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011;155:529–36.
[42] Lee J, Kim KW, Choi SH, et al. 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;16:1188–96.