

국내 사회복지 관련 메타분석 연구의 질 평가

조미경¹, 김희영^{2*}

¹을지대학교 간호대학, ²남서울대학교 아동복지학과

A Quality Assessment of Meta-Analyses Research in Social Work

Mi-Kyoung Cho¹, Hee-Young Kim^{2*}

¹College of Nursing, Eulji University

²Department of Child Welfare, Namseoul University

요약 본 연구의 목적은 사회복지관련 메타분석 연구를 대상으로 Assessment of Multiple Systematic Reviews (AMSTAR) 도구를 이용하여 질 평가를 하기 위한 서술적 조사연구이다. ‘메타분석’, ‘사회복지’, ‘사회사업’을 키워드로하여 2000년부터 2015년에 KISS, DBpia, RISS4U 데이터베이스를 이용하여 출판된 연구 중 42개의 메타분석 연구에 대해 질 평가를 실시하였다. 수집된 자료는 기술통계, 독립집단 t-test, 일원분산분석으로 분석되었다. 평균 AMSTAR 질 평가점수는 4.76±1.66점이었고, 낮은 수준이 19개(45.2%), 중간 수준이 22개(52.4%), 높은 수준이 1개(2.4%)이었다. 질 평가점수는 연구의 특성인 출판년도, 대상자, 분석논문수, DB의 수, 저널유형, 재정지원 유무, 연구의 질 평가유무, 분석대상논문을 추출흐름도 유무, 주제에 따라 차이가 분석되었다. 질 평가점수는 출판년도, DB의 수, 연구의 질 평가유무, 분석대상논문을 추출흐름도 유무에 따라 통계적으로 유의한 차이가 있었고, 4개 요인이 메타분석 연구의 질에 영향을 미치는 요인으로 51.9%의 설명력이 있었다. 본 연구는 메타분석 연구의 질을 향상시키기 위해 필요한 가이드라인을 제시하는데 기여하고 있다. 본 연구는 사회복지 분야 메타분석 연구의 질 향상과 분석 결과의 신뢰성 확보를 위하여 첫째, 메타분석 보고를 위한 공통적이고 구체적인 가이드라인을 마련하고 공유하는 것, 둘째, 근거기반의 실천적 개입을 위해서는 질 평가를 통해 근거강도가 높은 논문의 결과를 메타분석으로 합성하는 것, 셋째, 분석대상 논문의 연구특성을 제시하는 것, 마지막으로 문헌선택과 자료추출 과정에서 두 명이상의 연구자에 의한 독립적 실시와 최종 문헌추출을 위한 합의과정과 이해상충여부가 보고되어야 함을 제안하고 있다.

Abstract This study was conducted to evaluate the quality of meta-analysis of social work in South Korea using the Assessment of Multiple Systematic Review (AMSTAR). Electronic databases including the Korean Studies Information Service System (KISS), DBpia, and RISS4U were searched for 'meta-analysis', 'social work', and 'social welfare' from 2000 to December 2015, and 42 meta-analysis studies were included. Data were analyzed using descriptive statistics, t-tests, and ANOVA. The mean score for AMSTAR evaluation was 4.766±1.66, while 19 studies (45.2%) were classified at the low-quality level, and 22 (52.4%) were at the moderate-quality level. The scores of quality assessment were analyzed by publication year, participants, number of studies included, number of DB, reporting study quality, extraction diagram and topics. The findings indicated that the following changes should be implemented to improve the quality and reliability of meta-analysis results in social work research: 1) common reporting guidelines should be provided for the social work field, 2) quality analyses of each study should be conducted to achieve a high level of evidence of effectiveness of social work interventions, 3) the characteristics of the included studies should be provided, and 4) a consensus and procedure based on at least two independent data extractors should be reported.

Keywords : Meta-Analysis, Quality Assessment, Social Work

*Corresponding Author : Hee-Young Kim(Namseoul Univ.)

Tel: +82-41-580-3053 email: hkimjk@naver.com

Received July 11, 2016

Revised (1st August 12, 2016, 2nd September 1, 2016)

Accepted October 7, 2016

Published October 31, 2016

1. 서론

1.1 연구의 필요성

최근 사회복지실천에 있어서 개입 효과성 확보와 클라이언트 중심의 개입을 강조하면서 근거기반실천(evidence-based practice)모델에 근거한 접근을 강조하고 있다[1-2]. 이러한 사회복지 분야의 근거기반실천은 조사연구, 실험연구, 메타분석과 같은 과학적 지식의 활용을 강조하여 얻어진 연구결과들을 증거로 활용하여 사회복지실천개입을 체계적으로 실현할 것을 제시하고 있다[3]. 특히, 메타분석은 다양한 개별연구결과를 체계적이고 계량적으로 종합하고 분석하여 개입의 효과를 도출하는 가장 높은 수준의 과학적 근거를 제공하는 연구방법[2,4]이다. 또한 사회복지실천 개입의 효과성 평가를 도출하기 위하여 최근 사회복지연구 분야에서 많이 사용하고 있는 연구방법이기도 하다. 사회복지분야에서는 국외의 경우 1980년대 후반부터 메타분석에 대한 연구물이 나오기 시작하였고[5-6], 현재까지 317편의 학술지가 국제 데이터베이스인 EBSCOhost에서 검색되고 있다. 국내에서도 사회복지분야에서 발표된 연구 중 메타분석을 검색용어로 하였을 때, 2000년 이후 현재까지 15년 동안 국내 데이터베이스인 RISS에서만도 학위논문 16편, 학술지 게재논문 21편이 검색되었다. 이를 통해 국내외적으로 메타분석 연구가 사회복지분야에서도 지속적으로 팽창하고 있음을 알 수 있다.

메타분석은 특정 주제에 관련해 연구 대상의 표집, 측정도구, 자료분석 방법을 달리한 개별 연구들의 상이한 결과로부터 개입효과크기를 얻어내고 통계적으로 종합하여, 연구결과의 객관적인 일반화를 할 수 있도록 하는 연구의 통합(research synthesis) 기법이다[7]. 그러나 메타분석은 동질성과 타당성이 높거나 낮은 상이한 개별 연구들을 종합하는 방법이기 때문에, 연구물 선택의 편차 및 독립성 가정위반 등의 문제를 야기할 수 있고 이로 인한 왜곡된 결론을 도출할 가능성에 대한 비판을 받고 있다[7]. 메타분석의 연구결과로부터 도출된 결과의 적용과 해석에 대한 정확한 판단을 위해서는 메타분석의 연구방법과 결과에 대한 신뢰성과 타당성이 먼저 확보되어야 한다[2]. 따라서 메타분석의 보고지침이나 평가지침에 따른 자료 수집 절차와 방법, 통계적 분석 방법 등이 적절하게 수행되었는지에 대한 연구 방법론적 검증은 메타분석으로 인한 오류를 줄이면서 연구 간 불일치

를 해결하고 신뢰도가 높은 결과 확보를 위하여 필수적인 과정이라 할 수 있다.

국외의 경우, 메타분석 연구결과의 신뢰성 확보를 위하여 1990년 대 이후부터 메타분석 연구보고에 대한 질 평가 실시의 중요성을 보고하고 있으며, 메타분석의 질 평가 자체가 연구결과의 신뢰성 확보를 위한 도구로서 꾸준히 개발되고 제안되고 있다[8-9]. 지금까지 메타분석 보고의 질을 높이기 위한 다양한 메타분석의 질 평가 도구들로 무작위 실험연구의 질 평가를 위한 21개의 평가기준을 제시한 Moher 등[9]이 개발한 QUOROM (quality of reporting of meta-analysis of randomized controlled trials), 2009년 QUOROM의 수정판으로 보완된 27-35개 체크리스트 항목을 제시하고 있는 PRISMA(Preferred Reporting Items for Systematic Reviews and Meta-analysis), 미국심리학회에서 QUOROM과 PRISMA 등의 메타분석 보고기준을 검토하여 교육과 심리학분야에 적절하도록 보완 및 수정하여 제시한 APA (American psychological association) 보고기준이 있다[10].

본 연구에서는 메타분석의 방법론적 검증과 연구의 질 평가를 위한 11개의 항목들로 구성된 AMSTAR (assessment of multiple systematic reviews)를 사용하고 자 한다. 기존 도구에 비해 간결하고 사용이 용이하고 [11-12]. 기존 도구들이 체계적 문헌고찰의 보고기준으로만 활용되고 있는 반면, AMSTAR는 유일하게 신뢰도와 타당도가 검증된 도구이며, 보건 및 의료분야[13]뿐만 아니라 다양한 학문분야[11,14-15]에서 메타분석 연구의 객관적인 질 평가를 위해 활용되고 있다[11,16-17] 또한 국내의 경우 한국보건의료연구원에서 AMSTAR를 승인받아 번역한 한국어판 AMSTAR를 메타분석 연구의 질 평가도구로 사용할 수 있도록 제시하고 있다[18]. 사회복지분야의 경우에도 Lee, Kim, Yeum[19]이 APA에서 제시하는 기준들에 얼마나 부합하는지를 5개의 영역(제목 및 초록, 서론, 방법, 결과, 논의)을 APA (American psychological association) 보고기준에 따라 보고한 연구가 있으나, Lee, Kim, Yeum[19]의 연구는 메타분석 연구의 작성방법면에서 APA에서 제시하는 기준들에 얼마나 부합하는지를 알아보는 데에만 그쳤다.

따라서, 본 연구에서는 메타분석의 질 평가도구로써 타당도와 신뢰도를 확보한 AMSTAR 평가도구를 활용하여 국내학술지에 게재된 사회복지 분야 메타분석 연구

의 질 평가를 실시하고 사회복지실천연구에서 메타분석 연구의 질 향상 방안을 제시하며 더 나아가 근거기반실천에 기반을 둔 사회복지실천에 기여하고자 한다.

1.2 연구의 목적

본 연구의 목적은 사회복지관련 메타분석 연구를 대상으로 메타분석 질 평가도구인 Assessment of Multiple Systematic Reviews(AMSTAR) 평가기준에 따라 질 평가를 실시함으로써 메타분석 방법론적인 질 향상을 위한 방안을 제시하고자 하는 것이며 구체적인 목적은 다음과 같다.

- 첫째, 분석대상 메타분석 연구의 특성을 파악한다.
- 둘째, 분석대상 메타분석 연구의 질 평가 수준을 파악한다.
- 셋째, 분석대상 메타분석 연구의 특성에 따른 질 평가 점수를 파악한다.

2. 연구방법

2.1 연구설계

본 연구는 사회복지관련 메타분석 연구를 AMSTAR 평가도구를 활용하여 논문의 질을 평가하는 서술적 조사 연구이다.

2.2 분석대상 논문의 선정 및 자료수집

사회복지관련 메타분석 논문의 질 평가를 위해 한국 학술정보원(Korean studies information service system [KISS]), DBpia, RISS4U 데이터베이스를 이용하여 한국어로 ‘메타분석’, ‘사회복지’, ‘사회사업’을 주요어로 하여 2000년부터 2015년 12월까지 발간된 논문을 2명의 연구자가 각각 독립적으로 검색 및 자료추출을 실시하였다(Figure 1). 분석대상 논문의 선정기준은 첫째, 메타분석 연구, 둘째, 사회복지관련 효과크기를 분석한 메타분석 연구, 셋째, 한국어로 발표한 메타분석 연구이었다. 검색된 자료는 총 128편의 연구물이었고, 이중 연구자의 합의에 의해 분석에 이용된 메타분석 연구는 총 42 편이었다.

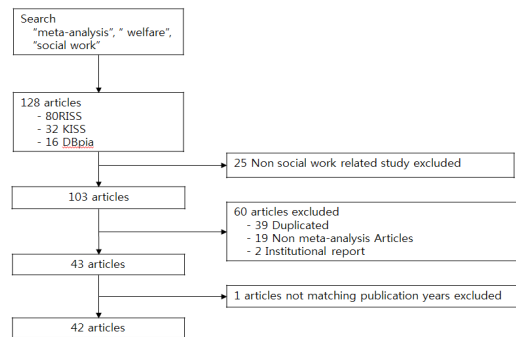


Fig. 1. Literature search strategy

2.3 연구도구

2.3.1 메타분석 연구의 질평가 도구

본 연구에서는 메타분석 연구의 질 평가를 위해 Shea 등[20]이 개발하고 Kim 등[18]이 국내에서 한국어로 번역하여 신뢰도와 타당도가 검증된 AMSTAR 한국어 버전을 사용하였다. 이 도구는 ‘사전에’ 체계적 문헌고찰의 계획이 수립되었는가?, 문헌선택과 자료추출을 여러 명이 수행하였는가?, 포괄적인 문헌검색을 하였는가?, 포함기준에 출판상태(예: 회색문헌)가 사용되었는가?, 포함 및 배제된 연구목록이 제시되었는가?, 포함된 연구의 특성이 제시되었는가?, 포함된 연구의 질이 평가되고 기술되었는가?, 포함된 연구의 질은 결론을 도출하는데 적절히 사용되었는가?, 개별연구의 결과를 결합한 방법이 적절하였는가?, 출판 비틀림의 가능성을 평가하였는가?, 이해상충이 기술되었는가? 의 11개 항목으로 구성되었으며, 각 항목에 대해 ‘예’, ‘아니오’, ‘대답할 수 없음’, ‘적용할 수 없음’으로 판단하여 체크하였다. 점수는 ‘예’는 1점을 주고, ‘아니오’, ‘대답할 수 없음’, ‘적용할 수 없음’은 0점으로 기록하였고, 11개 항목의 점수를 합산하여 총점을 산출하였다. 합산된 질 평가점수는 0-4점은 낮은 수준, 5-8점은 보통 수준, 9-11점은 높은 수준으로 해석하였다[18]. Shea 등[20]이 본 도구를 개발당시 Intra Class Coefficients는 0.91, 측정자간 일치도는 Cohen’s kappa=0.70이었고, 본 연구에서 Intra Class Coefficients는 0.95, 측정자간 일치도는 Cohen’s kappa=0.88이었다.

2.4 연구진행절차 및 자료수집

질 평가도구를 이용한 개별 논문에 대한 코딩은 논문 제목, 출판년도, 대상자, 주제, 저널유형, 데이터베이스

수, 분석대상 논문수, 재정지원유무, 자료추출흐름도 유무, 질 평가유무, 질 평가 각 항목의 점수와 총점을 기록할 수 있는 코딩북을 만들어서 공유한 후 연구자 2인이 개별적으로 실시하였다. 질 평가도구의 항목기준을 명확하게 이해하기 위해 AMSTAR도구를 이용하여 메타분석 논문의 질 평가를 수행한 경험이 있는 교수의 자문을 받아서 각 항목에서의 판단기준을 명확하게 하였다. 2편의 논문을 대상으로 AMSTAR도구를 이용하여 평가하고 항목에 대한 이해와 판단기준을 합의하는 사전훈련을 시행하였다. ‘사전에’ 체계적 문헌고찰의 계획이 수립되었는가? 항목은 연구문제나 목적이 제시되어 있고 분석대상 문헌의 포함기준이 있는 경우에, 문헌선택과 자료추출을 여러 명이 수행하였는가? 항목은 문헌선택과 자료추출을 2명이상 각각 한 경우에, 포괄적인 문헌검색을 하였는가?는 적어도 2개의 전자자료원을 이용하여 검색한 경우, 검색시 출판년도와 주제가 제시되어 있는 경우 ‘예’로 판단하였으며, 포함기준에 출판상태(예: 회색문헌)가 사용되었는가?는 문헌 검색이 학술논문이나 보고서 등까지 이루어졌다는 언급이 있으면 ‘예’로 판단하였다. 포함 및 배제된 연구목록이 제시되었는가?는 연구특성이 표로 제시되었거나 참고문헌으로 있는 경우에, 이해상충이 기술되었는가?는 분석대상 논문의 연구비 출처가 명확하게 제시되어 있는 경우에 ‘예’로 판단하였다.

2.5 자료분석

수집된 자료는 SPSS 21.0 프로그램을 사용하여 분석하였다. 분석대상 연구의 특성과 분석대상 연구의 질 평가를 위해 11개 평가항목에 대해 실수, 백분율, 평균과 표준편차로 분석하였다. 메타분석 연구의 특성에 따른 질 평가 점수는 독립집단 t-검정(independent t-test), 일원분산분석(one way ANOVA)으로 분석하였다. 메타분석 연구의 질 평가에 영향을 미치는 요인을 파악하기 위해 multiple regression(enter)으로 분석하였다. 모든 통계적 유의수준은 $p < .05$ 로 하였다.

3. 연구결과

3.1 분석대상 연구의 특성

분석대상 연구의 출판년도는 2000~2009년 11편,

2010~15년 31편이었고, 연구대상은 건강인이 30편, 환자 9편, 프로그램 3편이었다. 분석대상 논문의 수는 60편 미만인 24편, 60편이상인 18편이었으며, 데이터베이스 수는 4개 미만인 22편, 4개 이상이 19편이었다. 저널 유형은 연구재단등재지 및 등재후보지가 37편, 미등재가 1편, 학술대회 발표자료가 4편이었다. 재정지원을 받은 연구가 11편이었고, 연구의 질 평가가 이루어진 연구가 1편이었으며, 자료추출 흐름도가 있는 연구가 7편이었다. 주제는 다양하였고, 프로그램의 효과를 본 연구가 22편, 관련변인을 파악하기 위한 연구가 20편이었다 [Table 1].

Table 1. Characteristics of Studies (N=42)

ID	year	Participants	N. of study	N. of DB	support	study quality	diagram
1	2015	Adolescents	25	8	No	Yes	Yes
2	2015	Social worker	76	4	Yes	No	No
3	2015	Program	33	5	Yes	No	No
4	2014	Elderly	76	7	No	No	Yes
5	2014	Elderly patients with chronic disease	123	11	No	No	Yes
6	2014	Care helper	83	4	No	No	No
7	2014	Teacher	20	5	No	No	No
8	2014	Youth	84	6	Yes	No	No
9	2014	Persons with severe mental illness	172	6	No	No	Yes
10	2014	Adolescent	82	6	Yes	No	Yes
11	2014	ADHD children	31	2	No	No	No
12	2013	In-home elderly	155	3	No	No	No
13	2013	Program	30	3	No	No	No
14	2013	Persons with severe mental illness	67	4	Yes	No	No
15	2013	Social worker	111	4	No	No	No
16	2012	Elderly	21	3	Yes	No	No
17	2012	Child abuse experience	16	5	No	No	No
18	2012	All	31	3	Yes	No	No
19	2012	Older adult	41	Not mentioned	No	No	No
20	2012	Older adult	54	3	No	No	No
21	2012	Child abuse	41	6	Yes	No	No
22	2012	Depression patient	23	3	No	No	No
23	2012	People with schizophrenia	72	4	No	No	No
24	2012	The disabled	104	3	No	No	No
25	2012	The aged	102	2	No	No	No
26	2012	Adolescent	60	3	No	No	Yes
27	2012	Adolescent	101	4	No	No	No
28	2011	Program	25	2	Yes	No	No
29	2010	Children, student	43	3	No	No	Yes
30	2010	Alcoholics	22	4	No	No	No
31	2010	Children of parentless family	55	3	No	No	No

32	2009	Mentally ill	20	4	No	No	No
33	2009	Children, adolescent, parent	22	2	No	No	No
34	2008	People with severe disabilities	6	2	No	No	No
35	2008	Children's exposure to domestic violence	45	8	Yes	No	No
36	2005	Parent-children	65	1	No	No	No
37	2005	Adolescent	65	3	No	No	No
38	2004	Adolescent	39	2	No	No	No
39	2004	Student, adult	36	2	No	No	No
40	2003	Adolescent	93	2	No	No	No
41	2002	Children-adolescent	54	1	No	No	No
42	2000	Client	29	3	Yes	No	No

3.2 연구의 질평가 점수

AMSTAR도구를 이용한 분석대상 연구의 메타분석 질 평가 점수 평균은 4.76±1.66점(범위: 2~10점)이었고, 낮은 수준(0-4점)이 19편, 보통 수준(5-8점)이 22편, 높은 수준(9-11점)이 1편이었다. 연구의 질 평가 항목 중 “사전에 체계적 문헌고찰의 계획이 수립되었는가?”는 42편(100%) 모두 기술하고 있었고, “포함기준에 출판상태(예: 회색문헌)가 사용되었는가?”는 41편, “포괄적인 문헌검색을 하였는가?”는 38편, “개별연구의 결과를 결합한 방법이 적절하였는가?”에 대해서는 31편으로 대부분의 논문이 평가기준을 충족하고 있었다. “포함 및 배제된 연구목록이 제시되었는가?”는 15편, “문헌선택과 자료추출을 여러 명이 수행하였는가?”는 13편, “출판 비탈림의 가능성을 평가하였는가?”는 12편으로 나타났고, “포함된 연구의 특성이 제시되었는가?”는 5편, “포함된 연구의 질이 평가되고 기술되었는가, 포함된 연구의 질은 결론을 도출하는데 적절히 사용되었는가?”는 1편, “이해상충이 기술되었는가?”에 대해서는 1편 뿐이었다 [Table 2].

Table 2. Scores of Assessment of Multiple Systematic Reviews

no.	Items	Answer of yes n (%)	M±SD
1	Was an 'a priori' design provided?	42 (100.0)	
2	Was there duplicate study selection and data extraction?	13 (31.0)	
3	Was a comprehensive literature search performed?	38 (90.5)	
4	Was the status of publication (i.e. grey literature) used as an inclusion criterion?	41 (97.6)	

5	Was a list of studies (included and excluded) provided?	15 (35.7)	
6	Were the characteristics of the included studies provided?	5 (11.9)	
7	Was the scientific quality of the included studies assessed and documented?	1 (2.4)	
8	Was the scientific quality of the included studies used appropriately in formulating conclusions?	1 (2.4)	
9	Were the methods used to combine the findings of studies appropriate?	31 (73.8)	
10	Was the likelihood of publication bias assessed?	12 (28.6)	
11	Was any conflict of interest stated?	1 (2.4)	
Level of quality	Low (0-4)	19 (45.2)	3.32±0.75
	Moderate (5-8)	22 (52.4)	5.77±0.81
	High (9-11)	1 (2.4)	10.00±0.00
	Total	42 (100.0)	4.76±1.66

3.3 연구특성에 따른 질평가 점수비교

AMSTAR도구를 이용한 질 평가 점수는 분석대상 연구의 출판연도가 2010년 이전보다는 2010년 이후에 출판된 논문들이 질 평가점수가 높았고($t=-3.40, p=.002$), 분석에 이용된 데이터베이스의 수가 4개미만인 경우보다 4개 이상인 경우가 질 평가점수가 높은 것으로 나타났으며($t=-2.33, p=.025$), 통계적으로 유의한 차이를 보였다. 질 평가 점수를 보고한 논문($t=-3.63, p=.001$)과, 자료추출흐름도가 있는 경우가 그렇지 않은 경우보다 질 평가점수가 높은 것으로 나타났으며($t=-3.57, p=.001$)와, 통계적으로 유의한 차이를 나타냈다. 그러나, 대상자, 분석대상 연구의 수, 저널유형, 재정지원여부, 주제에 따라서는 차이를 보이지 않았다[Table 3].

Table 3. Comparison of Scores of Assessment of Multiple Systematic Reviews by Characteristics of Studies

Characteristics	Categories	n (%)	M±SD	t or F (p)
Publication year	2000-2009	11 (26.2)	3.45±1.04	-3.40 (.002)
	2010-2015	31 (73.8)	5.23±1.61	
Participants	Healthy people	30 (71.4)	4.73±1.74	0.19 (.829)
	Patients	9 (21.4)	4.67±1.32	
	Program	3 (7.1)	5.33±2.31	
Number of studies included	< 60	24 (57.1)	4.71±1.94	-0.24 (.813)
	≥ 60	18 (42.9)	4.83±1.25	
Number of DB	< 4	22 (52.4)	4.27±1.58	-2.33 (.025)
	≥ 4	19 (45.2)	5.42±1.57	
Journal type	KCI (E)	37 (88.1)	4.76±1.67	-0.05 (.957)
	The others	5 (11.9)	4.80±1.79	
Financial support	Yes	11 (26.2)	4.54±1.75	0.50 (.622)
	No	31 (73.8)	4.83±1.66	

Reporting study quality	Yes	1 (2.4)	10.00±0.00	-3.63 (.001)
	No	41 (97.6)	4.63±1.46	
Extraction diagram	Yes	7 (16.7)	6.57±1.62	-3.57 (.001)
	No	35 (83.3)	4.40±1.44	
Topics	Effect of program	22 (52.4)	4.95±1.86	0.78 (.438)
	Influencing factor	20 (47.6)	4.55±1.43	

4. 논의

본 연구는 국내 사회복지 분야에서 2000년부터 최근 2015년 12월까지 출간된 메타분석 연구물을 AMSTAR 평가도구를 사용하여 메타분석의 방법론적인 질향상 방안을 제시하고 근거기반 사회복지실천 향상에 기여하고자 수행되었다.

4.1 메타분석 연구의 질 평가 수준

AMSTAR의 평가 점수는 점수범위에 따라서 수준이 낮음(0-4), 보통(5-8), 높음(9-11)으로 구분되는데, 본 연구에서 총 42편의 메타분석연구를 분석한 결과, 논문 전체의 질 평가 수준은 45.2%가 '낮음', 52.4%가 '보통', 2.4%가 '높음'으로 11점 만점에 평균 4.76점으로 사회복지분야에서의 메타분석 연구의 질 수준이 낮게 나타났다. 이러한 결과는 동일한 AMSTAR평가도구를 사용한 간호학분야 메타분석 논문[13]의 질 평가결과(5.61점)보다 수준이 낮았으며, 국내 운동 관련 메타분석 논문[11]의 질 평가 결과(4.18점)와 노인관련 메타분석 연구[15]의 질 평가 결과(4.10점)와는 비슷한 수준이었다. 이는 다양한 학문과 주제 분야에서 근거기반 개입의 필요성을 인식하고 근거확보를 위한 메타분석연구가 실행되고 있으나 메타분석 연구의 질은 낮은 것을 알 수 있었다.

AMSTAR의 세부 항목별 점수를 살펴보면 다음과 같다.

첫째, '포함된 연구의 질이 평가되고 기술되었는가?'와 '포함된 연구의 질은 결론을 도출하는데 적절히 사용되었는가?'의 평가문항은 본 연구에 포함된 연구 중 1편(2.4%)에서만 실시한 것으로, 간호학 분야에서는 31.0%가 질평가를 실시하는 반면[13] 사회복지 분야에서는 질평가가 이루어지지 않음을 알 수 있었다. 운동[11]과 노인관련연구[15]에서 질평가 결과가 각각 12.1%와 0.13%로 본 연구와 비슷하게 낮았는데, 이러한

결과들은 사회복지분야 뿐 아니라 국내 메타분석 연구 전반에 있어서 아직까지 개별 연구의 질 평가가 중요하게 다루어지고 있지 않음을 시사한다. 메타분석 수행에 있어서 개별 연구의 질 평가는 연구 결과의 신뢰성을 확보하는 매우 중요한 요인으로, 개별연구 논문의 연구설계, 문헌의 질, 연구방법의 타당성 및 일관성 등을 고려한 질 평가를 메타분석 최종결과에 반영하여 제시할 때 연구결과의 신뢰성과 타당성이 높아진다고 할 수 있다.

무작위실험설계(randomized experimental design)를 이용한 개입효과성 연구가 많은 의료연구 분야와는 달리, 사회복지연구분야에서는 유사실험설계(quasi-experimental design)가 보편적으로 사용되는 만큼 적절한 질 평가도구를 선택하는 것 또한 중요하다. 비무작위 연구(non-randomized study)의 질 평가를 위한 RoBANS(risk of bias assessment tool for non-randomized study), GRADE(the grading of recommendations, assessment, development and evaluation), 그리고 MINORS(methodological items for non-randomized studies) 등과 같은 도구를 사용하여 사회복지분야에서도 메타분석 연구결과의 수준을 평가하고, 결과와 논의에서 질 평가가 연구결과에 미치는 영향들을 올바르게 해석하는 것이 필요하다.

둘째, '이해상충이 기술 되었는가?'에 관한 항목 또한 대상논문의 2.4%인 1편의 논문에서만 언급이 되었다. 즉, 개별 연구를 실행할 경우 체계적인 문헌고찰이나 포함된 연구들의 연구비 출처가 명확하게 제시 되었는지의 여부를 AMSTAR평가에서는 요구하지만 본 연구대상 논문들에서는 이해상충 여부를 보고한 연구는 거의 없었다. 이러한 결과는 국내 학술지의 경우 연구비 출처에 대한 명시를 하고는 있으나 외국의 학술지처럼 이를 반드시 보고해야 하는 지침이 없기 때문이다. 따라서 국내 학술지도 메타분석연구의 보고지침을 연구자들에게 명확하게 공지하는 것이 필요하다.

셋째, AMSTAR평가 문항 중 '포함된 연구의 특성이 제시되었는가?'에 대해서 개별연구의 연령, 인종, 성별 등과 같은 연구대상, 개입방법, 결과가 표 형태가 아니어도 연구내용에 제시되어야하나 대부분의 연구는 개별연구에 대한 구체적인 특성들을 제시하지 않았다. 이는 간호학 분야 97.6%, 노인관련 분야 46.2%, 국내 운동관련 분야의 69.7%가 평가기준을 충족한 것에 비하여 사회복지분야에서는 11.9%로 매우 낮았다. AMSTAR평가도구

를 사용한 질 평가 연구는 아니지만, Lee, Kim, Yeum[19]의 연구에서도 개별연구에 대한 표본크기, 연구결과를 포함한 표를 제공한 논문이 19.2%로 대다수의 연구가 개별연구의 특성들을 본 연구와 같이 제시하지 않은 것으로 나타났다. 개별연구의 특성을 제시하는 것은 자료의 객관성과 타당성을 확보하는 중요한 부분임에도 불구하고 사회복지분야 메타분석연구 보고에서는 일반적으로 간과되고 있다는 것을 알 수 있었다. 개별연구들의 효과크기 및 표본크기, 개입방법, 변수 등과 같은 개별연구의 특성들에 대한 정보제공은 다른 연구자가 메타분석과정을 반복했을 때 같은 연구결과를 보여줌으로서 연구결과에 대한 신뢰도를 높일 수 있게 하는 중요한 요인들이므로 추후 사회복지분야의 메타분석연구에서는 분석에 포함된 연구의 특성을 반드시 제시하여야겠다.

넷째, AMSTAR는 ‘개별연구의 결과를 결합한 방법이 적절하였는가?’와 ‘출판 편향(publication bias)의 가능성을 평가하였는가?’의 2개 문항에서 효과크기 산출과정을 평가하도록 제시하고 있다. 메타분석은 효과크기를 이용하여 개별연구 결과를 통합하고 대상 변수의 효과에 대한 방향과 크기를 제시하는 것으로서 효과크기를 어떠한 방법으로 결합하였는지 그리고 결합 과정이 적절하였는지 여부를 분석과정에 제시하는 것이 중요하다. 분석에 포함된 개별 연구 결과들이 한 모집단에서 추출된 결과인지 또는 다른 모집단에서 추출된 결과인지 여부를 카이제곱검증이나 I^2 와 같은 동질성 검증을 통하여 고정효과모형 (fixed effect model) 또는 무작위효과모형 (random effect model)을 선택하는 근거가 된다. 본 연구의 경우 73.8%의 논문이 동질성 평가를 제시하고 분석모형을 선택하는 통계적 근거와 기준을 제시했으나, 통계적으로 의미있는 연구만을 출판하는 출판편의가 사회복지 분야의 메타분석 연구결과에 미치는 영향을 통제했는가의 평가에서는 28.6%로 메타분석연구에서는 대체로 이루어지지 않는 것으로 나타났다.

AMSTAR평가에서는 출판편의의 가능성을 funnel plot과 같은 그림이나 통계적 검정인 Egger's test와 Fail-safe N 산출 등으로 확인하고 그 결과를 객관적으로 보고해야 함을 강조하였다. 또한 출판편의를 시사하는 이상치(outlier)는 메타분석 결과의 타당성을 저하시킬 수 있는 주요 요인이 될 수 있기 때문에 Trim & fill 분석이나 이상치 제거와 같은 방법으로 연구결과의 왜곡을 통제하려는 시도가 있었는지 여부를 확인해야 한다. 이는 메타분

석 결과의 타당성을 저하시킬 수 있는 주요 요인이 될 수 있기 때문이다.

다섯째, 본 연구 결과 ‘문헌선택과 자료추출을 여러 명이 수행하였는가?’에 대한 문항에서 대상 논문의 31.0%만이 기준을 충족한 것으로 나타났다. 메타분석 결과에 대한 정확성과 신뢰도에 영향을 미치는 중요한 오류 중 하나가 검색 오류이다. 메타분석에서는 문헌선택과 자료추출의 오류를 막기 위하여 사전에 정해진 연구 선택 기준에 따라서 적어도 2명의 연구자가 독립적으로 문헌을 검색하고 결과를 비교할 것을 권장한다. 뿐만 아니라 추출된 자료의 코딩과정에도 두 사람 이상이 코딩을 실시하고 불일치를 해결하는 과정을 통해 자료의 신뢰도를 높이도록 하고 있다[12]. 그럼에도 불구하고 본 연구에서 분석된 과반수의 논문들은 문헌선택과 자료추출과정에서 2명 이상의 연구자들이 참여했다고 보고하고 있으나 독립적인 문헌검색의 과정이나 문헌선정의 사결정과정을 자세히 기술하고 있지는 않았다. 1차 자료 분석으로 이루어지는 일반연구와 달리 2차 자료분석을 활용하는 메타분석연구의 경우 연구결과와 정확성과 신뢰도 확보를 위하여 사회복지분야 연구자들의 체계적이고 합리적인 자료추출 과정 및 합의과정의 실행뿐만 아니라 그 과정에 대해 분명하게 기술이 되어야 하겠다.

여섯째, ‘포함 및 배제된 연구 목록이 제시되었는가?’에 대한 항목의 경우, 포함된 연구목록을 참고문헌으로 또는 본문에 제시했는가를 평가한 항목은 본 연구대상 논문의 35.7%만이 해당기준을 충족한 것으로 나타났다. AMSTAR평가기준은 연구분석에 포함된 논문 뿐 아니라 배제된 연구 목록도 참고문헌으로 함께 제시할 것을 요구하였으나 APA나 PRISMA의 보고기준에서는 배제된 연구 목록의 경우는 배제된 이유와 개수만 간략하게 제시하도록 하고 있다. 이차자료에 의존하는 메타분석연구의 특성을 고려할 때 연구결과의 정확성을 위하여 분석에 포함된 연구 목록을 제시하는 것이 메타분석연구 게재를 위한 학술지의 공용 가이드라인이 제안될 수 있기를 기대한다.

일곱째, ‘사전에 체계적 문헌고찰 계획이 수립되었는가?’, ‘포괄적인 문헌검색을 하였는가?’, ‘포함기준에 출판상태가 사용되었는가?’와 같은 기준에 있어서는 대상 논문의 90.5%~100%가 충족한 것으로 나타났다. 특히 적어도 2개 이상의 자료원을 이용하여 검색하고, 검색연도와 주요 키워드 제시 여부에 대해서 ‘포괄적인 문헌

검색'과 회색문헌의 포함여부를 제시하는 '출판상태의 포함'의 기준항목은 간호학[13]의 64.3%~88.1%, 노인관련[15]의 62%~69%, 국내 운동관련[11]의 41.2%~54.5% 보다 본 연구결과가 높은 수준임을 알 수 있었다.

4.2 메타분석 연구 특성에 따른 질 평가

본 연구에서 연구특성에 따른 메타분석 연구의 질 평가 점수의 차이를 분석한 결과, 2010년 이전에 출판된 논문보다 이후에 출판된 논문의 질 평가 점수가 높았다. 간호학 분야와 노인관련 분야 연구에서 2010년 이후에 출판된 논문의 질이 높았다고 보고한 것과 동일한 결과이다. 이는 한국보건정보연구원의 PRISMA기준을 확산한 것이 기여를 했고, 교육과 사회과학 분야에서도 메타분석 연구물의 질을 평가하려는 시도[14,21]가 시작되었기 때문인 것으로 볼 수 있다. 또한, 분석에 이용된 데이터베이스의 수가 4개 미만인 경우보다 4개 이상인 경우가 질 평가 점수가 높은 것으로 나타났다.

AMSTAR의 경우는 2개 이상의 자료원 제공 여부를 평가하며, APA보고기준에서는 특별히 자료원의 개수를 제시하고 있지는 않으나, 선택편의(selection bias)를 제거하기 위하여 문헌검색에 있어서 다양한 자료원을 활용하는 것은 검색의 민감도(sensitivity)를 높이는 중요한 부분이다[18]. 질 평가 점수를 보고한 논문과 자료추출 흐름도가 있는 경우가 그렇지 않은 경우보다 질 평가 점수가 높은 것으로 나타났다. 메타분석을 실행하는 연구자들은 검색한 논문들의 근거를 평가하고 실행된 연구방법론의 수준을 결과에 제시함으로써 메타분석 결과의 타당성을 사회복지분야의 실무자들로 하여금 파악하고 적용할 수 있도록 하는 것이 필요하겠다. 특히 자료추출 흐름도는 다른 연구자들로 하여금 재분석(replication)을 통해 검증해 볼 수 있는 기회를 제공할 수 있기 때문에 메타분석 결과에 대한 신뢰성을 높이는 데 필수적이라 할 수 있다.

서 메타분석 연구의 타당성을 높일 수 있는 방안을 살펴 보았다. 연구에 포함된 42개 논문을 대상으로 질 평가를 실시한 결과 질 평가 점수의 평균은 4.76점(범위: 2~10 점)이었고, 낮은 수준(0-4점)이 19편, 보통 수준(5-8점)이 22편, 높은 수준(9-11점)이 1편이었다. 출판년도가 2010년 이후인 경우, 연구의 질 평가를 시행한 경우, 자료추출흐름도가 있는 경우, 데이터베이스의 수가 4개 이상인 경우가 그렇지 않은 경우보다 AMSTAR도구를 이용한 질 평가 점수에서 통계적으로 유의하게 높은 수준으로 나타났다.

본 연구는 사회복지분야에서 실시된 메타분석 연구의 질 평가를 시도한 것으로 앞으로의 사회복지연구의 메타분석 연구를 위한 제언은 다음과 같다.

첫째, 사회복지분야에서 근거기반의 사회복지를 강조하는 만큼 메타분석을 통한 연구결과와 근거를 확인하려는 연구가 증가하고 있는 반면, 메타분석 연구 질평가 수준은 낮은 것으로 나타났다. 따라서 사회복지연구 분야 안에서 메타분석을 위한 공통적이고 구체적인 가이드라인을 마련하고 공유하는 것이 필요하다. 둘째, 사회복지분야에서 근거기반의 실천적 개입을 위해서는 질 평가를 통한 근거강도가 높은 논문의 결과를 합성하는 것, 셋째, 분석대상 논문의 연구의 특성을 제시하는 것, 마지막으로 문헌선택과 자료추출 과정에서 두 명이상의 연구자에 의해 독립적으로 실시하고 최종문헌추출을 위한 합의과정을 제시하는 노력과 이해상충여부가 보고되어야 한다. 이를 통해 메타분석의 오류를 피하고 신뢰도와 타당성이 확보된 연구결과를 도출하리라 기대된다.

본 연구는 처음으로 사회복지 분야 메타분석 연구의 질평가를 AMSTAR 질 평가 도구를 사용하여 분석함으로써 사회복지 분야 메타분석 연구결과에 대한 신뢰도를 확인해 볼 수 있었다는 의의가 있다. 그러나 분석대상 논문에 국내에서 출판된 학술지만 포함하고 있으므로 본 연구결과로 도출된 결론은 제한점을 가지고 있다.

5. 결론 및 제언

본 연구는 국내 사회복지분야에서 출간된 42개 메타분석 논문의 질 평가를 실시하고 사회복지분야에서의 메타분석 연구의 특성을 파악하고 타당도와 신뢰도가 확보된 AMSTAR 질 평가 도구를 활용하여 사회복지분야에

References

- [1] Y. M. Kim, M. J. Park, "A study on social worker attitude toward the adoption of evidence-based practice", *Social Science Research Review*, vol. 23, no. 2, pp. 281-303, 2007.
- [2] S. D. Hwang, "Validation of evidence-based practice attitude scale", *Social Science Research Review*, vol. 27,

- no. 4, pp. 77-97, 2011.
- [3] Z. Gellis, W. J. Reid, "Strengthening evidence-based practice", *Brief Treatment and Crisis Intervention*, vol. 4, no. 2, pp. 155-165, 2004.
DOI: <http://dx.doi.org/10.1093/brief-treatment/mhh012>
- [4] G. V. Glass, "Primary, secondary, and meta-analysis of research", *Educational Researcher*, vol. 5, no. 10, pp. 3-8, 1976.
DOI: <http://dx.doi.org/10.3102/0013189X005010003>
- [5] L. Videka-Sherman, "Meta-analysis of research on social work practice in mental health", *Social Work*, vol. 33, no. 4, pp. 325-338, 1988.
- [6] K. Gorey, B. Thyer, D. Pawluck, "Differential effectiveness of prevalent social work practice models: a meta-analysis", *Social Work*, vol. 43, no. 3, pp. 269-278, 1998.
DOI: <http://dx.doi.org/10.1093/sw/43.3.269>
- [7] S. S. Oh, *Meta analysis : Theory and practice*, pp. 1-283, Konkuk publisher, Seoul, 2011.
- [8] D. J. Cook, D. L. Sackett, W. O. Spitzer, "Meteorologic guidelines for systematic reviews of randomized controlled trials in health care from the Postdam consultation on meta-analysis", *Journal of Clinical Epidemiology*, vol. 48, no. 1, pp. 167-171, 1995.
DOI: [http://dx.doi.org/10.1016/0895-4356\(94\)00172-M](http://dx.doi.org/10.1016/0895-4356(94)00172-M)
- [9] D. Moher, D. J. Cook, S. Eastwood, I. Olkin, D. F. Stroup, "Improving the quality of reports of meta-analyses of randomized controlled trials: the QUORUM statement", *The Lancet*, vol. 357, no. 27, pp. 1896-1900, 1999.
DOI: [http://dx.doi.org/10.1016/S0140-6736\(99\)04149-5](http://dx.doi.org/10.1016/S0140-6736(99)04149-5)
- [10] S. D. Park, I. S. Shin, "A review of meta-analysis research in early childhood education", *Journal of Early Childhood Education*, vol. 34, no. 2, pp. 393-412, 2014.
DOI: <http://dx.doi.org/10.18023/kjece.2014.34.2.018>
- [11] Y. S. Song, M. H. Gang, S. A. Kim, I. S. Shin, "Review of meta-analysis research on exercise in south Korea", *Journal of Korean Academy of Nursing*, vol. 44, no. 5, pp. 459-470, 2014.
DOI: <http://dx.doi.org/10.4040/jkan.2014.44.5.459>
- [12] B. J. Shea, C. Hamel, G. A. Wells, L. M. Bouter, E. Kristjansson, J. Grimshaw, D. A. Henry, M. Boers, "AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews", *Journal of Clinical Epidemiology*, vol. 62, no. 10, pp. 1013-1020, 2009.
DOI: <http://dx.doi.org/10.1016/j.jclinepi.2008.10.009>
- [13] J. H. Kim, A. K. Kim, "A quality assessment of meta-analyses of nursing in South Korea", *Journal of Koran Academy of Nursing*, vol. 43, no. 6, pp. 736-745, 2013.
DOI: <http://dx.doi.org/10.4040/jkan.2013.43.6.736>
- [14] I. S. Shin, E. Y. Park, "Review of the meta-analysis research in special education and related field", *Korean Journal of Physical and Multiple Disability*, vol. 54, no. 4, pp. 157-176, 2011.
- [15] Y. S. Song, M. H. Gang, S. A. Kim, "Methodological quality of meta-analyses on the elderly in Korea", *The Korean Association for Local Government & Administration Studies*, vol. 28, no. 4, pp. 349-370, 2014.
DOI: <http://dx.doi.org/10.18398/kjlgas.2014.28.4.349>
- [16] C. Mikton, A. Butchart, "Child maltreatment prevention: a systematic review of reviews", *Bulletin of the World Health Organization*, vol. 87, no. 5, pp. 353-361, 2009.
DOI: <http://dx.doi.org/10.2471/BLT.08.057075>
- [17] J. Á. Bellón, P. Moreno-Peral, E. Motrico, A. Rodríguez-Morejón, A. Fernández, A. Serrano-Blanco, E. Zabaleta-del-Olmo, S. Conejo-Cerón, "Effectiveness of psychological and/or educational interventions to prevent the onset of episodes of depression: a systematic review of systematic reviews and meta-analyses", *Preventive Medicine*, vol. 76, pp. 522-532, 2015.
DOI: <http://dx.doi.org/10.1016/j.ypmed.2014.11.003>
- [18] S. Y. Kim, J. E. Park, H. J. Seo, H. S. Seo, H. J. Sohn, C. M. Shin, Y. J. Lee, B. H. Jang, D. S. Heo, "NECA's guidance for undertaking systematic reviews and meta-analyses for intervention", *National Evidence-based Healthcare Collaborating Agency*, pp. 1-287, 2011.
- [19] S. D. Lee, S. J. Kim, D. M. Yeum, "Discussion and proposal for metal-analysis papers of social welfare field", *Journal of Korea Contents Association*, vol. 14, no. 1, pp. 209-222, 2014.
DOI: <http://dx.doi.org/10.5392/JKCA.2014.14.01.209>
- [20] B. J. Shea, J. M. Grimshaw, G. A. Wells, M. Boers, N. Andersson, C. Hamel, A. C. Porter, P. Tugwell, D. Moher, L. M. Bouter, "Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews", *BioMed Central Medical Research Methodology*, pp. 7-10, 2007.
DOI: <http://dx.doi.org/10.1186/1471-2288-7-10>
- [21] M. O. Suh, "The review of meta-analysis and research methodology proposed", *Korean Journal of Educational Research*, vol. 49, no. 2, pp. 1-23, 2011.

조미경(Mi-Kyoung Cho)

[정회원]



- 2000년 8월 : 서울대학교 간호학과 (간호학석사)
- 2008년 8월 : 서울대학교 간호학과 (간호학박사)
- 1994년 5월 ~ 2009년 2월 : 서울대학교병원 간호사
- 2009년 3월 ~ 2011년 2월 : 울산대학교 의과대학 임상전문간호학 전공 교수
- 2011년 3월 ~ 현재 : 을지대학교 간호학과 교수

<관심분야>

혈액투석, 자가간호, 인터넷일탈, 노인, 기초간호과학

김 희 영(Hee-Young Kim)

[정회원]



- 2008년 5월 : Louisiana State University, School of Social Work (M.S.W)
- 2002년 5월 : Louisiana State University, School of Social Work (Ph.D.)
- 2014년 3월 ~ 현재 : 남서울대학교 아동복지학과 교수

<관심분야>

아동청소년복지, 행위중독, 청소년비행