

대한한방부인과학회지에 게재된 체계적 문헌고찰의 방법론 및 보고의 질 평가

세명대학교 한의과대학 한방부인과교실
서지혜, 강소현, 허다희, 이동녕, 김형준

ABSTRACT

A Methodological and Reporting Quality Assessment of the Systematic Reviews in the Journal of Korean Obstetrics and Gynecology

Ji-Hye Seo, So-Hyeon Kang, Da-Hee Hur,
Dong-Nyung Lee, Hyeong-Jun Kim
Dept. of Korean Obstetrics and Gynecology,
College of Korean Medicine, Se-Myung University

Objectives: The purpose of this study is to assess the methodological and reporting quality of the systematic reviews in the Journal of Korean obstetrics and gynecology.

Methods: Systematic reviews in the Journal of Korean obstetrics and gynecology were selected from January 1987 to February 2021, by searching Koreanstudies Information Service System (KISS) and the search system of the Society of Korean medicine obstetrics and gynecology. The methodological and reporting quality was assessed using A MeaSurement Tool to Assess systematic Reviews 2 (AMSTAR 2) tool and Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.

Results: Total of 19 systematic reviews were finally included for this study. The levels of methodological quality were low in 63.2% and very low in 36.8%. As a result of evaluating the reporting quality, 22 to 15 (average 18.6) items out of 27 items (81.5 to 55.6%) were reported.

Conclusions: The results of a systematic review are helpful in giving a high level of evidence. In order to improve the quality of systematic reviews published in the Journal of Korean Obstetrics and Gynecology, it is necessary to perform and report according to the the PRISMA checklist and AMSTAR 2.

Key Words: Systematic Reviews, Methodological Quality, Reporting Quality, PRISMA, AMSTAR 2, Korean Obstetrics and Gynecology.

Corresponding author(Hyeong-Jun Kim) : Korean Medical Hospital, Se-Myung University,
Sinwol-dong, Jecheon-si, Chungcheongbuk-do, Korea
Tel : 043-649-1819 E-mail : khjoongy@hanmail.net

I. 서 론

체계적 문헌고찰(systematic review)이란 특정 연구 질문에 대하여 신뢰성 있는 정보를 제공하기 위해 명확하고 체계적인 방법을 적용하여 관련된 일차 연구들의 근거를 종합하고 분석하는 연구 방법이다¹⁾. 근거중심의학(evidence-based medicine)에서 체계적 문헌고찰은 가장 높은 수준의 근거로 분류되며 의학적 치료법의 안전성과 유효성에 관한 근거를 종합하는 가장 유용한 도구가 된다. 따라서 체계적 문헌고찰은 임상 진료 지침 개발에 있어 근거 수준을 결정하는 방법이며, 보건정책에도 영향을 미치는 주요한 연구 방법이다²⁾.

체계적 문헌고찰은 메타 분석과 함께 의료 분야에서 점점 더 중요해지고 있으며, 그 숫자도 점점 늘어나고 있다. 하지만 그 방법론과 보고의 질에 있어서 연구마다 다양할 수 있으며, 제대로 수행되지 않은 체계적 문헌고찰은 그 결과도 편향되거나 잘못된 결론으로 이어질 수 있다³⁾. 따라서 체계적 문헌고찰의 방법론과 보고의 질을 개선하는 데 도움을 주기 위해 다양한 방법론 및 보고의 지침이 개발되었다⁴⁻⁷⁾. 체계적 문헌고찰 보고의 질을 평가하기 위해 개발된 도구 중 첫 번째로 점검표와 흐름도를 제시한 Quality of Reporting of Met-analyses statement(QUOROM)가 있다⁴⁾. 이후 2009년에 QUOROM은 세부항목에 보고의 투명성에 관한 항목이 통합되어 Preferred Reporting Items for Systematic reviews and Meta-analyses statement(PRISMA)로 발표되었으며, 이후 프랑스, 독일, 일본, 한국 등 각 나라의 언어로 번역본도 제공되고 있다⁵⁾. 보고의 질 평가 도구인

PRISMA 점검표는 저자들이 연구의 질을 개선하도록 도우며 심사자나 편집자들의 평가를 돕는다⁵⁾. 체계적 문헌고찰의 방법론적 질을 평가하는 도구 중 신뢰도와 구성 타당도가 검증된 도구로는 Assessment of Multiple Systematic Reviews(AMSTAR)가 있다⁶⁾. AMSTAR는 비무작위 연구에 대한 비뚤림 위험(risk of bias)의 평가를 포함하지 않았다는 한계점이 있어, 이를 보완하고 더 각 문항에 대하여 간단한 응답 범주로 수정하여 AMSTAR 2로 개정되었다⁷⁾. AMSTAR 2로 검토한 평가 결과는 전반적인 문헌의 방법론적 질을 매우 낮음(Critically low), 낮음(Low), 중등도(Moderate), 높음(High)으로 분류하게 된다⁷⁾.

국내 의약학 분야에서는 체계적 문헌고찰 연구가 1990년대부터 초반부터 시행되었으며, 이를 대상으로 논문의 질을 평가하는 연구가 이루어졌다⁸⁾. 국내 한의학 분야에서 체계적 문헌고찰을 대상으로 시행한 질 평가 연구로는 2015년도에 시행된 연구와 대한한방소아과학회지의 연구⁹⁾, 대한예방한의학회지의 연구¹⁰⁾ 등이 있었으나 다른 한의학 관련 학회지의 연구는 보고된 바가 없었다. 체계적 문헌고찰의 결과는 임상가의 의사결정을 도와주는 근거를 제공해주지만, 그 근거의 가치와 명확도는 방법론과 보고의 질에 따라 달라질 수 있다¹¹⁾. 이에 본 연구는 대한한방부인과학회지에 게재된 체계적 문헌고찰을 대상으로 AMSTAR 2와 PRISMA 도구를 사용하여 방법론과 보고의 질을 평가하고자 하였다. 이를 통해 향후 체계적 문헌고찰 시행 시 개선할 점을 살펴보고 추후 높은 수준의 근거생성에 도움이 되고자 한다.

II. 방 법

1. 대상 문헌 검색 및 선정

대한한방부인과학회지에 게재된 대상 문헌을 검색하기 위하여 대한한방부인과학회 홈페이지의 학회지 논문 검색 시스템과 한국학술정보(Koreanstudies Information Service System, KISS) 데이터베이스에서 검색을 시행하였다. 창간호가 발간된 시기인 1987년부터 2021년 2월까지 발간된 문헌을 대상으로 하였으며, 검색은 2021년 3월 15일에 시행하였다. 검색어는 “체계적 문헌고찰”, “문헌고찰”, “메타분석”으로 제목에서 검색을 시행하였다. 한국학술정보에서는 추가적으로 간행물명에 “대한한방부인과학회지”를 추가하여 검색하였다. 대한한방부인과학회의 검색 시스템에서는 포괄적인 검색을 위해 “고찰”을 검색어로 추가하였다.

검색 결과 중복된 문헌을 제거 후에 제목과 초록 검토를 통해 1차 선정을 하였으며, 1차 선정된 문헌의 원문을 검토한 후에 최종적으로 체계적 문헌고찰을 수행한 문헌을 선정하였다. 완전한 논문 형태가 아닌 초록이나 학회발표, 연구동향 또는 현황 분석 연구, 서술적 문헌고찰 연구들은 배제하였다. 문헌 선정 과정은 독립적으로 2명의 저자가 시행하였으며, 결과의 불일치가 있는 경우 서로 논의를 통해 해결하였다. 문헌 선정 과정은 선정 흐름도에 정리하였다.

2. 평가 도구

1) AMSTAR 2

최종 선정된 체계적 문헌고찰의 방법론적 질 평가 도구로 AMSTAR 2를 사

용하였다. AMSTAR 2는 신뢰도와 구성 타당도가 검증된 도구로 16개 항목으로 구성되어있다. 각 항목에 대하여 잘 수행되었을 경우에는 ‘예(Y)’, 잘 수행되지 않았거나 정보가 없는 경우에는 ‘아니오(N)’, 문항에서 제시한 최소한의 수행 기준을 만족한 경우에는 ‘부분적 예(PY)’로 평가하였다.

2) PRISMA

보고의 질 평가는 PRISMA 점검표 한국어판(Korean version of PRISMA checklist)을 사용하여 평가하였다. 각 문헌의 제목, 초록, 서론, 방법, 결과, 고찰, 연구비 출처 섹션에 대하여 총 27개 항목을 참고하여 해당 항목에 보고가 되어 있는지를 평가하였다. 보고가 된 경우에는 ‘예(Y)’, 보고가 되지 않은 경우에는 ‘아니오(N)’, 해당 항목을 적용할 수 없거나 해당되지 않은 경우에는 ‘해당 없음(N/A)’으로 평가하였다.

각각의 평가 과정은 독립적으로 2명의 저자가 시행하였으며, 평가 결과의 불일치가 있는 경우 서로 논의를 통해 해결하였다.

3. 분석 방법

개별 문헌에 대한 AMSTAR 2, PRISMA 도구를 활용한 평가 결과는 실수와 백분율로 나타내었다. AMSTAR 2의 16개 항목, PRISMA의 27개 항목에 대하여 각각 ‘예(Y)’와 ‘부분적 예(PY)’로 평가된 문헌의 백분율을 구하여 보고율 또는 수행률을 나타내었다.

AMSTAR 2로 평가한 결과는 중요한 항목(critical domains)의 누락을 조사하여 각 문헌의 방법론적 질을 매우 낮음(Critically low), 낮음(Low), 중등도

(Moderate), 높음(High)으로 평가하였다. AMSTAR 2의 16개 항목 중 중요 항목(critical domains)은 표에 제시하였다(Table 1). 비중요 항목(non-critical domains)의 누락된 항목이 한 개거나 없는 경우에는 높음,

비중요 항목이 두 개 이상 누락된 경우에는 중등도, 중요 항목이 한 개 누락된 경우에는 낮음, 중요 항목이 두 개 이상 누락된 경우에는 매우 낮음으로 평가하였다.

Table 1. The Critical Domains of AMSTAR 2

No.	Critical domains
Item 2	Protocol registered before commencement of the review
Item 4	Adequacy of the literature search
Item 7	Justification for excluding individual studies
Item 9	Risk of bias from individual studies being included in the review
Item 11	Appropriateness of meta-analytical methods
Item 13	Consideration of risk of bias when interpreting the results of the review
Item 15	Assessment of presence and likely impact of publication bias

Ⅲ. 결 과

1. 선정 결과 및 일반적 특성

검색 결과 총 317편의 논문이 검색되었으며 중복 논문 87편을 제외하고 230편의 논문에서 제목과 초록을 통하여 190편의 논문을 제외하였다. 1차 제외 후 남은 40편의 논문 전문을 확인하여 체계적 문헌고찰이 아닌 논문 21편을 제외하여 19편¹²⁻³⁰⁾을 최종적으로 분석하였다(Fig. 1). 대한한방부인과학회지에 게재

된 체계적 문헌고찰 연구는 2014년부터 발행되었고, 2014년도에 1편, 2016년도에 1편, 2017년도에 6편, 2018년도에 3편, 2019년도에 4편, 2020년도에 4편이 발행되었다. 포함된 논문의 대상 질환은 월경이나 자궁과 관련된 질환 등 부인과 질환이었으며, 중재로는 침 또는 전침이 6편, 한약이 7편이었으며 논문별 대상 질환과 중재, 포함된 무작위 대조군 연구의 수에 대해서는 표로 정리하였다(Table 2).

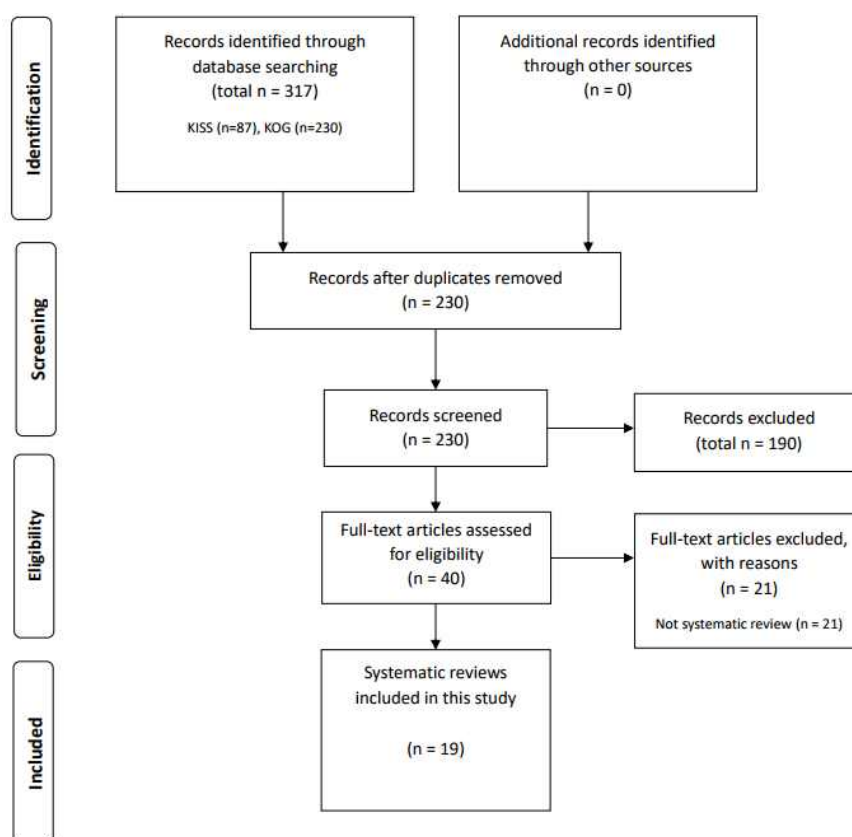


Fig. 1. Flow chart of the study selection.

Table 2. Characteristics of the Included Studies

First author (year)	Patient	Intervention	Number of included RCTs
Kim SH (2014) ¹²⁾	Polycystic Ovary Syndromes	Acupuncture	3
Baek SE (2016) ¹³⁾	Atrophic Vaginitis	Fumigation Therapy	5
Bae JE (2017) ¹⁴⁾	Polycystic Ovary Syndrome Subfertility Patient	Integrated Medicine	8
Jeong SY (2017) ¹⁵⁾	Mastitis	Acupuncture and Moxibustion	6
Kim YJ (2017) ¹⁶⁾	Postpartum Hypogalactia	Acupuncture	5
Lee HI (2017) ¹⁷⁾	Primary Dysmenorrhea	Moxibustion	10
Park NC (2017) ¹⁸⁾	Primary Dysmenorrhea	Electroacupuncture	5
Song MH (2017) ¹⁹⁾	Climacteric Hot Flush	Acupoint Catgut Embedding	1
Park EJ (2018) ²⁰⁾	Dysfunctional Uterine Bleeding	Traditional Herbal Medicine	16
Park KD (2018) ²¹⁾	Mycotic Vaginitis	Fumigation Therapy	5
Yoon JG (2018) ²²⁾	Polycystic Ovarian Syndrome	Electroacupuncture	5
Bae JE (2019) ²³⁾	Endometriosis	Acupoint Sticking Therapy	2
Jung JW (2019) ²⁴⁾	Adenomyosis	Traditional Herbal Medicine	12
Nam EY (2019) ²⁵⁾	Atrophic Vaginitis	Traditional Korean Medicine	18
Park EJ (2019) ²⁶⁾	Atrophic Vaginitis	Traditional Herbal Medicine	26
Ha SJ (2020) ²⁷⁾	Overactive Bladder	Electroacupuncture	5
Kim SG (2020) ²⁸⁾	Premenstrual Syndrome	Traditional Herbal Medicine	8
Lee SJ (2020) ²⁹⁾	Cervical Intraepithelial Neoplasia	Traditional Herbal Medicine	10
Park HR (2020) ³⁰⁾	Stress Urinary Incontinence	Moxibustion	11

2. 방법론적 질 평가 결과

1) 선정된 논문의 평가 결과

선정된 19편의 논문을 방법론적 질 평가를 위해 AMSTAR 2를 적용하여 평가하였다. 평가 결과 '높음'과 '중등도'로 분류된 논문은 없었으며, '낮음'으로 분류된 논문이 12편(63.2%), '매우 낮음'으로 분류된 논문은 7편(36.8%)이었다.

2014년도와 2016년도에 발행된 각각의 1편의 논문은 모두 '낮음'으로 평가되었다. 2017년도에 발행된 논문은 3편(50%)이 '낮음', 3편(50%)이 '매우 낮음'으로, 2018년도에 발행된 논문은 2편(66.7%)이 '낮음', 1편(33.3%)이 '매우 낮음'으로, 2019년도에 발행된 논문은 3편(75%)이 '낮음', 1편(25%)이 '매우 낮음'으로, 2020년도에 발행된 논문은 2편(50%)이 '낮음', 2편(50%)이 '매우 낮음'으로 평가되었다. AMSTAR 2의 항목별로 개별논문을 평가한 상세 결과는 표로 정리하였다(Table 3).

2) AMSTAR 2의 문항별 평가 결과

AMSTAR 2의 항목 중 핵심질문(Patient, Intervention, Comparison, Outcome, PICO)의 구성 요소(Item 1)와 선택한 연구의 연구 디자인 종류(Item 3)에 대해서는 포함된 모든 논문에서 보고되었다. 포괄적 검색전략(Item 4)과 선택한 연구의 상세한 기술(Item 8)과 관련한 문항에서는 최소한 2개의 데이터베이스를 검색하였고 키워드 또는 검색전략을 기술하였으며, 선택한 연구에 대한 정보를 표로 제시하였기에 모두 '부분적 예'로 평가하였다. 또한, 모든 논문에서 포함된 연구의 편향 위험 평가는 Cochrane의 비플림 위험 평가 도구(risk-of-bias tool, RoB)로 적절한 도구(Item 9)를 사용하여 수행되었다. 비플림 위험도가 결과에 미칠

수 있는 영향(Item 13)에 대한 논의는 1편을 제외한 논문에서 모두 기술하였다. 1편을 제외한 논문에서 모두 체계적 문헌고찰 방법을 명시하였으나 사전에 확립되었거나 프로토콜을 등록했다고 명시적으로 기술(Item 2)한 논문은 없어 '부분적 예'로 평가하였다.

문헌의 선택과정과 데이터 추출 과정에서 중복으로 수행했는지에 대한 문항(Item 5, 6)에서는 각각 12편(63.2%)과 10편(52.6%)에서 명확한 기술이 없었으며 나머지 논문에서는 2명의 저자가 각각 중복하여 수행했음을 기술하였다. 배제된 문헌 목록 제공에 관한 문항(Item 7)에서는 모든 문헌에서 배제한 사유와 문헌의 수는 기술하였으나 목록을 제공한 논문은 없어 모두 '아니오'로 평가하였다. 포함된 문헌의 연구비 출처에 대한 문항(Item 10)에서도 포함된 문헌의 연구비 출처에 관한 정보를 제시한 논문이 없었다. 연구 결과 이질성에 대한 논의 문항(Item 14)에서는 4편(21.1%)에서 이질성의 원인이나 이질성이 연구 결과에 미치는 영향을 기술하지 않았다.

19편 중 메타분석을 시행한 논문은 6편으로 메타분석이 시행되었을 경우 평가하는 문항들(Item 11, 12, 15)에 대해서 평가하였으며, 나머지 13편은 'No meta-analysis conducted(No MA)'로 평가하였다. 메타분석을 시행한 6편 논문 모두 메타분석시 비플림 위험의 잠재적 영향(Item 12)과 출판 비플림(Item 15)에 대한 평가를 시행하지 않았으며, 모두 적절한 메타분석 방법을 사용하였으나 2편(10.5%)에서는 이질성이 관찰되었으나 이질성의 원인을 조사하지 않았다(Item 11).

연구 자금을 포함하여 이해상충의 잠재적인 원인에 관한 문항(Item 16)에서 6편은 연구자금에 관한 기술이 있었으나, 19편 모두 이해상충에 관한 기술이

없어 모두 '아니오'로 평가하였다.

AMSTAR 2의 각 문항별 평가 결과는 그래프로 나타내었다(Fig. 2).

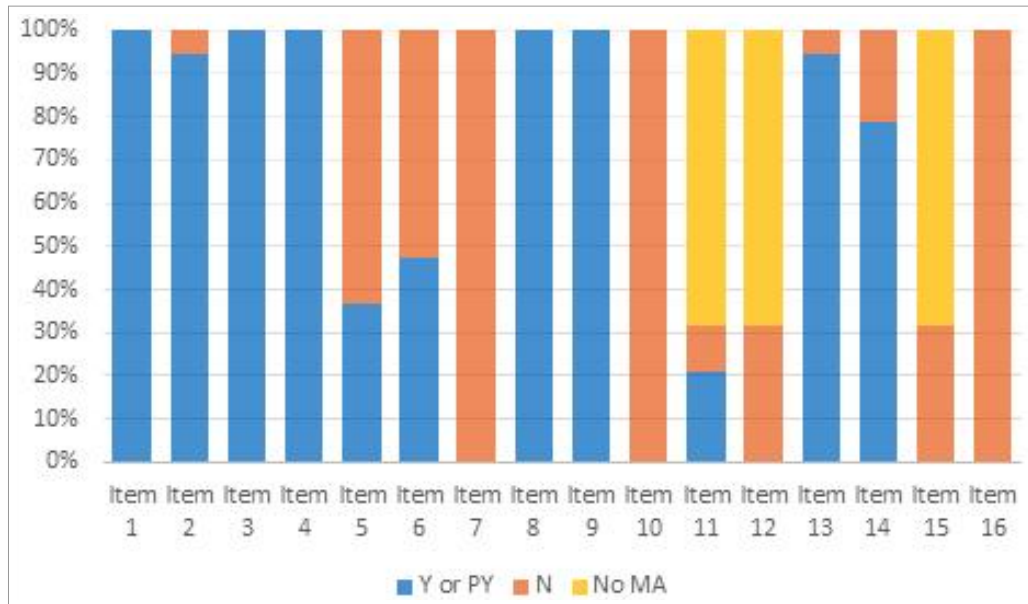


Fig. 2. Evaluation result for each item of AMSTAR 2.

Table 3. Methodological Quality Assessment of Included Studies with AMSTAR 2

First author (year)	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10	Item 11	Item 12	Item 13	Item 14	Item 15	Item 16	Overall quality
Kim SH ¹² (2014)	Y*	PY*	Y	PY	N	Y	N	PY	Y	N*	No MA [§]	No MA	Y	Y	No MA	N	Low
Baek SE (2016) ¹³	Y	PY	Y	PY	N	Y	N	PY	Y	N	No MA	No MA	Y	Y	No MA	N	Low
Bae JE (2017) ¹⁴	Y	PY	Y	PY	Y	Y	N	PY	Y	N	Y	N	Y	Y	N	N	Critically Low
Jeong SY (2017) ¹⁵	Y	PY	Y	PY	N	Y	N	PY	Y	N	N	N	Y	N	N	N	Critically Low
Kim YJ (2017) ¹⁶	Y	PY	Y	PY	N	N	N	PY	Y	N	No MA	No MA	Y	Y	No MA	N	Low
Lee HI (2017) ¹⁷	Y	PY	Y	PY	N	Y	N	PY	Y	N	N	N	Y	N	N	N	Critically Low
Park NC (2017) ¹⁸	Y	PY	Y	PY	N	N	N	PY	Y	N	No MA	No MA	Y	N	No MA	N	Low
Song MH (2017) ¹⁹	Y	PY	Y	PY	Y	N	N	PY	Y	N	No MA	No MA	Y	Y	No MA	N	Low
Park EJ (2018) ²⁰	Y	PY	Y	PY	Y	N	N	PY	Y	N	No MA	No MA	Y	Y	No MA	N	Low
Park KD (2018) ²¹	Y	PY	Y	PY	Y	Y	N	PY	Y	N	No MA	No MA	Y	Y	No MA	N	Critically Low
Yoon JG (2018) ²²	Y	PY	Y	PY	N	N	N	PY	Y	N	No MA	No MA	Y	N	No MA	N	Low
Bae JE (2019) ²³	Y	PY	Y	PY	Y	Y	N	PY	Y	N	No MA	No MA	Y	Y	No MA	N	Low
Jung JW (2019) ²⁴	Y	PY	Y	PY	N	N	N	PY	Y	N	No MA	No MA	N	Y	No MA	N	Critically Low
Nam EY (2019) ²⁵	Y	N	Y	PY	N	N	N	PY	Y	N	No MA	No MA	Y	Y	No MA	N	Low
Park EJ (2019) ²⁶	Y	PY	Y	PY	Y	N	N	PY	Y	N	No MA	No MA	Y	Y	No MA	N	Low
Ha SJ (2020) ²⁷	Y	PY	Y	PY	N	N	N	PY	Y	N	Y	N	Y	Y	N	N	Critically Low
Kim SG (2020)	Y	PY	Y	PY	N	N	N	PY	Y	N	No MA	No MA	Y	Y	No MA	N	Low
Lee SJ (2020) ²⁸	Y	PY	Y	PY	N	Y	N	PY	Y	N	Y	N	Y	Y	N	N	Critically Low
Park HR (2020) ³⁰	Y	PY	Y	PY	Y	Y	N	PY	Y	N	No MA	No MA	Y	Y	No MA	N	Low

*Y : yes, †PY : partial yes, #N : no, §No MA : no meta-analysis conducted

3. 보고의 질 평가 결과

1) 선정된 논문의 평가 결과

보고의 질 평가는 PRISMA 점검표로 평가하였다. PRISMA 점검표의 경우 보고의 질을 분류하는 명확한 기준이 PRISMA statement에 제시되지 않았기에 평가 결과를 ‘예(Y)’, ‘아니오(N)’, ‘해당 없음(N/A)’의 숫자로 나타내었다. 총 27개 문항 중 22개 문항을 보고한 논문 3편, 20개 문항을 보고한 논문 1편, 19개 문항을 보고한 논문 5편, 18개 문항을 보고한 논문 8편, 16개 문항을 보고한 논문 1편, 15개 문항을 보고한 논문 1편이었다. PRISMA 점검표의 항목별로 개별논문을 평가한 상세 결과는 표로 정리하였다(Table 4).

2) PRISMA의 항목별 평가 결과

PRISMA 점검표의 항목별 평가 결과는 그래프로 제시하였다(Fig. 3). 구조화된 요약(2번), 연구의 필요성(3번), 선정 기준(6번), 문헌선택(9번), 자료수집 과정(10번), 자료추출 항목(11번), 연구선택(17번), 연구특성(18번), 개별연구의 비뚤림 위험 평가방법(19번), 개별연구의 결과(20번), 근거요약(24번), 제한점(25번), 결론(26번)에 대한 13개의 항목은 모든 논문에서 보고하였다.

제목(1번)은 1편의 논문이 ‘체계적 문헌고찰’이 아닌 ‘고찰’로만 명시하였으며, 나머지 논문에서는 모두 제목에 ‘체계적

문헌고찰’을 명시하였다. 연구목적(4번) 항목에서 2편(10.5%)의 논문이 PICO를 명시하지 않았다. 정보원(7번) 항목에서는 8편(42.1%)이 최종 검색일을 명시하지 않았으며, 검색(8번) 항목에서 7편(37.8%)이 검색전략을 제시하지 않았다. 개별연구의 비뚤림 위험 평가 방법(12번)에 있어서 1편(5.3%)이 결과는 제시하였으나 방법에 평가방법을 기술하지 않았으며, 요약 측정치(13번)와 자료합성 방법(14번)에 대해서는 각각 6편(31.6%)과 12편(63.2%)이 명확하게 기술하지 않았다. 연구비 출처(27번)에 대해서는 13편(68.3%)이 기술하지 않았다.

결과합성(21번)에 대한 항목은 메타분석에 대한 항목으로 13편(68.3%)은 메타분석을 시행하지 않아 ‘해당 없음(N/A)’으로 평가하였으며, 메타분석을 시행한 6편의 논문 모두 메타분석 후 신뢰구간과 일관성 측정치를 제시하였다.

프로토콜과 등록(5번)은 모든 논문에서 보고하지 않았다. 전체적인 비뚤림 위험(15번), 추가분석(16번)에 대한 2개의 항목도 모든 논문에서 보고되지 않았으며, 이 항목과 관련된 결과는 보고될 수 없으므로 전체적 비뚤림 위험 결과(22번), 추가분석 결과(23번)에 대해서는 ‘해당 없음(N/A)’으로 평가하였다.

Table 4. Reporting Quality Assessment of Included Studies with PRISMA Checklist

No.	Topic	Kim SH (2014)	Baek SE (2016)	Bae JE (2017)	Jeong SY (2017)	Kim YJ (2017)	Lee HI (2017)	Park NC (2017)	Song MH (2017)	Park EJ (2018)	Park KD (2018)	Yoon JG (2018)	Bae JE (2019)	Jung JW (2019)	Nam EY (2019)	Park EJ (2019)	Ha SJ (2020)	Kim SG (2020)	Lee SJ (2020)	Park HR (2020)	
1	Title	Y*	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N*	Y	Y	
2	Structured summary	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
3	Rationale	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
4	Objectives	N	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
5	Protocol and registration	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	
6	Eligibility criteria	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
7	Information sources	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	N	N	N	N	N	
8	Search	N	Y	Y	N	Y	Y	N	Y	Y	Y	N	Y	Y	N	Y	N	N	Y	Y	
9	Study selection	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
10	Data collection process	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
11	Data items	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
12	Risk of bias in individual studies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	
13	Summary measures	Y	Y	Y	Y	N	Y	Y	N	Y	Y	Y	N	Y	N	Y	Y	N	Y	N	
14	Synthesis of results	Y	N	Y	Y	N	Y	N	N	N	N	N	N	N	N	N	Y	N	Y	N	
15	Risk of bias across studies	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	
16	Additional analyses	Y	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	
17	Study selection	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
18	Study characteristics	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
19	Risk of bias within studies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
20	Results of individual studies	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
21	Synthesis of results	N/A*	N/A	Y	Y	N/A	Y	N/A	N/A	N/A	Y	N/A	N/A	N/A	N/A	N/A	Y	N/A	Y	N/A	
22	Risk of bias across studies	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
23	Additional analysis	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	
24	Summary of evidence	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
25	Limitations	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
26	Conclusions	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	
27	Funding	N	N	Y	N	N	Y	N	N	Y	Y	N	N	N	N	Y	N	N	N	Y	
	Total	Y=19 N=5 N/A=3	Y=18 N=6 N/A=3	Y=22 N=3 N/A=2	Y=19 N=6 N/A=2	Y=18 N=6 N/A=3	Y=22 N=3 N/A=2	Y=18 N=6 N/A=3	Y=18 N=6 N/A=3	Y=18 N=6 N/A=3	Y=19 N=5 N/A=3	Y=22 N=3 N/A=2	Y=18 N=6 N/A=3	Y=18 N=6 N/A=3	Y=18 N=6 N/A=3	Y=14 N=10 N/A=3	Y=19 N=5 N/A=3	Y=19 N=6 N/A=2	Y=15 N=9 N/A=3	Y=20 N=5 N/A=2	Y=18 N=6 N/A=3

*Y : yes, *N : no, *N/A: not applicable

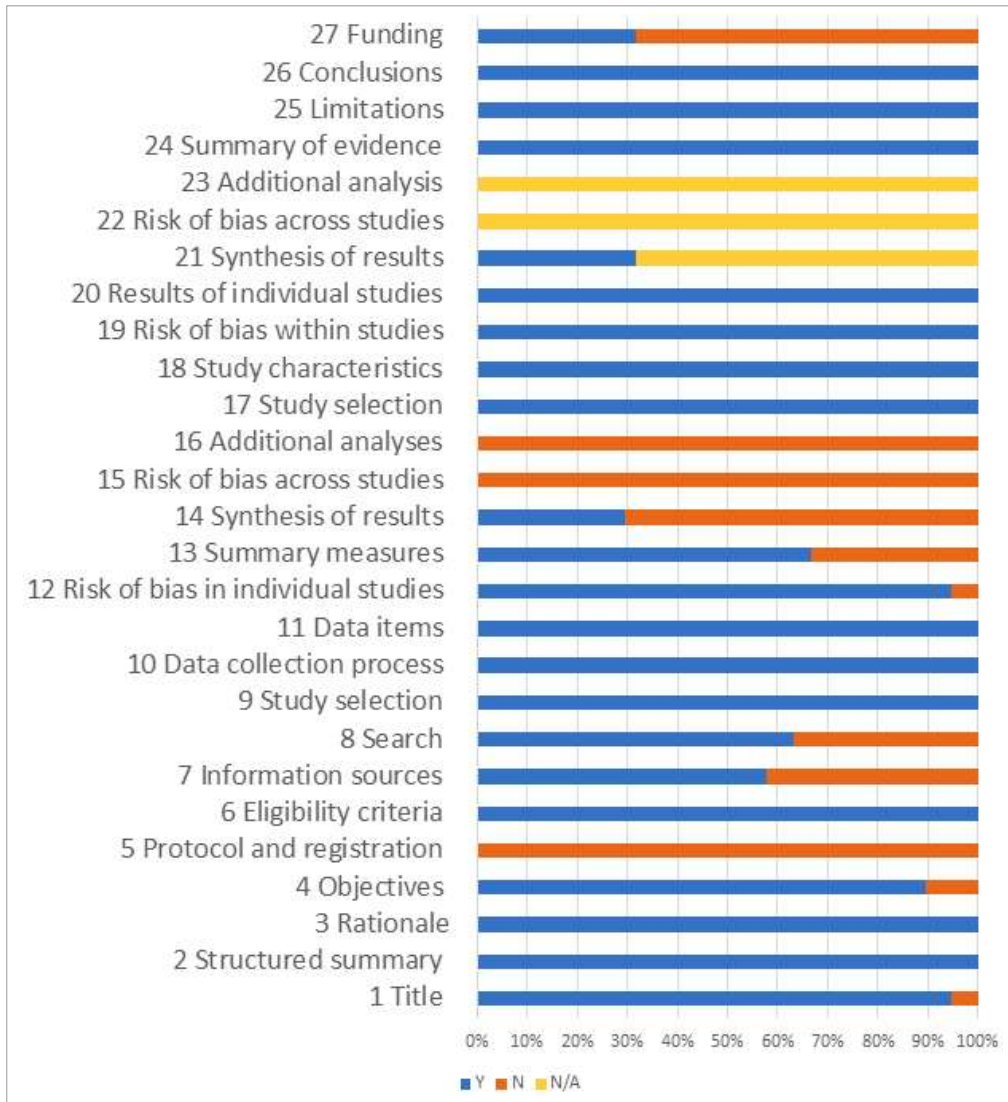


Fig. 3. Evaluation result for each item of PRISMA checklist.

IV. 고찰

본 연구는 대한한방부인과학회지에 게재된 체계적 문헌고찰 논문들을 대상으로 보고 및 방법론적인 질 평가를 수행하였다. 최종 선정된 19편의 체계적 문헌고찰은 2014년부터 발표되었으며 2017년부터는 매해 3~5편씩 꾸준히 발표되었다. 선정된 체계적 문헌고찰의 방법론적 질 수준은 낮음(63.2%) 또는 매우 낮음(36.8%)이었으며, 보고 질 평가 결과

PRISMA 점검표의 27개의 문항 중 22~15 (81.5~55.6%)개를 보고하였으며, 13편 (68.4%)이 19~18(70.4~66.7%)개의 문항을 보고하였으며 평균 18.6개의 문항을 보고하였다.

체계적 문헌고찰은 특정한 주제와 연관된 연구들의 광범위한 검색, 체계적인 선택과 분석, 비뚤림 위험도 평가 등을 시행하며 체계적인 방법론에 따라 객관적이고 분명한 연구방법과 목적을 가지는 연구방법으로 근거중심의학에서 가장

높은 근거 수준으로 분류된다³¹⁾. 임상 의들에게 많은 연구 정보들을 요약하고 종합 평가한 결과를 제공해준다는 이점이 있으나, 같은 주제로 시행한 체계적 문헌고찰과 메타분석 중 서로 다른 결론을 내리고 있는 연구들이 존재하며 체계적 문헌고찰에도 편향의 위험이 있다는 한계점이 있다^{32,33)}. 또한, 체계적 문헌고찰의 결과는 임상가의 의사결정뿐만 아니라 보건의료 정책에에도 영향을 미칠 수 있기 때문에 질 평가가 필요하며, 이를 위한 다양한 도구들이 개발되었다⁴⁻⁷⁾. 본 연구에서는 보고의 질 평가도구로 널리 쓰이고 있는 PRISMA 점검표와 신뢰도와 구성 타당도가 검증된 AMSTAR 2를 활용하여 체계적 문헌고찰의 방법론 및 보고의 질을 평가하였다.

체계적 문헌고찰의 질 평가가 필요함에 따라 국내외에서 선행된 연구가 여럿 시행되었다^{9,10,34-6)}. 2017년도에 시행된 중국과 미국의 체계적 문헌고찰의 방법론 및 보고의 질 평가 연구³⁴⁾에서 AMSTAR로 평가한 결과 18%가 낮음, 60%가 중등도, 22%가 높음으로 평가되었으며, PRISMA로 평가한 결과 중국의 문헌은 평균 21.2개, 미국의 문헌은 평균 20.6개를 보고하였다. 또한, 국내 대한한방소아과학회지의 4건의 문헌을 AMSTAR 2로 평가한 결과 3건은 중등도, 1건은 매우 낮음으로 평가되었으며, PRISMA 점검표를 적용한 결과 평균 19.3건을 보고하였다⁹⁾. 해외 선행연구³⁴⁾와 비교하여 볼 때, 본 연구에 포함된 체계적 문헌고찰은 방법론 및 보고의 질이 상대적으로 낮았다. 하지만, 선행연구에서 포함된 문헌 중 약 6-10%만이 한의학적 중재에 대한 연구로 AMSTAR 2를 사용하여 평

가한 침 또는 한약과 관련된 선행연구에서 보고된 질은 본 연구의 결과와 유사하게 매우 낮았다^{35,36)}. 또한, 국내 대한한방소아과학회지의 연구⁹⁾에 포함된 4건의 체계적 문헌고찰은 AMSTAR 2의 중요한 항목인 7번과 15번을 누락 하였으므로 기준에 따라 4건 모두 매우 낮음으로 평가되어야 하며, 그 결과 본 연구의 질 평가 결과와 유사하다고 볼 수 있다. 따라서, 향후 대한한방부인과과학회지의 체계적 문헌고찰 뿐만 아니라, 국내외 한의학 관련 체계적 문헌고찰은 방법론 및 보고의 질 개선이 필요할 것으로 보인다.

방법론적인 측면에서 AMSTAR 2의 7개의 중요한 항목은 체계적 문헌고찰의 결과에 영향을 미칠 수 있으며 보고의 질 또한 체계적 문헌고찰의 타당도와 신뢰도를 평가함에 있어 중요한 요소다⁷⁾. 본 연구 결과에서 대한한방부인과과학회지에 게재된 체계적 문헌고찰의 방법론 및 보고의 질 측면에서 공통으로 누락된 항목을 살펴봄으로써 향후 게재되는 체계적 문헌고찰의 개선방안을 모색해볼 수 있다. AMSTAR 2의 중요한 항목 중 공통으로 누락된 항목은 배제된 문헌 목록 제공에 관한 항목(Item 7)으로 배제된 이유와 함께 배제된 연구 목록을 제공해야 하나 모든 연구에서 배제된 기준이나 이유에 대하여 기술하였지만 목록을 제시한 연구는 없었다. 배제된 연구들은 체계적 문헌고찰의 결과에 영향을 미칠 수 있기 때문에 목록을 제시해야 한다⁷⁾. 메타분석이 시행된 경우 출판 비뚤림(Item 15)에 대한 평가도 중요한 항목으로 메타분석을 시행한 논문에서 모두 누락되었다. 출판 비뚤림은 통계적으로 의미 있는 결과를 보고한 연구만 선택적으

로 출판되는 현상으로 체계적 문헌고찰 결과에 영향을 미칠 수 있기에 funnel plot이나 그래프 등으로 보고해야 한다⁷⁾. 이는 PRISMA의 전체적인 비뮌림 위험(15번)과 전체적 비뮌림 위험 결과(22번) 항목과 공통된 항목이다. 이 외에 AMSTAR 2의 비중요 항목으로 포함된 문헌의 연구비 출처에 대한 문항(Item 10)과 연구자금을 포함하여 이해상충의 잠재적인 원인에 관한 문항(Item 16)이 모두 누락되었다. 이는 PRISMA에서도 연구비 출처(27번)에 대한 보고와 동일한 문항이다. 포함된 문헌의 연구비 출처는 연구의 결과에 영향을 미칠 수 있으므로 그 출처에 대한 정보도 제공함을 권장하고 있으며 모든 연구와 마찬가지로 체계적 문헌고찰에도 연구 자금과 이해상충의 잠재적 원인에 대하여 명시해야 연구의 편향을 파악하는 데 도움이 된다^{5,33)}. 이 외에도 문헌의 선택과정과 데이터 추출과정에서 중복으로 수행했는지에 대한 문항(Item 5, 6)에 있어서 비교적 높은 비율로 누락되어 있었다. 문헌선택과 데이터 추출과정에서 오류를 줄이기 위해 이 부분에 대하여도 향후 개선이 필요하다. 추가적으로 PRISMA 점검표 중 공통으로 누락된 항목은 프로토콜과 등록(5번)과 추가분석(16번)이다. 사전에 체계적 문헌고찰 연구 계획을 수립하여 이를 프로토콜로 출판하거나 등록하여 정보를 제시하는 것은 연구의 편향을 줄일 수 있다⁵⁾. 체계적 문헌고찰 연구에 있어서 프로토콜 등록에 대한 인지가 부족하거나 중요성을 간과한 경우가 많으므로 향후 이에 대한 실천이 필요할 것으로 보인다³⁷⁾.

본 연구를 통해 대한한방부인과학회지

에 게재된 체계적 문헌고찰 논문의 질을 평가하고 향후 개선점을 제시하였으나 몇 가지 한계점이 있었다. 기존의 한의학 관련 학회지에 보고된 선행연구에 비해 선정된 분석대상 논문 수는 많았지만, 국외에 비해서는 적은 편이었다. 또한, AMSTAR 2와 PRISMA 점검표 도구에 따라 평가를 시행하였으나, 평가 기준에 대하여서 연구자간의 주관적인 관점을 배제할 수는 없었다. 이에 저자 2명이 독립적으로 평가 후 결과의 불일치가 있는 경우 서로 논의를 통해 해결하였으나, 선행 연구의 결과와 비교하여 해석하는 데는 제한점이 있을 수 있다. 본 연구는 국내 대한한방부인과학회지에 게재된 논문만을 통해 질 평가 결과를 제시하였기에 국내의 한방부인과학과 관련된 주제로 해외학회지에 게재된 논문은 포함하지 않았으므로 전체 한방부인과학과 관련된 체계적 문헌고찰의 질 평가 결과와는 다를 수 있다. 본 연구에서 공통으로 누락된 항목 중 배제된 문헌 목록 제시나 이해상충에 관한 항목 등은 학회지의 분량 제한이나 투고규정과 관련이 있을 수 있으며, PRISMA 점검표를 투고시 필수적으로 제출하지 않아도 되는 시스템의 영향도 있을 수 있다. 향후 한방부인과학과 관련 해외학회지에 발표된 체계적 문헌고찰과 연구의 질을 비교해 본다면 국내의 제한점과 개선점에 대해서 좀 더 자세히 제시해 볼 수 있을 것이다.

향후 대한한방부인과학회지에 게재된 체계적 문헌고찰 논문의 정확도와 신뢰도를 높이고 질을 향상시키기 위해서는 체계적 문헌고찰 논문 게재 시 PRISMA 점검표를 활용하고 AMSTAR 2를 적용하여 점검하는 것이 필요할 것이다.

V. 결 론

본 연구는 2021년 2월까지 대한한방부인과학회지에 게재된 체계적 문헌고찰 논문 19편을 검토하여 AMSTAR 2와 PRISMA 점검표를 활용하여 방법론 및 질적 평가를 시행하여 다음과 같은 결론을 얻었다.

1. AMSTAR 2로 평가한 결과 선정된 체계적 문헌고찰의 방법론적 질 수준은 낮음(63.2%) 또는 매우 낮음(36.8%)이었다.
2. PRISMA 점검표로 보고의 질 평가 결과 27개의 문항 중 최대 22개에서 최소 15개(81.5~55.6%)의 문항을 보고하였으며, 평균 18.6개의 문항을 보고하여 국외 선행연구에 비해 보고의 질이 대체적으로 낮았다.
3. 19편의 논문에서 공통적으로 누락된 내용으로는 배제된 문헌 목록 제공과 출판 비फल을 포함한 전체적인 비फल 위험 평가, 연구비 출처와 이해상충의 잠재적 원인에 대한 내용이 있었으며, 프로토콜과 등록과 추가분석에 대해서도 모두 누락되었다.
4. 향후 대한한방부인과학회지에 게재되는 체계적 문헌고찰 논문의 질을 높이기 위해서 PRISMA 점검표와 AMSTAR 2의 항목에 따라 수행하고 보고하는 것이 필요하며, 이는 근거수준이 높은 근거기반 형성에 도움이 될 것이다.

Received : Apr 16, 2021

Revised : Apr 19, 2021

Accepted : May 28, 2021

References

1. Mulrow CD. Rationale for systematic reviews. *BMJ*. 1994;309(6954):597-9.
2. Pearson A, Jordan Z, Munn Z. Translational science and evidence-based healthcare: a clarification and reconceptualization of how knowledge is generated and used in healthcare. *Nurs Res Pract*. 2012;2012(1):e792519.
3. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet*. 2009;374(9683):86-9.
4. Moher D, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. *Quality of Reporting of Meta-analyses*. *Lancet*. 1999;354(9193):1896-900.
5. Liberati A, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*. 2009;6(7):e1000100.
6. Shea BJ, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol*. 2009;62(10):1013-20.
7. Shea BJ, 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.
8. Kim JH, Kim AK. A quality assessment

- of meta-analyses of nursing in South Korea. *J Korean Acad Nurs.* 2013; 43(6):736-45.
9. Shim SB, Lee JA, Lee HL. Evaluation of the Reporting and Methodological Quality of the Systematic Review from the Journal of Pediatrics of Korean Medicine. *J Pediatr Korean Med.* 2020;34(1):26-36.
10. Song EH, et al. Methodological and Reporting Quality of Systematic Reviews Published in Journal of Society of Preventive Korean Medicine. *J Prev Med Public Health.* 2019; 23(2):67-76.
11. Haidich AB. Meta-analysis in medical research. *Hippokratia.* 2010;14(Suppl 1):29-37.
12. Kim SH, et al. Recent Acupuncture Therapy for Polycystic Ovary Syndromes : Systematic Review. *J Korean Obstet Gynecol.* 2014;27(2):71-82.
13. Baek SE, et al. Systematic Review of Fumigation Therapy for Atrophic Vaginitis. *J Korean Obstet Gynecol.* 2016;29(1):92-101.
14. Bae JE, Yoon YJ. The Effect of Integrated Medicine for Polycystic Ovary Syndrome Subfertility Patients: A Systematic Review and Meta-Analysis. *J Korean Obstet Gynecol.* 2017;30(4):59-76.
15. Jeong SY, et al. The Effectiveness of Acupuncture and Moxibustion Treatment for Mastitis: A Systematic Review. *J Korean Obstet Gynecol.* 2017;30(3):29-39.
16. Kim YJ, et al. The Clinical Effect on Acupuncture Treatment of Postpartum Hypogalactia: A Systematic Review of Randomized Controlled Trials. *J Korean Obstet Gynecol.* 2017;30(3):79-91.
17. Lee HI, et al. The Effectiveness of Moxibustion Treatment for Primary Dysmenorrhea: A Systematic Review and Meta-Analysis. *J Korean Obstet Gynecol.* 2017;30(2):107-19.
18. Park NC, et al. Electroacupuncture for Treatment of Primary Dysmenorrhea: Systematic Review of Randomized Controlled Trials. *J Korean Obstet Gynecol.* 2017;30(2):81-92.
19. Song MH, Hong KK, Choi CM. Systematic Review of Acupoint Catgut Embedding Therapy for Climacteric Hot Flush. *J Korean Obstet Gynecol.* 2017;30(2):71-80.
20. Park EJ, Jo HG, Choi JH. Traditional Herbal Medicine for Dysfunctional Uterine Bleeding versus Western Medicine: A Systematic Review of Randomized Controlled Trials. *J Korean Obstet Gynecol.* 2018;31(4):110-27.
21. Park KD, Bae JE, Yoon YJ. Systematic Review of Herbal Medicine Fumigation Treatment for Mycotic Vaginitis (Candida Vaginitis). *J Korean Obstet Gynecol.* 2018;31(3):20-32.
22. Yoon JG, et al. Electroacupuncture for Treatment of Polycystic Ovarian Syndrome: Systematic Review of Randomized Controlled Trials. *J*

- Korean Obstet Gynecol. 2018;31(4):39-53.
23. Bae JE, Park KD, Yoon YJ. Acupoint Sticking Therapy for Endometriosis : Systematic Review. J Korean Obstet Gynecol. 2019;32(3):162-73.
24. Jung JW, Yoon YJ. The Effect of Herbal Medicine for Adenomyosis patients : A Systematic Review. J Korean Obstet Gynecol. 2019;32(4):87-101.
25. Nam EY. Review on Clinical Studies of Traditional Korean Medicine Treatment for Atrophic Vaginitis. J Korean Obstet Gynecol. 2019;32(3):245-62.
26. Park EJ, Jo HG. Application of Traditional East Asian Herbal Medicine for Atrophic Vaginitis: A Systematic Review of Randomized Controlled Trials. J Korean Obstet Gynecol. 2019;32(3):227-44.
27. Ha SJ, et al. Electroacupuncture for Women with Overactive Bladder : A Systematic Review and Meta-analysis of Randomized Controlled Trials. J Korean Obstet Gynecol. 2020;33(3):1-19.
28. Kim SG, Yoon YJ. Review of Randomized Controlled Trials on the Therapeutic Effects of Herb Medicine for Premenstrual Syndrome. J Korean Obstet Gynecol. 2020;33(4):36-55.
29. Lee SJ, et al. Herbal Medicine for Cervical Intraepithelial Neoplasia After Loop Electrosurgical Excision Procedure: A Systematic Review and Meta-Analysis. J Korean Obstet Gynecol. 2020;33(2):44-62.
30. Park HR, Jo HG. The Efficacy of Moxibustion for Female Stress Urinary Incontinence: a Systematic Review. J Korean Obstet Gynecol. 2020;33(4):1-22.
31. Petrie A, Bulman JS, Osborn JF. Further statistics in dentistry Part 8: Systematic reviews and meta-analyses. Br Dent J. 2003;194(2):73-8.
32. Sauerland S, Seiler CM. Role of systematic reviews and meta-analysis in evidence-based medicine. World J Surg. 2005;29(5):582-7.
33. Whiting P, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. Recent Prog Med. 2018;109(9):421-31.
34. Tian J, et al. The methodological and reporting quality of systematic reviews from China and the USA are similar. J Clin Epidemiol. 2017;85:50-8.
35. Tian ZY, et al. An Overview of Systematic Reviews and Meta-analyses on Acupuncture for Post-acute Stroke Dysphagia. Geriatrics (Basel). 2019;4(4):68.
36. Li YX, et al. Effectiveness and Safety of Acupuncture for Migraine: An Overview of Systematic Reviews. Pain Res Manag. 2020;2020:3825617.
37. Tawfik GM, et al. Protocol registration issues of systematic review and meta-analysis studies: a survey of global researchers. BMC Med Res Methodol. 2020;20(1):213.