

간호연구 질 평가 도구 및 보고지침 고찰

김경희¹ · 김주현² · 임경춘³ · 이경숙⁴ · 정재심⁵ · 최명애⁶ · 채영란²

¹중앙대학교 간호대학 교수, ²강원대학교 의과대학 간호학과 교수, ³성신여자대학교 간호대학 부교수, ⁴강릉원주대학교 원주캠퍼스 간호학과 교수, ⁵울산대학교 임상대학원 부교수, ⁶서울대학교 간호대학 교수

Quality Assessment Tools and Reporting Standards in Nursing Research

Kyunghee Kim¹, Joo Hyun Kim², Kyung-Choon Lim³, Kyung-Sook Lee⁴, Jae-Sim Jeong⁵, Myoung-Ae Choe⁶, Young Ran Chae²

¹Professor, Chunang University College of Nursing, Seoul; ²Professor, Department of Nursing, Kangwon National University College of Medicine, Chuncheon; ³Associate Professor, Shungshin Women's University College of Nursing, Seoul; ⁴Professor, Department of Nursing, Gangneung Wonju National University Wonju Campus, Wonju; ⁵Associate Professor, Department of Nursing, University of Ulsan, Ulsan; ⁶Professor, Seoul National University College of Nursing, Seoul, Korea

Purpose: Quality of nursing research should be evaluated before it is applied as an evidence for evidence-based nursing practice. This study attempted to analyze and to compare tools for the quality assessment and reporting standards of nursing research using CONSORT and STROBE checklist by types of research design. **Methods:** We searched the tools for quality assessment in nursing research based on the National Evidence-based Healthcare Collaborating Agency (NECA) publication. Then, we analysed and compared the tools for quality evaluation by types of research design. **Results:** According to the analysis using CONSORT checklist, ROB shows coherence in 17 items, Jadad shows coherence in 3 items, SIGN (for RCT) shows coherence in 26 items, and Downs & Black shows coherence in 24 items. According to the analysis using STROBE checklist, MINORS shows coherence in 25 items, NOS shows coherence in 21 items, SIGN (for Cohort studies & Case-control studies) shows coherence in 29 items, and RoBANS shows coherence in 21 items. **Conclusion:** Based on our analysis, we recommend that nursing researchers should report according to the reporting standards of tools for quality evaluation. We hope that our analysis can be helpful to develop evidence-based nursing.

Key Words: Quality assessment; Reporting guideline; Nursing research; Evidence-based nursing

국문주요어: 질 평가, 보고지침, 간호연구, 근거기반 실무

서 론

1. 연구의 필요성

과학의 발전은 검증을 거친 이론을 실제에 적용하는 응용과정을 통해 지식체를 확대시킴으로써 가능해진다. 자연과학의 한 분야인 간호학에서도 간호 지식체를 개발하고 확대시키기 위해 이론 또는 실무중심의 다양한 간호연구가 이루어지고 있다(Lee, Park, Cho,

& Park, 2011). 간호실무 현장에서 임상적 결정을 내릴 때 최상의 근거를 활용하고자 하는 근거기반실무에 대한 관심이 고조되고 있다(Sackett, Straus, Richardson, Rosenberg, & Haynes, 2000). 국내에서는 2004년 한국간호과학회 주관으로 개최된 추계학술대회에서 근거기반간호(evidence based nursing)가 소개되면서, 간호학 분야별 학회 및 간호사회 등 간호계의 전문가 단체를 중심으로 근거기반간호 활성화를 위한 노력이 지속적으로 전개되고 있다(Gu et al., 2012). 간호실무 현장에서 근거기반실무가 이루어지기 위해서는 활용 가능한 근거 생성이 우선되어야 하는데, 이러한 근거는 간호연구를 통해서 이루어진다. 즉, 간호사가 대상자에게 근거기반간호를 제공하기 위해서 활용하는 최상의 근거는 과학적인 연구방법론을 통해 얻어진 결과 중에서 가장 타당하고 신뢰할 수 있는 결과를 의미한다.

이와 같이 근거는 간호연구를 통해 생성되고, 연구의 질적인 수준에 따라 근거의 등급이 나누어지므로(Park, 2006; Sackett, Straus,

Corresponding author:

Joo Hyun Kim, Professor, Department of Nursing, Kangwon National University College of Medicine, Gangwondaehakkyo-gil 1, Chuncheon 200-701, Korea
Tel: +82-33-250-8882 Fax: +82-33-242-8840 E-mail: joohkim@kangwon.ac.kr

투고일: 2012년 11월 5일 심사완료일: 2012년 11월 6일 게재확정일: 2012년 11월 19일

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Richardson, Rosenberg, & Haynes, 1997), 근거의 수준을 평가하는 작업은 매우 중요하다. 근거수준의 평가는 연구설계에 따라 타당도가 검증된 도구를 사용하여 질 평가를 실시하고 이에 따라 근거수준을 등급화하여 여러 중재 결과의 질평가를 통합하여 전체 근거수준을 결정하게 된다(Kim et al., 2011). 이러한 과정으로 근거의 수준이 결정되면 임상현장 적용을 원하는 정도인 권고의 강도를 결정하게 된다(Atkins et al., 2004; Kim et al., 2011).

그러므로 근거를 찾아내는 전체 과정에서 연구의 질 평가는 매우 기본적이고 또한 중요한 과정이며, 연구의 질을 올바르게 평가할 수 있는 도구의 선정 또한 중요하다. 연구의 질 평가를 위한 도구는 연구설계에 따라 개발되었다. 무작위대조연구(randomized controlled trial, RCT)에 사용되는 연구 질 평가도구는 Cochrane's assessment of Risk of Bias (ROB), Jadad scale, Scottish Intercollegiate Guidelines Network (SIGN-RCT) 및 Downs & Black 도구 등이 타당도가 검증된 척도 혹은 체크리스트로 국제적으로 많이 활용되고 있다(Kim et al., 2011; Lim et al., 2011). 또한 비무작위연구(non randomized study, NRS)에 사용되는 근거의 질 평가도구는 Methodological Index of Non-Randomized Studies (MINORS), Newcastle & Ottawa Scale (NOS), SIGN 및 Risk of Bias Assessment tool for Non-randomized Study (RoBANS) 등이 타당도가 검증된 척도 혹은 체크리스트로 국제적으로 많이 활용되고 있다(Kim et al., 2011; Lim et al., 2011).

이러한 연구의 질 평가 도구는 특히 방법론적 질 평가에 초점을 두며 근거의 수준은 연구의 방법론적 질을 평가함으로써 결정된다. 이와 더불어 질 높은 연구수행임을 충분히 잘 나타내는 연구보고 또한 중요한데, 이것은 연구 보고의 질을 나타내는 부분으로 연구보고에 따라 질 높은 연구를 수행하였음을 평가받을 수 있기 때문이다. 이에 여러 연구그룹에서는 연구보고의 질을 향상시키기 위한 목적으로 연구결과와 보고에 필수적으로 포함되어야 할 문항들을 정리하여 연구 설계별로 표준화된 연구 보고지침을 개발하였다. 무작위 대조연구의 보고지침은 Consolidated Standards of Reporting Trials (CONSORT), 비무작위연구의 보고지침은 Transparent Reporting of Evaluations with Nonrandomized Designs (TREND), 관찰 연구의 보고지침은 Strengthening the Reporting of Observational Studies in Epidemiology (STROBE), 체계적 고찰의 보고지침은 Quality of Reporting of Meta-analysis (QUOROM), 관찰연구 메타분석의 보고지침은 Meta-analysis of Observational Studies in Epidemiology (MOOSE), 진단연구의 보고지침은 Standards for the Reporting of Diagnostic Accuracy Studies (STARD)이다(Lee, 2009).

최근 국내에서 간호연구가 수적으로 매우 증가되고 있고 또한 근거기반간호에 대한 관심이 증가하고 있으므로 근거수준을 향상시

킬 수 있는 전략이 요구된다. 이를 위해 연구의 질 평가 도구와 연구의 보고지침을 잘 알고 연구수행의 모든 단계에서 이를 활용하는 것이 중요하다. 객관적인 지표가 되는 생물학적 측정을 상대적으로 가장 많이 활용하고 있는(Kim et al., 2006) 기초간호자연과학회지의 연구들은 상대적으로 근거로 활용 가능성이 높다고 하겠다. 그러나 이런 연구들이 근거에 활용되려면 높은 등급의 평가를 받아야 하고 이를 위해서는 연구의 질 평가 항목들을 숙지하고 연구 설계부터 적용하려고 노력해야 할 것이다. 이에 본 연구는 무작위 연구와 비무작위 연구의 질 평가 도구와 보고지침을 고찰하고, 보고지침을 기준으로 연구 질 평가도구의 구성요소를 비교하고자 하였다. 본 연구는 근거의 질 평가 도구 및 보고지침에 대한 고찰을 통해 추후 간호연구 수행 및 보고를 위한 방향을 제시함으로써, 간호실무 현장에서 실제로 활용 가능한 최상의 근거를 생성하는 간호연구수행에 도움이 될 것이다.

2. 연구 목적

본 연구는 연구설계에 따른 질평가 도구와 보고지침을 살펴보고 보고지침을 이용하여 질 평가 도구를 비교하고자 시도되었으며, 구체적인 연구목적은 다음과 같다.

- 1) 무작위 연구와 비무작위 연구의 질평가 도구를 고찰한다.
- 2) 무작위 연구와 비무작위 연구의 연구보고지침을 고찰한다.
- 3) 연구보고지침을 이용하여 연구의 질 평가도구를 비교한다.

연구 방법

본 연구는 무작위 연구와 비무작위 연구의 질을 평가하는 도구들과 연구보고지침에 대한 전반적인 검토 및 비교를 통해 근거기반 실무의 기초가 될 간호연구의 질을 평가하는데 가장 적절한 도구를 선정 및 실제로 적용하기 위한 고찰 연구이다.

무작위 연구와 비무작위 연구의 질을 평가하는 도구들과 연구보고지침에 대해 한국보건 의료연구원(NECA)에서 공식적으로 발표한 내용들을 기본으로 하여 문헌고찰을 통해 본 연구의 분석에 포함하였다. 간호학 연구에 활용가능한 질 평가 도구들의 원래 도구를 찾고, 한글판 도구가 있는지를 확인하였으며, 구체적인 평가방법을 고찰하였다. 이 과정을 통해 무작위대조연구(RCT)에 사용되는 연구 질 평가도구인 ROB, Jadad scale, SIGN-RCT 및 Downs & Black 도구 등이 타당도가 검증된 척도 혹은 체크리스트로 국제적으로 많이 활용되고 있음을 확인하였다. 또한 비무작위연구(NRS)에 사용되는 근거의 질 평가도구는 MINORS, NOS, SIGN 및 RoBANS 등이 타당도가 검증된 척도 혹은 체크리스트로 국제적으로 많이 활

용되고 있음을 고찰하였다. 이에 따라 수차례의 연구자 회의를 통해 연구 질 평가 도구들을 검토하고 비교 분석하였다.

질 평가 도구와 마찬가지로의 방법으로 연구보고지침들을 검색하고 검토하였다. 무작위 대조연구를 위한 CONSORT, 비무작위 연구를 위한 TREND, 관찰연구를 위한 STROBE, 체계적 고찰을 위한 QUOROM, 관찰연구 메타분석을 위한 MOOSE, 진단연구를 위한 STARD가 개발되었음을 확인하였고, 이 논문에서는 CONSORT와 STROBE를 분석에 활용하였다.

연구 결과

1. 무작위 연구(RCT)의 질 평가도구

무작위 연구의 질을 평가하기 위하여 국내·외에서 많이 이용되고 있는 도구는 ROB, Jadad 척도, SIGN 및 Downs & Black 도구 등 (Table 1)이며, 이들은 국제적으로 타당도가 검증된 RCT 연구의 질 평가도구이다.

1) ROB

ROB는 체크리스트 방식과 영역평가 방식을 결합한 형태의 도구로 연구의 비뚤림 위험을 평가하는데 초점을 두고 있다. 코크란 연합(cochrane collaboration)에서 제공하는 2011년 개정된 ROB (원도구 출처: www.cochrane-handbook.org)의 평가영역은 6개 영역으로 무작위배정순서 생성(random sequence generation)의 적절성, 무작위배정순서 은폐(allocation concealment)의 적절성, 연구참여자 및 연구자에 대한 눈가림(blinding of participants and personnel)의 적절성, 결과평가에 대한 눈가림(blinding of outcome assessment)의 적절성, 불충분한 결과자료(incomplete outcome data)의 완결성, 선택적 보고(selective reporting)의 위험성 및 기타 비뚤림 유입 가능성이다. 평가방법은 6개 영역별 각 문항에 대하여 비뚤림 위험이 '높음(high risk of bias)', '낮음(low risk of bias)', '불확실(uncertain risk of bias)'로 판정하게 되는데, 판정기준에 대한 구체적인 가이드라인이 제시되어 있다. ROB로 평가한 내용을 해석할 때 연구의 질평가 결과를 하나로 요약하는 것은 일반적으로 추천되지 않고, 한편의 연구에서도 각 영역의 비뚤림 위험이 다를 수 있으므로 영역별로 비뚤림 위험을 요약하는 것이 추천된다. 서로 다른 연구에서 동일한 하나의 결과에 대해 비뚤림 위험을 요약하는 것은 메타분석이나 근거등급 평가에서 주로 이용된다. 비뚤림 위험은 모든 결과들에서 동일하지 않고 임상상황에 따라 다를 수 있으므로 모든 연구의 모든 결과에 대해 비뚤림 위험을 요약하여 해석하지 않도록 한다(Higgins & Green, 2011; Kim et al., 2011).

2) Jadad Scale

Jadad 척도(원도구 출처: www.ncbi.nlm.nih.gov/books/NBK37429)는 Jadad scoring 또는 the Oxford quality scoring system으로도 알려져 있으며 RCT 연구의 질 평가도구이다(Jadad et al., 1996). Jadad 척도의 평가영역은 무작위배정(randomization), 이중눈가림(double-blind), 탈락(withdrawal)의 3개 영역이고, 문항은 총 5개로, 무작위배정에 관한 2문항, 눈가림에 관한 2문항 및 탈락에 관한 1문항으로 구성되어 있다. 무작위배정 영역의 평가방법은 무작위배정을 언급하면 1점, 무작위배정이 적절하면 1점 추가, 잘못된 무작위 배정은 -1점을 부여하여 0-2점의 점수분포를 갖는다. 이중눈가림 영역에 대한 평가는 이중눈가림이 언급된 경우 1점, 이중 눈가림이 적절하면 1점 추가, 잘못된 이중눈가림은 -1점을 부여하여 무작위 배정에 관한 평가와 마찬가지로 점수분포는 0-2점이다. 탈락 영역에 대한 평가는 탈락에 관한 내용이 언급되었을 경우 1점의 점수를 부여한다. 이와 같은 점수를 부여하여 Jadad 도구로 평가한 내용은 총 5점 만점으로 평가한다. 최종 연구의 질 평가는 평가점수가 0-2점 사이는 연구의 질이 낮은 것으로, 3-5점 사이는 연구의 질이 높은 것으로 평가한다(Lim et al., 2011).

3) SIGN for RCT

SIGN (원도구 출처: www.sign.ac.uk/guidelines/fulltext/50/checklist2.html)은 연구설계 유형별로 무작위연구용과 비무작위연구용으로 구분되며, 연구수행의 방법론적인 질을 평가한 후 잠재적인 비뚤림 유입의 가능성에 따라 등급을 매겨서 연구설계와 연구수행의 질을 결합하여 평가한다. 연구수행의 방법론적인 질 평가는 영역별 문항을 체크리스트 방식의 도구로 판정하고 판정기준에 대한 구체적인 가이드라인이 제시되어 있다(Health Insurance Review and Assessment Service & Hallym University Industry Academic Cooperation Foundation, 2009). SIGN-RCT의 평가영역은 내적타당도, 전체 연구의 전반적인 평가, 연구의 서술 등 3개의 영역으로 구성된다. 척도의 문항은 내적타당도에 관한 문항 10개(1.1-1.10), 전체 연구의 전반적인 평가에 관한 문항 4개(2.1-2.4), 연구의 서술에 관한 문항 15개(3.1-3.15)로 총 29개 문항이다(SIGN, 2011). 내적타당도 평가방법은 18문항 '분석에 포함된 집단이나 개인의 비율은 얼마인가?'만 제외하고 각 문항은 잘 수행됨, 적절히 다루어짐, 빈약하게 다루어짐, 다루어지지 않거나 언급이 미흡함, 보고되지 않음, 해당 없음의 6개 범주로 판정하고 판정기준에 대한 구체적인 가이드라인이 있다. 전체 연구의 전반적인 평가의 평가방법은 1++, 1+, 1-로 판정한다. '1++'는 평가기준이 모두 또는 대부분 충족되고 충족되지 않은 기준들이 연구의 결론을 바꿀 가능성은 매우 적다고 판단되는 경우이다. '1+'

Table 1. Summary of the Items and Area of Quality Assessment Tools

Tools	Research design	Areas (No. of items)	No. of items	Method of assessment
ROB	RCT	6 areas (each 1 item) Random sequence generation Allocation concealment Blinding; participants & personnel Blinding; outcome assessment Incomplete outcome data Selective reporting Other bias	7	Risk of bias: high, low, uncertain Revealed all areas' scores
Jadad Scale	RCT	3 areas Randomization (2) Double-blind (2) Withdrawal (1)	5	Random assignment & Double blinding Reported = 1 Appropriately = +1 Not appropriately = -1 Mortality Reported = 1 Total score: 0-2; low, 3-5; high,
Downs & Black	RCT	5 areas Reporting (10) External validity (3) Internal validity-bias (7) Internal validity-confounding (6) Power (1)	27	Reporting 9 items; yes = 1, no = 0, 1 item; yes = 2, partly yes = 1, no = 0 Areas except reporting & power: yes = 1, no = 0, uncertain = 0 power: 0-5 Total score: 0-32 (High score means high quality)
SIGN	RCT	3 areas Internal validity (10) Overall assessment (4) Description (15)	29	Internal validity area - 6 categories (well covered, adequately addressed, poorly addressed, not addressed, not reported, not applicable) Overall assessment - ++, +, -
	Cohort study	3 areas Internal validity (14), Overall assessment (4) Description (15)	33	General description - 3 items (V) - 12 items (description)
	Case-control study	3 areas Internal validity (11), Overall assessment (4) Description (11)	26	
MINORS	With control group study		12	Not reported = 0 Reported, inadequate = 1
	Without control group study		8	Reported, adequate = 2 Total score 0-24 (with control) 0-16 (without control) High score means high quality.
NOS	Case-control study	Selection (4) Comparability (1) Exposure (3)	8	Selection & Exposure/Outcome High quality : '★' Each item can get maximum 1 '★' Comparability (1)
	Cohort study	Selection (4) Comparability (1) Outcome (3)	8	1 item that can be compared can get maximum 2 '★'. Total score : 0-9 High score means high quality.
RoBANS	NRS. Cohort study Case-control study, before-after study	6 areas Selection of participants Confounding variables Measurement of intervention Blinding for outcome assessment Incomplete outcome data Selective outcome reporting	6	Risk of bias: low, high unclear

ROB = Cochrane's assessment of risk of bias; SIGN for RCT = Scottish Intercollegiate Guidelines Network for randomized controlled trial; MINORS = Methodological Index of Non-Randomized Studies; NOS = Newcastle & Ottawa Scale; RoBANS = Risk of Bias Assessment tool for Non-randomized study.

는 일부 평가기준이 충족되고 충족되지 않거나 적절히 기술되지 않은 기준들이 연구의 결론을 바꿀 가능성이 적다고 판단되는 경우이다. 이에 비하여 '1'은 모든 또는 거의 모든 기준이 충족되지 않고 연구의 결론이 바뀔 가능성이 있거나 매우 높다고 판단되는 경우이다. 연구의 서술에 대한 평가방법은 3개의 문항(3.1, 3.3, 3.4)은 해당 사항에 'V'로 표기하고 나머지 12개 문항은 각 문항에서 요구하는 핵심적인 내용을 구체적으로 자세하게 서술하여 빈칸을 채운다. SIGN 도구로 평가한 연구 설계와 질 평가, 전체 근거의 일관성, 임상에 미치는 영향 및 적용가능성에 대한 영역별 평가를 근거로 평가 대상 연구에 대한 근거의 등급과 권고의 등급이 결정된다(SIGN, 2011).

4) Downs & Black

Downs & Black 도구(원도구 출처: www.ncbi.nlm.nih.gov/books/NBK37429)는 역학연구의 기본원칙, 연구 설계의 검토 및 RCT 연구 평가를 위한 기존의 체크리스트 등을 기반으로 개발된 평가도구이다(Downs & Black, 1998). 이 도구의 평가영역은 연구의 전반적 서술(reporting), 외적타당도, 내적타당도-비뿔림, 내적타당도-혼동변수(선택 비뿔림), 검정력의 5개 영역으로 구성되어있다. 영역별 문항은 연구의 전반적 서술에 관한 문항 10개, 외적타당도에 관한 문항 3개, 내적타당도-비뿔림에 관한 문항 7개, 내적타당도-혼동변수에 관한 문항 6개, 검정력에 관한 문항 1개로 총 27개 문항으로 구성되어있다. 평가방법을 살펴보면, 연구의 전반적 서술 영역의 10개 문항 중 9개 문항은 각 문항의 평가기준에 대해 '예=1점', 또는 '아니오=0점'으로 판정하고, 예외적으로 5번 문항은 '예=2점', '부분적=1점', '아니오=0점'으로 판정한다. 외적타당도 영역의 3개 문항, 내적타당도-비뿔림 영역의 7개 문항, 내적타당도-혼동변수(선택 비뿔림)영역의 6개 문항은 각 문항의 평가기준에 대해 '예=1점', '아니오=0점', '모름=0점'으로 판정한다. 검정력 문항은 우연에 의한 차이가 나타날 확률이 5% 미만인 상황에서, X%와 Y%의 차이를 확인하여 산출한 표본수를 근거로 임상적으로 중요한 효과를 규명할 수 있는 통계적 검정력의 정도를 확인하여 0-5점까지 점수를 부여할 수 있다. 따라서 Downs & Black 도구의 점수 범위는 최소 0점에서 최대 32점까지이다. 평가대상 연구의 질 평가는 총점이 높을수록 질적인 논문으로 평가한다(Lim et al., 2011).

2. 비 무작위 연구(NRS)의 질 평가도구

국내 · 외에서 근거 및 문헌의 질 평가도구로 많이 활용하고 있는 MINORS, NOS, SIGN: cohort studies, SIGN: Case-control studies 및 RoBANS 등(Table 1)은 국제적으로 타당도가 검증된 비무작위 연구

의 질 평가도구이다.

1) MINORS

MINORS (원도구 출처: <http://onlinelibrary.wiley.com/doi/10.1046/j.1445-2197.2003.02748.x/pdf>)는 비무작위연구의 질평가를 위하여 개발되었으며, 내적일관성, 외적타당도, 평가자 간 일치도, 검사-재검사법으로 신뢰도와 타당도가 검증된 평가도구이다(Slim et al., 2003). MINORS 도구의 평가항목은 12개의 문항으로 구성되어있으며, 비교군이 있는가, 없는가에 따라 평가 문항이 달라진다. 12개의 문항 중에서 1) 분명한 목적이 있는가, 2) 환자가 연속적으로 포함되었는가, 3) 자료가 전향적으로 수집되었는가, 4) 연구목적에 적절한 결과인가, 5) 연구결과가 비뿔림 없이 평가되었는가, 6) 추적기간은 적절한가, 7) 탈락이 5% 미만인가, 8) 연구크기가 전향적으로 계산되었는가, 9) 8개의 문항은 공통문항이다. 그리고 비교군이 있는 비무작위연구 경우 9) 적절한 대조군이 있는가, 10) 대상자군 모집이 동시에 이루어졌는가, 11) 대상자군의 기저상태가 동질한가, 12) 통계 분석이 적절한가의 4개 문항이 추가된다. 다시 말해 MINORS 도구의 평가방법은 비교군이 없는 비무작위 연구는 1번에서 8번까지 8개의 공통문항만 평가하며, 비교군이 있는 비무작위 연구는 9번에서 12번까지 4개 문항을 추가하여 12문항 모두를 평가한다. 각 문항은 0점에서 2점까지 3점 척도이다. '0점=연구에서 보고되지 않음', '1점=보고되었으나 부적절함', '2점=적절하게 보고됨'으로 판정한다. 따라서 비교군이 없는 비무작위 연구의 점수범위는 최소 0점에서 최대 16점이고 비교군이 있는 비무작위 비교연구의 점수범위는 최소 0점에서 최대 24점이다. 평가대상 연구의 질 평가는 총점이 높을수록 질적인 논문으로 평가한다(Kim & Kim, 2009).

2) NOS

NOS 도구(원도구 출처: www.ohri.ca/programs/clinical_epidemiology/nosgen.pdf)는 비무작위 관찰연구의 질 평가를 위해서 개발된 도구로, 환자-대조군 연구(case-control studies)의 질 평가도구와 코호트 연구(cohort studies)의 질 평가도구를 구분하여 개발하였다(Wells et al., 2009). 환자-대조군 연구의 질 평가영역은 환자-대조군 선택, 환자-대조군 비교가능성(comparability) 및 노출확인(exposure)의 3개 영역으로 구성되어있다. 영역별로 선택에 관한 문항 4개, 비교가능성에 관한 문항 1개, 노출확인에 관한 문항 3개 등 총 8개의 문항이다. 코호트 연구의 질 평가 영역은 코호트 선택, 코호트 비교가능성(comparability) 및 결과평가(outcome) 등 3개 영역으로 구성되어있다. 영역별로 선택에 관한 문항 4개, 비교가능성에 관한 문항 1개, 결과평가에 관한 문항 3개 등 총 8개의 문항이다. 평가방법

은 각 문항에 대해 근거의 질이 높은 경우에 '★'로 표시하도록 한다. 선택영역의 4개 문항은 각 문항에 최대 한 개의 '★'을 줄 수 있고, 비교가능성 영역의 1개 문항은 최대 두 개의 '★'을 줄 수 있다. 노출확인 및 결과평가 영역의 3개 문항은 최대 한 개의 '★'을 줄 수 있다. 세 영역의 총 8개 문항에서 '★'의 점수범위는 최소 0점에서 최대 9점까지이다. 평가대상 연구의 질 평가는 '★'의 수가 많을수록 질적인 논문으로 평가한다(Lim et al., 2011; Wells et al., 2009).

3) SIGN

SIGN은 비무작위연구 중 코호트 연구(원도구 출처: www.sign.ac.uk/guidelines/fulltext/50/checklist3.html)와 환자-대조군 연구(원도구 출처: www.sign.ac.uk/guidelines/fulltext/50/checklist4.html)의 질 평가를 위한 연구이다. 평가영역은 무작위연구에서와 마찬가지로 내적타당도, 전체 연구의 전반적인 평가, 연구의 서술 등 3개의 영역으로 구성된다.

코호트 연구의 영역별 문항은 총 33개로, 내적타당도 영역은 연구질문 기술의 명확성과 적절성에 관한 문항 1개(1.1), 대상자 선택에 관한 문항 5개(1.2-1.6), 자료수집에 관한 문항 6개(1.7-1.12), 혼동변수에 관한 문항 1개(1.13), 통계분석에 관한 문항 1개(1.14) 등 14개 문항이다. 전체 연구의 전반적인 평가 영역에 관한 문항은 4개(2.1-2.4)이고, 연구의 서술 영역에 관한 문항은 15개(3.1-3.15)이다(SIGN, 2011). 이에 비하여 환자-대조군 연구의 문항은 내적타당도 영역은 연구질문 기술의 명확성과 적절성에 관한 문항 1개(1.1), 대상자 선택에 관한 문항 6개(1.2-1.7), 자료수집에 관한 문항 2개(1.8-1.9), 혼동변수에 관한 문항 1개(1.10), 통계분석에 관한 문항 1개(1.11) 등 11개 문항이다. 전체 연구의 전반적인 평가 영역에 관한 문항은 4개(2.1-2.4)이며, 연구의 서술 영역에 관한 문항은 11개(3.1-3.11)로 총 26개 문항이다(SIGN, 2011).

내적타당도 영역 평가방법은 각 문항별로 잘 수행됨, 적절히 다루어짐, 빈약하게 다루어짐, 다루어지지 않음, 보고되지 않음, 해당 없음의 6개 범주로 판정한다. 이렇게 평가하지 않는 예외 문항이 있는데, 코호트 연구에서는 문항1.5 '분석에 포함된 집단이나 개인의 비율은 얼마인가?'와 문항1.14 신뢰구간이 제시되었는가?의 2개 문항을 제외하고, 환자-대조군 연구에서는 문항1.4 '연구에 참여한 환자군과 대조군의 비율은 얼마인가?'와 문항1.11 신뢰구간이 제시되었는가? 등 2개 문항을 제외한다. 전체 연구의 전반적인 평가 영역의 평가방법은 2++, 2+, 2-로 판정하는 것으로 코호트연구와 환자-대조군 연구에서 동일하다. ++, +, - 평가기준은 무작위연구에서 나타난 것과 같으며, 단지 무작위연구는 부호 앞에 숫자를 1로, 비무작위연구는 2로 표기한다. 연구의 서술영역에 대한 평가방법 또한 코호트

연구와 환자-대조군 연구에서 동일하며, 무작위 연구에서 나타난 것과 같은 방법으로 평가한다.

4) RoBANS

RoBANS (원도구 출처: www.neca.re.kr/kctic/resource/)는 비무작위 연구의 질 평가를 위해 2009년 개발된 도구로 타당도가 검증되었으며(Park et al., 2011), 체크리스트 방식과 영역평가 방식이 결합된 형태이다. 비무작위배정 비교임상시험연구, 코호트 연구, 환자-대조군 연구, 전후연구의 질 평가에 활용된다. RoBANS는 대상군 선정, 혼동변수, 중재[노출]측정, 결과평가 눈가림, 불완전한 결과자료, 선택적 보고 등 6개 영역의 비뚤림 위험을 평가할 수 있도록 구성되어 있다. 각 영역에서는 코호트 연구, 비교임상시험연구, 환자-대조군 연구, 전후연구로 구분하여 해당 영역의 비뚤림 위험을 판정하는 기준을 제시하고 있다. 평가방법은 영역별로 비뚤림 위험 '낮음', '높음', '불확실'의 세 가지로 판정한다. 비뚤림 위험 '낮음'은 해당 영역의 비뚤림 가능성이 낮음을 의미하며, 비뚤림 위험 '높음'은 해당 영역의 비뚤림 가능성이 높음을 의미한다. 비뚤림 위험 '불확실'은 비뚤림에 대한 위험을 판단하기 어려운 경우이다. RoBANS 평가도구의 해석방법은 연구 내 모든 결과에 대해 하나의 비뚤림 위험으로 평가하고, 한편의 연구에서는 결과에 따라 비뚤림 위험을 평가하며, 서로 다른 연구에서 동일한 하나의 결과에 대해 비뚤림 위험을 평가하고, 모든 연구의 모든 결과에 대해 비뚤림 위험을 평가하는 등, 네 가지 방식으로 해석한다(Kim et al., 2011).

3. 무작위 연구의 보고지침과 보고지침을 이용한 질 평가도구 비교

본 연구에서는 위에서 살펴본 RCT 연구의 질 평가도구인 ROB, Jadad, SIGN-RCT 및 Downs & Black을 연구의 보고지침을 기준으로 비교하고자 한다.

1996년 CONSORT 그룹에 의해 개발된 CONSORT 체크리스트(원도구 출처: www.consort-statement.org/consort-statement/)는 무작위 대조연구의 보고지침이다(Begg et al., 1996). 2001년 CONSORT 그룹은 CONSORT 체크리스트의 양식과 문항을 부분적으로 수정하여 1차 개정하였고(Moher, Schulz, & Altman, 2001), 2010년에는 CONSORT 체크리스트 문항을 수정하면서 추가적인 문항을 개발하여 2차 개정하였다(Schulz, Altman, & Moher, 2010). 2010년 개정된 CONSORT 체크리스트는 제목과 초록, 서론, 연구방법, 연구결과, 논의 및 기타 정보의 6개 영역으로 구성되어 있다. 각 영역별 문항을 살펴보면 제목과 초록영역 2문항[제목 1문항(문항번호-1a), 초록 1문항(1b)], 서론영역 2문항[배경 1문항(2a), 목적 1문항(2b)], 연구방법영역 17문항[연구설계 2문항(3a, 3b), 대상자 2문항(4a, 4b), 중재 1문항

(5), 결과 2문항(6a, 6b), 대상자 수 2문항(7a, 7b), 배정순서 2문항(8a, 8b), 배정은폐 1문항(9), 배정실행 1문항(10), 눈가림 2문항(11a, 11b), 통계방법 2문항(12a, 12b)이다. 연구결과영역은 10문항으로 참가자 흐름도 2문항(13a, 13b), 모집 2문항(14a, 14b), 연구시점 자료 1문항(15), 분석된 대상자 수 1문항(16), 결과와 추정치 2문항(17a, 17b), 부가적 분석 1문항(18), 부작용 1문항(19)이다. 논의영역은 3문항으로 제한점 1문항(20), 일반화 가능성 1문항(21), 해석 1문항(22)이다. 기타 정보 영역은 3문항으로 등록 1문항(23), 프로토콜 1문항(24), 연구비 지원 1문항(25)이다. 따라서 2010년 개정된 CONSORT는 총 37문항으로 구성되어 있으며 무작위대조연구에 기술되어야 하는 지침을 제목과 초록, 서론, 연구방법, 연구결과, 논의 및 기타 정보의 순서로 제시하고 있으며(Schulz, Altman, & Moher, 2010), 웹사이트(<http://www.consort-statement.org/consort-statement/>)를 통해 CONSORT체크리스트 파일과 체크리스트 문항의 평가방법과 기준 및 근거에 대한 상세한 지침을 제공하고 있다.

본 연구에서 RCT 연구의 질평가도구인 ROB, Jadad, SIGN:RCT 및 Downs & Black의 구성항목을 CONSORT 체크리스트의 평가기준 및 세부항목들과 비교해 본 결과는 Table 1과 같다. CONSORT 체크리스트 총 37문항 중 ROB 평가도구의 문항은 17개 문항(6a, 8a, 9, 11a, 12b, 13a, 13b, 14a, 14b, 15, 16, 17a, 17b, 18, 19, 20, 24)이 부합하였다. Jadad 평가도구의 문항은 3개(8a, 11a, 13b), SIGN-RCT 평가도구의 문항은 26개(1b, 2b, 3a, 3b, 4a, 4b, 5, 6a, 8a, 8b, 9, 11a, 11b, 12a, 12b, 13a, 13b, 14a, 15, 16, 17a, 18, 19, 21, 24, 25), Downs & Black 평가도구의 문항은 24개(2b, 3a, 4a, 4b, 5, 6a, 7a, 8a, 9, 11a, 11b, 12a, 12b, 13a, 13b, 14a, 15, 16, 17a, 17b, 18, 19, 23, 24) 문항이 부합하였다(Table 1).

따라서 SIGN-RCT 평가도구가 CONSORT 체크리스트 보고지침과 가장 잘 부합되고, Jadad 평가도구의 부합도가 가장 낮은 것으로 나타났다. CONSORT 체크리스트의 문항 중제목과 초록영역에서 연구설계, 연구방법, 연구결과 및 결론에 대한 구조화된 요약에 관한 문항(1b), 서론영역에서 과학적 배경과 연구의 필요성에 관한 문항(2a), 연구방법영역에서 연구진행 중 변경된 것과 그 이유에 관한 문항(6b), 적용이 가능한 경우 중간평가와 연구중단 지침에 대한 설명에 관한 문항(7b), 배정순서 생성자, 참여자 모집자, 중재 배정자에 관한 문항(10), 논의영역에서 연구결과에 대한 일관성 있는 해석, 이로운 점과 해로운 점의 균형 및 또 다른 관련근거에 대한 고려에 관한 문항(22)등 6개 문항은 ROB, Jadad, SIGN:RCT 및 Downs & Black 평가도구와 부합되지 않았다(Table 2).

4. 비무작위 연구의 보고지침과 보고지침을 이용한 질 평가도구 비교

2003년 STROBE 그룹에 의해 개발된 STROBE 체크리스트는 코

Table 2. CONSORT 2010 Checklist of Information to Include When Reporting a Randomized Trial

Section/Topic	Item No	Cochrane	Jadad	SIGN	Downs & Black
Title and abstract	1a 1b			v	
Introduction					
Background	2a			v1.1	v 1
objectives	2b				
Methods					
Trial design	3a 3b			v3.1 v1.9	v 23
Participants	4a 4b			v1.5 v1.5	v 3 v 11
Interventions	5			v3.3	v 4, 19
Outcomes	6a 6b	v 6		v3.6	v 2
Sample size	7a 7b				v 27
Randomization:					
Sequence generation	8a 8b	v 1	v 1	v1.2 v3.9, 3.13	v 11, 23
Allocation concealment mechanism	9	v 2		v1.3	v 24
Implementation	10				
Blinding	11a 11b	v 3, 4	v 2	v 1.4, 3.9 v1.6	v 14, 15 v 19
Statistical methods	12a 12b			v1.6, 3.13, 3.14 v3.14	v 18 v5, 17, 25
Results					
Participant flow (a diagram is strongly recommended)	13a 13b	v 5 v 5	v 3	v3.13 v1.8, 1.9, 3.13	v 18, 27 v 9, 26
Recruitment	14a 14b	v 7 v 7		v 3.10, 3.11	v 17, 22
Baseline data	15	v 7		v3.12	v 5, 21
Numbers analysed	16	v 5		v1.9, 3.13, 3.14	v 18, 27
Outcomes and estimation	17a 17b	v 5 v 5		v 3.13, 3.14	v 7, 20, 27 v 27
Ancillary analyses	18	v 5, 6		v 3.13, 3.14	v 5, 16, 17, 25
Harms	19	v 6		v3.7	v 8
Discussion					
Limitations	20	v 7			
Generalisability	21			v2.3	
Interpretation	22				
Other information					
Registration	23				v 11
Protocol	24	v 6		v1.10	v 13
Funding	25			v3.1	

호트 연구, 환자-대조군 연구, 단면연구 등과 같은 관찰연구의 보고 지침이다. 2007년 STROBE 체크리스트를 개정하여 제4판을 발행하였으며(von Elm et al., 2008), 2007년 개정된 STROBE 체크리스트는 제목과 초록, 서론, 연구방법, 연구결과, 논의 및 기타 정보 등 6개 영역으로 구성되어 있다. 각 영역별 문항은 제목과 초록영역에 2문항

으로 제목 1문항(문항번호-1a), 초록 1문항(1b)이고, 서론영역은 2문항으로 배경 1문항(2), 목적 1문항(3)이다. 연구방법영역은 14문항으로 연구설계 1문항(4), 연구환경 1문항(5), 대상자 2문항(6a, 6b), 변수 1문항(7), 자료원/측정 1문항(8), 비뿔됨 1문항(9), 표본수 1문항(10), 양적변수 1문항(11), 통계적 방법 5문항(12a, 12b, 12c, 12d, 12e)이다. 연구결과영역은 11문항으로 대상자 3문항(13a, 13b, 13c), 기술자료 3문항(14a, 14b, 14c), 결과 자료 1문항(15), 주요결과 3문항(16a, 16b, 16c), 추가 분석 1문항(17)이다. 논의영역은 4문항으로 주요결과 논의 1문항(18), 제한점 1문항(19), 해석 1문항(20), 일반화 가능성 1문항(21)이다. 기타 정보 영역은 1문항으로, 연구비 지원에 관한 1문항(22)이다. STROBE 체크리스트의 23개 문항(1a, 1b, 2, 3, 4, 5, 7, 8, 9, 10, 11, 12a, 12b, 12c, 12e, 16a, 16b, 16c, 17, 18, 19, 20, 21)은 연구설계와 상관없이 공통적으로 적용되는 문항이고, 4개 문항(6a, 6b, 12d, 15)은 연구설계에 따라 문항별 평가기준에 차이가 있다. 그리고 7개 문항(8, 13a, 13b, 13c, 14a, 14b, 14c)은 환자-대조군 연구에서는 환자군과 대조군에 각각 따로 적용되고, 코호트 연구에서는 노출군과 대조군에 각각 따로 적용되는 문항이 포함되어 있다. STROBE 체크리스트는 총 34문항으로 구성되며 코호트 연구, 환자-대조군 연구, 단면 연구 등과 같은 관찰연구에 기술되어야 하는 지침을 제목과 초록, 서론, 방법, 결과, 고찰 및 다른 정보의 순으로 제시하고 있다(Lee, 2009; von Elm et al., 2008). STROBE 체크리스트도 웹사이트(<http://www.strobe-statement.org/>)를 통해 체크리스트 파일과 평가방법, 기준 및 근거에 대한 상세한 지침을 제공하고 있다.

본 연구에서 NRS의 질 평가도구인 MINORS, NOS, SIGN, RoBANS의 구성항목을 NRS의 보고지침으로 활용되는 STROBE 체크리스트의 평가기준 및 세부문항들과 비교해 본 결과는 Table 2와 같다. STROBE 체크리스트 총 34문항 중 MINORS의 문항은 25개 문항(1a, 2, 3, 4, 5, 6a, 6b, 7, 8, 9, 10, 12a, 12c, 12d, 13a, 13b, 13c, 14a, 14b, 14c, 15, 16a, 16c, 18, 19)이 부합하였다. NOS의 문항은 21개(1a, 4, 5, 6a, 6b, 7, 8, 9, 11, 12a, 12d, 13a, 13b, 13c, 14a, 14b, 14c, 15, 16a, 16b, 16c), SIGN 문항은 29개(1a, 3, 4, 5, 6a, 6b, 7, 8, 9, 10, 12a, 12b, 12c, 12d, 13a, 13b, 13c, 14a, 14b, 14c, 15, 16a, 16b, 16c, 18, 19, 20, 21, 22), RoBANS의 문항은 21개(1a, 4, 5, 6a, 6b, 7, 8, 9, 10, 12a, 12c, 12d, 13a, 13b, 13c, 14a, 14b, 14c, 15, 16a, 17) 문항이 부합하였다(Table 2).

따라서 SIGN이 STROBE 체크리스트 보고지침과 가장 잘 부합되는 것으로 나타났다. STROBE 체크리스트의 문항 중에서 제목과 초록영역의 연구목적, 방법 및 연구결과를 요약한 초록에 관한 문항(1b), 연구방법영역에서 민감도 분석에 관한 문항(12e) 등 2개 문항은 MINORS, NOS, SIGN, RoBANS 평가도구 모두와 부합되지 않았다(Table 3).

Table 3. STROBE Statement - Checklist of Items That Should Be Included in Reports of Observational Studies

Section/Topic	Item No	MINORS	NOS	SIGN	RoBANS
Title and abstract	1a	v	v	v	v
	1b				
Introduction	Background	2	v 1		
	objectives	3	v 1	v1.1	
Methods					
Study design	4	v	v	v	v
Setting	5	v 2,6,10	v 1-1,1-2,1-3	v 3.2,3.3,3.4, 3.10	v 1,3
Participants	6a	v 2,6,10,11	v 1-2,1-4	v1.2,3.5	v 1
	6b	v 2,9,11	v 1-3,1-4	v1.2,3.5	v 1
Variable	7	v 4,8	v 1-2,2-1	v1.7,3.8	v 2,6
Data sources/ measurement	8	v 3	v 1-1,1-2,1-32-1	v3.12, 3.13	v 1,3
Bias	9	v 5	v	v1.9,2.1	v 2
Study size	10	v 8		v1.3	v 5
Quantitative variables	11		v 2-1		
Statistical methods	12a	v 11	v 2-1	v1.13,2.1	v 2
	12b			v2.1	
	12c	v 7		v2.1	v 5
	12d	v 7	v 2-1	v2.1	v 2
	12e				
Results					
Participants	13a	v 4,6,7,10	v 3-2	v 1.5,3.10,3.11	v 5
	13b	v 4	v 3-3	v1.6	v 5
	13c	v 4	v 3-3	v2.3	v 5
Descriptive data	14a	v 11	v 2-1	v3.4,3.12	v 2,5
	14b	v 4,7	v 3-3	v3.13	v 5
	14c	v 4	v 3-3	v3.11	v 5
	14c	v 4	v 3-3	v3.11	v 5
Outcome data	15	v 4,8	v 3-1,3-3	v 1.8,1.10,2.1	v 5,6
Main results	16a	v 6,8,11,12	v 2-1	v 1.13,1.14,2.1	v 2
	16b		v 2-1	v2.1	
	16c	v 12	v 2-1	v2.1	
	16c	v 12	v 2-1	v2.1	
Other analyses	17				v 6
Discussion					
Key results	18	v 4		v2.4	
Limitations	19	v 5		v1.9	
Interpretation	20			v1.11	
Generalisability	21			v2.3	
Other information					
Funding	22			v3.1	

논 의

간호실무현장에서 근거기반실무가 활성화되기 위해서는 우선적으로 연구를 통한 근거가 생성되어야 한다. 근거기반실무는 최신이면서 최상의 연구결과에서 나온 임상지식에 대한 근거의 질, 근거의

등급 및 권고의 강도를 결정하는 평가과정을 거치므로, 실무현장에서 적용하고 있는 연구 활용(research utilization)과는 차별화된다(Park, 2006).

근거의 평가과정은 먼저, 연구설계 유형에 따라서 표준화된 도구를 사용하여 연구결과로 생성된 근거의 질을 평가한다. 연구결과와 타당성을 확보하여 근거의 질이 보장되기 위해서는 연구목적, 연구설계, 연구대상의 선정 및 수행과정, 분석대상, 분석방법, 분석결과 및 결과에 대한 해석 등이 정확하게 보고되어야 한다. 연구결과와 정확한 보고로 타당성을 확보하기 위해 여러 연구그룹에서는 연구설계 유형별로 연구결과에 필수적으로 포함되어야 하는 항목들을 정리하여 연구결과 보고지침을 개발하였다(Lee, 2009).

이러한 연구설계 유형별 연구결과 보고지침의 개발은 연구의 질을 향상시키는데 기여하고 있는 것으로 보고되면서(Moher, Jones, & Lepage, 2001), 국내에서도 의학 학술지를 중심으로 연구유형별 보고지침에 따라 연구결과를 보고할 것을 학술지 투고규정에 추가하고 있다(Lee, 2009). 이에 근거의 질이 보장된 연구결과를 보고함으로써 간호실무 현장에서 실제적으로 활용 가능한 근거를 생성하는데 기여하기 위해서는 적용할 근거를 찾기 위해 그에 맞는 좋은 연구 질 평가 도구를 적용하는 것이 중요하기에 본 연구에서는 무작위 연구와 비 무작위 연구에 적용할 수 있는 연구설계 유형별 질 평가 도구들에는 어떠한 것들이 있으며 어떤 항목으로 평가하고 있는지를 고찰해보고 각 도구의 항목들을 보고 지침의 구성 항목에 따라 비교 분석하였다. 분석결과, ROB, Jadad, SIGN:RCT 및 Downs & Black 도구 등 RCT 연구의 질 평가도구의 구성항목을 RCT 연구의 보고지침으로 활용되는 CONSORT 체크리스트의 평가기준 및 세부항목들과 비교해 본 결과, CONSORT 체크리스트 총 37문항 중 ROB 평가도구의 문항은 17개 문항이 부합하였다. Jadad 평가도구의 문항은 3개, SIGN:RCT 평가도구의 문항은 26개, Downs & Black 평가도구의 문항은 24개 문항이 부합하였다. 따라서 SIGN:RCT 평가도구가 CONSORT 체크리스트 보고지침과 가장 잘 부합되고, 분석대상 RCT 연구의 질 평가도구는 CONSORT 체크리스트의 문항과 잘 부합되었으나, Jadad 평가도구의 부합도가 가장 낮은 것으로 나타났다. Jadad 척도는 평가문항이 간단하여 국내·외에서 그동안 사용되어 왔지만 RCT 연구의 질 평가의 중요한 항목인 무작위배정순서의 할당 은폐에 대한 문항이 없어서 코크란 연합에서는 권장하지 않은 평가도구로 보고되고 있다(Lim, 2011).

MINORS, NOS, SIGN:Cohort studies, SIGN:Case-control studies, RoBANS 도구 등 NRS 연구의 질 평가도구의 구성항목을 NRS 연구의 보고지침으로 활용되는 STROBE 체크리스트의 평가기준 및 세부항목들과 비교해 본 결과, STROBE 체크리스트 총 34문항 중 MI-

NORS 평가도구의 문항은 25개 문항이 부합하였다. NOS 평가도구의 문항은 21개, SIGN:Cohort studies 및 Case-control studies 평가도구의 문항은 29개, RoBANS 평가도구의 문항은 21개문항이 부합하였다. 따라서 STROBE 체크리스트 보고지침과 분석대상 NRS 연구의 질 평가도구는 일반적으로 잘 부합되었고, SIGN:Cohort studies 및 Case-control studies 평가도구가 가장 잘 부합되었다.

여기에서 부합된다는 것은 지침에 해당하는 항목이 있느냐의 여부인데 평가항목이 많을수록 부합도는 올라가게 된다. 따라서 질 평가 도구와 보고지침이 반드시 일치하는 것은 아니다. 실제로 SIGN 도구는 부합도가 높지만 National Evidence-based healthcare Collaborating Agency (NECA) 체계적 문헌고찰(Huh, 2011)에서 권장하지는 않고 있다.

주의해야 할 사항은 분석결과 부합되는 사항이 많다고 해서 반드시 좋은 평가도구인 것은 아니라는 점이다. 간호연구들은 무작위 연구보다는 비 무작위연구들이 많기 때문에 간호연구를 위해 어떤 질 평가도구가 적합한가를 더 숙고해 볼 필요가 있다.

결론 및 제언

본 연구에서는 연구결과들이 실무에 적용할 근거로 활용될 수 있도록 연구의 질을 평가하는 다양한 도구들과 보고 지침 들을 분석하였다. 실험 연구의 경우, 대상군 선정, 혼동변수, 중재(노출)측정, 결과평가 눈가림, 불완전한 결과자료, 선택적 보고 등의 영역에서의 비틀림 위험을 평가해야 하므로 연구 설계 단계부터 이를 고려해야 한다. 또한 연구결과와 타당성을 확보하여 근거의 질을 보장 받을 수 있도록 연구설계 유형별 연구결과 보고지침을 적극 활용하도록 권장할 필요가 있다. 궁극적으로 간호연구의 질 평가는 연구보고의 질을 향상시키고 근거기반간호의 정착화에 기여할 것이다. 추후 연구보고지침에 대해 포괄적인 분석을 제언한다.

REFERENCES

- Atkins, D., Best, D., Briss, P. A., Eccles, M., Falck-Ytter, Y., Flottorp, S., et al. (2004). Grading quality of evidence and strength of recommendations. *British Medical Journal*, 328, 1490-1494.
- Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., et al. (1996). Improving the quality of reporting of randomized controlled trials: The CONSORT statement. *The Journal of the American Medical Association*, 276, 637-639.
- CONSORT: transparent reporting tirals [Internet]. [cited 2012 May 17]. Available from: <http://www.consort-statement.org/consort-statement/>
- Downs, S. H., & Black, N. (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *Journal of Epidemiology & Community Health*,

- 52, 377-384.
- Gu, M. O., Cho, M. S., Cho, Y. A., Jeong, J. S., Eun, Y., Jeong, I. S., et al. (2012). A prioritizing for the evidence-based nursing practice guidelines development. *Journal of Korean Clinical Nursing Research*, 18(1), 39-51.
- Health Insurance Review and Assessment Service, & Hallym University Industry Academic Cooperation Foundation. (2009, March). Development of an appraisal instrument for assessing the quality of clinical research.
- Higgins, J. P. T., Green, S., (editors). *Cochrane Handbook for Systematic Reviews of Interventions* Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
- Huh, D. S. (2011). *NECA's guidance for undertaking systemic reviews and meta-analysis for intervention*. Seoul: National Evidence-based healthcare Collaborating Agency.
- Jadad, A. R., Moore, R. A., Carroll, D., Jenkinson, C., Reynolds, D. J., Gavaghan, D. J., et al. (1996). Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials*, 17(1), 1-12. [http://dx.doi.org/10.1016/0197-2456\(95\)00134-4](http://dx.doi.org/10.1016/0197-2456(95)00134-4)
- Kim, H. J., & Kim, S. Y. (2009). Quality Assessment of Non-Randomized Studies in the Korean Journal of Family Medicine. *Korean Journal of Family Medicine*, 30, 129-137.
- Kim, J. H., Choe, M. A., Kim, Y. K., Kim, J. H., Kim, H. S., Park, M. J., et al. (2006). An analysis on the researches using biological measurement in major Korean nursing journals. *Journal of Korean Biological Nursing Science*, 8(2), 61-72.
- Kim, S. Y., Ji, S. M., Lee, S. J., Lee, Y. J., Park, J. E., Nam, M. H., et al. (2011). Guidance for development of clinical practice guidelines Ver 1.0. National Evidence-based Healthcare Collaborating agency [NECA].
- Lee, J. U. (2009). Reporting guidelines of medical research. *Korean Journal of Family Medicine*, 30, 1-14.
- Lee, K. E., Park, Y. R., Cho, K. J., & Park, M. J. (2011). Research trends in the Korean Biological Nursing Science -Based on analysis of the research papers published in the Journal of Korean Biological Nursing Science from 1999 to 2010. *Journal of Korean Biological Nursing Science*, 13(2), 81-93.
- Lim, S. M., Shin, E. S., Lee, S. H., Seo, K. H., Jung, Y. M., Jang, J. E. (2011). Tools for assessing quality and risk of bias by levels of evidence. *Journal of the Korean Medical Association*, 54, 419-429. <http://dx.doi.org/10.5124/jkma.2011.54.4.419>
- Moher, D., Jones, A., & Lepage, L. (2001). Use of the CONSORT statement and quality of reports of randomized trials : a comparative before and after evaluation. *The Journal of the American Medical Association*, 285, 1992-1995.
- Moher, D., Schulz, K. F., & Altman, D. G. (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *Annals of Internal Medicine*, 134, 657-662.
- Park, J., Lee, Y., Seo, H., Jang, B., Son, H., Kim, S. Y., et al. (2011). Risk of bias assessment tool for non-randomized studies (RoBANS): Development and validation of a new instrument. Paper presented at the meeting of the 19th Cochrane colloquium VI international conference on patient safety. Madrid.
- Park, M. H. (2006). *Understanding and application of evidence based nursing*. Seoul: Koonja.
- Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (1997). *Evidence-Based Medicine: How to Practice and Teach EBM*. (1th Ed.). London: Churchill Livingstone.
- Sackett, D. L., Straus, S. E., Richardson, W. S., Rosenberg, W., & Haynes, R. B. (2000). *Evidence-Based Medicine: How to Practice and Teach EBM*. (2nd Ed.). London: Churchill Livingstone.
- Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomized trials. *Annals of Internal Medicine*, 152, 1-8.
- Scottish Intercollegiate Guidelines Network. (2011). SIGN 50: a guideline developers' handbook (Revised ed.) 2011 [Internet]. Edinburgh : Scottish Intercollegiate Guidelines Network; 2011 [cited 2012 May 12]. Available from: <http://www.sign.ac.uk/pdf/sign50.pdf>.
- Slim, K., Nini, E., Forestier, D., Kwiatkowski, F., Panis, Y., Chipponi, J. (2003). Methodological index for non-randomized studies (minors-MINORS): development and validation of a new instrument. *ANZ Journal of Surgery*, 73, 712-6.
- STROBE statement [Internet]. [cited 2012 May 17]. Available from : <http://www.strobe-statement.org/>
- von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenberg, J. P. (2008). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology*, 61, 344- 349.
- Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., et al. (2009). The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses [Internet]. Ottawa: Ottawa Hospital Research Institute; c1996-2010 [cited 2012 May 12]. from: http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.