

국내 학술지에 게재된 환자보고형 건강평가 측정도구에 대한 체계적 고찰 연구의 현황

채덕희¹, 이지연², 이은현³

¹전남대학교 간호대학 교수

²연세대학교 간호대학 교수

³아주대학교 보건대학원 교수

Current status of systematic review studies on patient-reported outcome measures published in Korean journals

Duck-Hee Chae¹, Jiyeon Lee², Eun-Hyun Lee³

¹Professor, College of Nursing, Chonnam National University, Gwangju, Korea

²Professor, College of Nursing, Yonsei University, Seoul, Korea

³Graduate School of Public Health, Ajou University, Suwon, Korea

Purpose: This study aimed to explore the current status of systematic review studies on patient-reported outcome measures (PROMS) published in Korean journals.

Methods: Studies on the systematic review of PROMs, which applied the COnsensus-based Standards for selecting health Measurement Instruments (COSMIN) methodology. There were searched from the KoreaMed. The studies were evaluated based on the eight steps suggested by the COSMIN systematic review processes.

Results: A total of eight systematic review studies were published in seven Korean journals. Patient-reported outcome measures (PROMS) measure for post-traumatic growth, social adjustment, physical impairment in post-intensive care syndrome, sleep quality, constipation, discharge readiness, post-traumatic stress disorder, and eHealth literacy.

Conclusions: No studies adhered to the steps of the COSMIN systematic review guidelines. Among the steps, the weakest points were searching strategies and evaluating measurement properties.

Keywords: Patient reported outcome measures; Questionnaire; Systematic review

서론

연구의 필요성

간호연구자 및 실무자는 임상이나 지역사회에서 대상자의 신체, 정신 및 사회적 기능 등에 대한 여러 측면들을 측정한다. 측정 대상이

되는 개념에는 체중과 같이 기구나 장비를 사용해서 직접적으로 측정할 수 있는 것이 있고, 자기효능감 같은 추상적인 개념이 있는데 이는 주로 질문을 통한 간접적 방법으로 측정한다. 간접 측정이란 추상적 개념이 가지고 있는 속성들을 도출해서 이를 질문 형태로 만들어서 심리측정 속성 또는 심리계량적 속성(psychometric prop-

주요어: 환자보고형 건강평가, 질문지, 체계적 고찰

Received: January 3, 2025; **Accepted:** January 3, 2025

Corresponding author: Eun-Hyun Lee

Graduate School of Public Health, Ajou University, 164 World cup-ro, Yeongtong-gu, Gyeonggi-do, Suwon 16499, Korea

Tel: +82-31-219-5296 Fax: +82-31-219-5025 E-mail: ehlee@ajou.ac.kr

erties) 검정을 통해 만든 질문지 유형의 측정도구를 사용하는 것이다. 이렇게 만들어진 도구를 오래전부터 자가보고형 측정도구(self-reported instrument)라고 불려왔다. 의료계에서는 환자(개인) 자신이 인지한 건강 상태가 어느 정도인지를 파악하기 위해 사용하는 질문지를 환자보고형 건강평가 도구(Patient-Reported Outcome Measures: PROMs)로 부른다[1]. 그래서 최근 의학관련 학술지를 보면 자가보고형 질문지라는 표현 이외에도 PROMs라는 표현을 많이 볼 수 있다(두 가지 중에서 어떤 것을 사용해도 상관없으나 여기서는 PROMs이라는 용어를 사용한다).

건강과 관련된 추상적 개념을 측정하기 위해서 수많은 PROMs가 개발되고 있다. 연구자나 실무자의 고민은 많은 PROMs 중에서 어떤 것을 사용해야 하는가? 라는 것이다. 다시 말해서 심리측정 속성이 충족된 도구인지를 확인하고 사용해야 하는데, 모든 연구자나 실무자가 심리측정 속성을 평가할만한 충분한 지식을 갖고 있지는 않다. 이에 PROM 선정 시, 근거에 기반하는 것이 아니라 다른 사람들이 자주 사용하는 것을 따라 사용하는 경우가 빈번하다.

연구자나 실무자에게 근거 기반의 가장 적합한 PROM이 어떤 것인지에 대한 정보를 제공하기 위한 목적으로 2005년 COnsensus-based Standards for the selection of health Measurement Instruments (COSMIN) 운영위원회가 조직되었다. 이후 2006년 프로토콜 개발을 위한 델파이연구[2]를 시작으로 PROMs체계적 고찰을 위한 COSMIN 가이드라인 버전 1.0이 완성되었으며[3], 최근 COSMIN 가이드라인 버전 2.0, COSMIN Risk of Bias checklist버전 3.0, 그리고 측정 속성 결과 기준표(criteria for good measurement properties) 버전 2.0이 업데이트 되었다[4].

연구의 목적

COSMIN을 적용한 PROM체계적 고찰 연구는 2015년 이후 해마다 100편 이상 국제 학술지에 게재되고 있다[4]. 이러한 세계적인 흐름 속에서, 현재 국내 학술지에 게재된 PROM체계적 고찰 연구의 현주소가 어떠한 지 알아볼 필요가 있다. 따라서 본 연구의 목적은 국내 의료·보건분야 학술지에 게재된 PROMs 체계적 고찰 연구 현황을 파악하기 위함이다.

연구방법

연구설계 및 대상

국내 의료·보건분야 학술지에 게재된 PROMs 체계적 고찰 연구의 현황을 파악하기 위한 고찰연구다. 국내 의료·보건분야 논문 데이터베이스인 KoreaMed에서 COSMIN을 적용한 PROMs 체계적 고찰연구가 그 대상이다.

평가 틀

COSMIN 가이드라인에 따르면, PROMs 체계적 고찰 과정은 아래와 같이 8 단계로 이루어진다[4]. 각 단계에 대한 수행과정을 평가의 틀로 사용했다.

① 연구문제 정하기

연구문제에는 일반적으로 다음과 같은 4가지 기본 요소를 포함한다: 관심 구성개념, 표적 모집단, 도구 유형(자가보고형 질문지), 측정속성(모든 또는 일부 측정속성). 이러한 4가지 기본요소에 대한 결정은 다음 2, 3 단계에도 지속적으로 영향을 미친다. 구체적인 연구문제의 유형으로는 특정 PROM에 대한 모든 심리측정 연구를 대상으로 수행한 체계적 고찰 연구, 특정 모집단에서 특정 개념을 측정하는 모든 PROMs에 관한 심리측정 연구를 대상으로 수행한 체계적 고찰 연구, 그리고 모집단을 특정하지 않고 특정 개념을 측정하는 모든 PROMs에 관한 심리측정 연구를 대상으로 수행한 체계적 고찰 연구가 있다.

② 선정 기준 정하기

1단계에서 언급된 4가지 기본 요소(구성개념, 대상자, 측정도구 유형 및 측정속성)에 관한 구체적 선정기준을 정한다.

③ 문헌 검색 전략 설정

문헌 검색은 MEDLINE (PubMed) 및 EMBASE를 기본적으로 사용하고, SCOPUS, Web of Science, CINAHL 또는 PsycINFO 등 다른 데이터베이스를 부가적으로 사용한다. 그리고 각 데이터베이스에 맞게 4가지 기본 요소(구성개념, 대상자, 측정도구 유형 및 측정속성)에 대한 검색 필터를 만든다[5].

④ 문헌검색 시행

만들어진 검색 필터를 불리언 논리연산자(Boolean operator) “AND”로 연결한다. 그리고 배제기준 필터를 “NOT”로 연결해서 검색한다[5]. 문헌검색을 시행하려면, 최소 3명의 연구자가 필요하다. 이 중에서 2명의 연구자는 서로 독립적으로 검색된 논문을 스크린하고 이견이 생기면 나머지 1명과 상의해서 대상 연구논문을 선정한다. 이 과정을 PRISMA-COSMIN flow diagram [6]으로 정리해서 보고한다.

⑤ PROMs 특성 자료 도출

문헌 검색을 통해 선정된 PROM의 특성에 대한 자료를 도출한다. 즉, PROM의 이름, 사용언어, 국가, 문항 수, 하부척도, 응답 반응, 응답 관련 회상 기간(recall period), 그리고 도구에서 측정하고자 하는 구성개념 정의 등을 도출한다.

⑥ 측정속성(measurement properties) 평가

PROM 측정속성(내용타당도, 구성타당도, 내적일관성, 교차문화 타당도/측정동일성, 측정-재측정 신뢰도, 측정오차, 수렴타당도, 판별타당도, 준거타당도 및 반응도) 평가는 아래와 같이 다시 6개의 과정으로 이루어진다.

⑥-1 연구에 포함된 대상자 특성(대상자, 국가, 사용언어, 표본 크기, 성별, 나이 및 질병 특성 등) 자료를 도출한다.

⑥-2 연구방법의 질(methodological quality of studies)을 평가한

다. 측정속성 결과에 대한 자료를 도출하고 COSMIN Risk of Bias checklist를 이용하여, 각 연구에서 수행된 측정속성별 연구 설계나 분석 방법에 ‘very good’, ‘adequate’, ‘doubtful’, ‘inadequate’로 평가한다.

⑥-3 측정속성에 관한 연구결과를 평가(rating of the study results)한다. 측정속성 결과 기준표(criteria for good measurement properties)에 준해서, 각 연구에서 측정속성 별로 결과를 ‘sufficient (+)’, ‘insufficient (-)’, ‘indeterminate (?)’로 평가한다.

⑥-4 PROM 별로 각 측정 속성에 대해서 질적/양적으로 요약한다. 각 PROM에 대한 연구가 여러 개 있으면, 이 연구 결과들을 측정속성별로 요약한다. 측정속성 결과의 범위를 질적으로 요약(qualitatively summary)하거나 통계적 관점에서 양적으로 요약(quantitatively pooling)한다.

⑥-5 요약된 결과에 대한 종합적 평가를 한다. PROM 별로 각 측정속성에 관한 결과가 양적/질적으로 요약되었으면, 이 결과를 측정속성 결과 기준표(criteria for good measurement properties)에 준해서 PROM 별로 각 측정속성에 대해서 종합적으로 ‘sufficient (+)’, ‘insufficient (-)’, ‘inconsistent (±)’, ‘indeterminate (?)’로 평가한다.

⑥-6 근거의 질(quality of evidence)을 평가한다. PROM 별로 요약된 결과가 얼마나 신뢰할 만한 것인지를 판정한다. 비뮴립위험(risk of bias), 불일치성(inconsistency), 부정확성(imprecision), 및 간접성(indirectness) 요인에 대해 GRADE 방법을 적용해서 근거의 질을 ‘high’, ‘moderate’, ‘low’, 및 ‘very low’로 등급을 매긴다.

⑦ 추천되는 PROM 선정

위 단계를 모두 수행한 후에, 연구자는 결론을 내려야 한다. 즉, 어떤 PROM의 사용을 추천할 것인지, 적합한 도구의 부재라고 할 것인지, 또는 도구 사용에 대한 추천 보류라고 할 것인지에 대한 의견을 밝힌다[7].

⑧ 보고서 작성

PRISMA-COSMIN 지침을 참조하여 보고서를 작성한다[6].

연구결과

본 고찰에 포함된 체계적 고찰연구는 총 8편이었다[8-15]. 외상 후 성장, 사회적응, 집중치료를 증후군의 신체적 장애, 수면의 질, 퇴원준비, 외상 후 스트레스 장애, 및 인터넷 헬스리터러시(eHealth literacy: eHL) 개념을 측정하는 PROMs 체계적 고찰 연구였다.

Table 1은 본 조사에서 선정된 8편에 대해서 PROMs 체계적 고찰의 8단계에 따라 평가한 것이다. 1단계 연구문제는 8편 모두에서 명확하게 제시되었다. 2단계 선정기준에서는 4편 만이 COSMIN 지침을 충족하는 것으로 나타났다[8,9,12,15]. 3 단계 검색전략에서는 8편 모두 사용된 검색필터에 문제가 있는 것으로 나타났다.

예를 들어, 측정속성(measurement properties)을 검색하기 위해 사용된 검색어는 “validation OR development” 뿐이었다[12]. 4 단계 문헌검색에서는 3편의 연구가 지침을 만족하지 못했다. 기본적으로 사용해야 할 데이터베이스를 사용하지 않았거나[15] 검색을 하는데 필요한 최소한의 연구자 수를 충족하지 못했다[8,11]. 5단계 PROMs 특성에 대한 자료는 8편의 연구에서 모두 보고했다. 6단계 측정속성 평가는 COSMIN을 적용한 체계적 고찰 연구의 핵심이며 가장 까다로운 분석단계다. 이 단계는 다시 6개의 순차적인 세부 과정으로 구성된다. 8편의 연구 중에서, 6개 세부과정을 모두 보고한 연구는 없었다. 7단계 추천 PROM 선정과 관련하여 2편 [8,11]의 연구에서는 불명확하게 보고되었으며, 나머지 6편에서는 추천하는 PROM이 없다고 결론지었다. 8단계 보고서 작성에서는 모든 논문이 원고를 작성했고, Korean Journal of Adult Nursing (2 편), Child Health Nursing Research (1편), Journal of Korean Critical Care Nursing (1편), Journal of the Korean Society of Maternal and Child Health (1편), Psychiatry Investigation (1편), Journal of Neurogastroenterology and Motility (1편), Research in Community and Public Nursing(1편)에 게재되었다.

논의

국내 의료보건의분야 학술지에 게재된 논문 중에서 COSMIN을 적용한 PROM 체계적 고찰연구는 총 8편이었다. 연구문제 유형을 보면, 특정 PROM (Pittsburgh Sleep Quality Index)에 대한 심리측정 연구를 대상으로 수행한 체계적 고찰 연구가 1편[11], 모집단을 특정하지 않고 특정 개념(변비 증상)에 대한 심리측정 연구를 대상으로 체계적 고찰을 수행한 연구가 1편[12], 나머지는 특정 모집단에서 특정 개념에 대한 심리측정 연구를 대상으로 실시한 체계적 고찰 연구였다. 따라서 모든 논문의 연구문제 유형은 COSMIN에서 제시한 범위 내에 포함된 것으로 볼 수 있다.

COSMIN 지침에 의하면, PROMs 체계적 고찰에서 제외해야 할 연구 유형이 있다. 첫째, 연구목적이 심리측정 평가가 아닌 경우다. 상관관계연구나 실험연구에서 PROMs이 사용되는데, 연구에 사용된 PROMs의 측정속성 (예, Cronbach's alpha)에 대한 정보를 서술한다. 체계적 고찰연구에서는 PROMs이 이런 목적으로 사용된 논문은 배제해야 한다. 둘째, 연구자가 체계적 고찰을 하려는 관심 개념을 측정하는 PROM이 다른 측정도구의 타당성을 검증하기 위한 연구에서 비교대상 PROM으로 사용된 경우다. 본 조사에 의하면, 4편에서 이런 선정 기준을 충족하지 못했다[10-14].

체계적 고찰연구에서 문헌검색은 매우 중요하다. 4가지 기본 요소(구성개념, 대상자, 측정도구 유형 및 측정속성)에 대한 검색 필터를 연구자가 사용하려는 각 데이터베이스에 맞게 만들어 사용해야 한다. 본 조사에서는 이러한 문헌 검색 조건을 충족하는 연구를 발견할 수 없었다. 이는 관심 개념을 측정하는 심리측정 연구들 중에서 일부만을 검색해서 체계적 고찰을 수행했음을 의미한다.

본 조사에 포함된 논문 중에서, 2024년 게재된 eHL에 대한 체계

Table 1. Systematic Review Studies Published in Korean Journals

Author	Concept	Population	Aim of the systematic review ①	Eligibility criteria ②	Searching strategies: searching filters for 4 key elements [†] per searching database ③	Conducting the literature search: used databases ④
Kim & Lee [†] [8]	Posttraumatic growth	Breast cancer	Systematic review of all available studies on the quality of a selection of PROMs that aim to measure a posttraumatic growth in breast cancer	○	△ Primitive searching terms were used.	△ PubMed, EMBASE, CINAHL, PsycINFO et al. Screened by two researchers.
Oh et al. [†] [9]	Social adjustment	Youth after treatment of childhood cancer	Systematic review of all available studies on the quality of a selection of PROMs that aim to measure a social adjustment in youth after treatment of childhood cancer	○	△ Primitive searching terms were used for population.	○ Medline, EMBASE et al.
Kang et al. [10]	Physical impairment in post-intensive care syndrome	Patients with experience administered in an intensive care unit	Systematic review of all available studies on the quality of a selection of PROMs that aim to measure a physical impairment in post-intensive care syndrome of intensive care unit survivors	△ Cross-sectional, experimental, and cohort studies were included.	△ Primitive searching terms were used.	○ PubMed, EMBASE, CINAHL, Cochrane Library et al.
Kim & Lee [11]	Sleep quality	Adults with trauma experience	Systematic review of all available studies on the quality of the Pittsburgh Sleep Quality Index	△ Cross-sectional and randomized controlled studies were included.	△ Primitive searching terms were used.	△ PubMed, EMBASE, CINAHL, Cochrane Library et al. Screened by two researchers.
Lee et al. [12]	Constipation symptoms	Patients	Systematic review of all available studies on the quality of a selection of PROMs that aim to measure constipation symptoms in patients	○	△ Primitive searching terms were used.	○ PubMed, EMBASE, and PsycINFO
Kim et al. [13]	Discharge readiness	Parents of high-risk infants	Systematic review of all available studies on the quality of a selection of PROMs that aim to measure discharge readiness in the parents of high-risk infants	△ Cross-sectional, randomized controlled, experimental, and cohort studies were included.	△ Primitive searching terms were used.	○ Medline, EMBASE, and PsycINFO et al.
Kim et al. [14]	Posttraumatic stress disorder symptoms	North Korean defectors	Systematic review of all available studies on the quality of a selection of PROMs that aim to measure posttraumatic stress disorder symptoms in North Korean defectors	△ All kinds of study designs were included.	△ Searching filters only for construct and population were used.	○ Medline, EMBASE et al.
Ahn & Kim [15]	eHealth literacy	Community population	Systematic review of all available studies on the quality of a selection of PROMs that aim to measure eHealth literacy within community settings	○	△ Primitive searching filters were used for population and type of instrument.	△ PubMed, CINAHL et al.

(Continued to the next page)

Table 1. Continued

Author	PROM characteristics ⑤	Evaluation of measurement properties ⑥					Selection of the best PROM ⑦	Writing report (published journal) ⑧
		Characteristics of the study population ⑥-1	Methodological quality of studies based on COSMIN Risk of Bias checklist ⑥-2	Rating of the study results against criteria for good measurement properties ⑥-3	Summarizing the results per a PROM: quantitative summary or quantitative pooling ⑥-4	Rating the summarized results against criteria for good measurement properties ⑥-5		
Kim & Lee [†] [8]	○ PROM name, Language, Number of subscales/items, Administration mode	△ Participants, Languages,	○	-	-	-	Unclear	Korean Journal of Adult Nursing
Oh et al. [†] [9]	○ PROM name, Number of subscales/items, Response options,	△ Language	○	○	-	△	Recommendation against its use.	Child Health Nursing Research
Kang et al. [10]	○ PROM name, Administration mode, Country, Number of subscales/items	△ Disease characteristics	-	○	-	-	Recommendation against its use.	Journal of Korean Critical Care Nursing
Kim & Lee [11]	○ PROM name, Administration mode, Country, Number of subscales/items, Response options	△ Participants, Sex, Age, Sample size	-	○	-	-	Unclear	Korean Journal of Adult Nursing
Lee et al. [12]	○ PROM name, Administration mode, Country, Language, Number of subscales/items, Response options, recall period	○ Participants, Location of study, Sample size, Type of constitution	○	○	△ Psychometric results of the Bowel Function Index were not summarized.	-	Not enough evidence.	Journal of Neurogastroenterology and Motility
Kim et al. [13]	○ PROM name, Country, Number of subscales/items	○ Participants, Sample size	-	○	-	-	Recommendation against its use.	Journal of the Korean Society of Maternal and Child Health
Kim et al. [14]	○ PROM name, Country, Number of subscales/items	○ Participants, Sample size, Age, Spent time in South Korea	○	○	-	-	Recommendation against its use.	Psychiatry Investigation
Ahn & Kim [15]	○ PROM name, Country, Number of subscales/items, Response options	○ Participants, Age, Country, Sample size	○	○	-	-	Unclear	Research in Community and Public Nursing

[†] construct, population, type of instrument, and measurement properties.

[‡] An old guideline before the version 1.0 was applied.

①-⑧: Eight steps of COSMIN systematic review.

○ sufficient information; △ Information was provided, but not enough or partially incorrect; - No information.

적 고찰 연구[15]의 문헌 검색에 대해 살펴보자. 이 연구에서는 PubMed (n=413), CINAHL (n=32), Google Scholar (n=269), RISS (n=13), other sources (n=16)를 검색해서 중복을 제외하고 47편이 도출되었으며, 최종적으로 분석에 포함된 논문은 21편이었다. Lee, Lee, & Chae는 2021년에 동일 개념인 eHL를 측정하는 PROMs에 대한 체계적 고찰 연구를 수행했다[16]. 연구자들은 COSMIN 지침에 따라서 각 데이터베이스에 맞는 검색 필터를 만들어서 사용했다. 그 결과 PubMed (n=783), CINAHL (n=402), EMBASE (n=932), PsycInfo (n=238) 검색에서 중복을 제외하고 1,868편이 도출되었으며, 최종적으로 41편 논문에서 보고한 57개의 심리측정 연구를 대상으로 분석했다. 이 두 편의 비교는 문헌검색에 사용된 검색 필터와 데이터베이스의 중요성을 보여주는 사례라고 할 수 있다.

문헌검색에 사용된 상세한 검색 필터는 supplementary file로 제공해서 다른 연구자들이 확인할 수 있도록 제공하는 것이 권장된다. 예를 들어, Chae 등[17]이 수행한 Patient Health Questionnaire-9에 대한 체계적 고찰 연구를 살펴보면, PubMed, EMBASE, CINAHL, PsycInfo 및 Cochrane 등의 문헌 데이터베이스별로 만들어서 사용한 검색필터를 supplementary file로 제시한 것을 볼 수 있다. 때로는 학술지 논문 심사과정에서 검색 필터를 제공하라는 심사평을 하기도 하며, 더 나아가 심사자가 실제로 제공된 필터를 사용하여 검색을 확인하는 경우도 있다.

PROM 체계적 고찰의 6번째 단계인 측정속성에 대한 평가는 COSMIN을 적용한 체계적 고찰의 하이라이트라고 할 수 있다. 측정속성에 대한 평가는 다시 여섯 과정(⑥-1부터 ⑥-6)을 거쳐야 하지만, 본 조사에 포함된 8편 중에서 이 과정을 모두 수행한 연구는 없었다. 여섯 과정 중에 가장 먼저 실시해야 하는 것은 PROM 연구에 참여한 대상자 특성에 관한 자료를 정리하는 것이다(⑥-1). 이런 특성 중에서 각 연구에 사용된 표본 크기 정보를 확보하는 것은 매우 중요하다. 이는 각 PROM 연구에서 측정속성에 대한 방법론적 질을 평가할 때(⑥-2), PROM 별로 결과를 요약할 때(⑥-4) 및 근거의 질에 대한 등급을 매길 때(⑥-6) 반영되기 때문이다. 그럼에도 불구하고, 2편에서는 표본크기에 대한 정보를 찾아볼 수 없었다[8,9]. 본 조사에 포함된 8편 중에서, Lee 등[12]의 연구가 측정속성 평가 과정을 가장 많이 수행한 것으로 나타났다(⑥-1 부터 ⑥-4). 연구자는 문헌검색을 통해서 변비증상을 측정하는 23개 PROMs를 찾았다. 그리고 23개 중, 21개 PROMs는 심리측정 평가 논문이 각각 한편씩만 존재하여 평가를 더 진행하지 않았다고 보고했다. 보통 이런 경우, 각 PROM별 결과 요약은 큰 의미가 없기 때문에 측정속성 평가를 더 이상 진행하지 않는다. 국외 학술지에 게재된 Depression Anxiety Stress Scales-21[18] 및 General Health Questionnaire-9(PHQ-9)[17]에 대한 체계적 고찰 연구에서도 심리측정 연구가 한번만 시행된 PROM에 대해서는 요약된 결과에 대한 전반적 평가와 근거의 질에 대한 등급을 평가하지 않았다. 하지만, Lee 등[12]의 연구에서 23편의 변비증상 PROMs중에 Bowel Function Index (BFI)에 대한 심리측정 연구는 3편이 있었다. 그럼

에도 불구하고 연구자가 이 3편에 대한 각 측정속성을 양적/질적으로 요약하지 않았다는 점은 매우 아쉽다. 만약에 요약을 했다면, 그 다음과정(⑥-5 및 ⑥-6)도 평가할 수 있었을 것이다. 특히 BFI의 내적일관성(Cronbach's alpha) 속성의 경우는 메타분석을 적용한 양적 요약도 가능하다. 내적일관성을 양적으로 요약한 사례로는 당뇨병 자가관리[19], 당뇨병 자기효능감[20], eHL[16], 및 PHQ-9[17]에 대한 체계적 고찰 연구 등이 있다.

Oh 등[9]이 수행한 측정속성 평가를 살펴보면, 각 PROM에 대한 전반적인 근거의 질이 보고되었다. 하지만 어떻게 이런 결과를 도출했는지 의문이 남는다. 예를 들어, 연구자는 사회 적응 측정 도구 Minneapolis-Manchester Quality of Life-Adolescent (MMQL-AF)에 대한 3편(미국, 일본 및 한국 연구 각 1편씩)의 심리측정 연구를 합쳐서 전반적인 MMQL-AF 측정속성에 대한 근거의 질을 제시했다. 하지만 미국, 일본 및 한국의 MMQL-AF는 하부척도의 수 또는 문항의 수가 서로 달랐다. COSMIN에 의하면, 이런 경우는 같은 측정 도구라고 볼 수 없으며, 각각 따로 취급해서 분석해야 한다. 예를 들어, Lee 등[16]의 eHL 관련 PROMs체계적 고찰 연구를 살펴보자. 이 연구에서 찾은 eHealth Literacy Scale (eHEALS) 심리측정 연구는 38편이 있었는데, 하부척도와 문항의 묶임(clustering)에 따라서 12 종류로 분류하고 12 종류 별로 측정속성을 평가했다.

측정속성에 대한 평가가 끝났으면, 이제 남은 단계는 평가된 PROMs 중에서 어떤 것을 사용하라고 추천할 것인지를 결정하고 보고서를 작성하는 일이다. 본 조사에 포함된 체계적 고찰 연구 중에서 외상 후 성장, 사회 적응, 집중치료 후 증후군의 신체적 장애, 수면의 질, 퇴원준비, 외상 후 스트레스 장애 PROMs를 다룬 연구자들은 실무나 연구에서 사용을 추천할 만한 도구가 없다는 입장을 밝혔고, 나머지 수면의 질 및 eHL을 다룬 연구자들의 결정은 무엇인지 명확하게 기술되지 않았다. 앞서 말한 것처럼, 본 조사에 포함된 체계적 고찰 연구 8편 중에는 측정속성을 평가하는 여섯 과정(⑥-1부터 ⑥-6)을 모두 수행한 연구는 없었다. 그렇다면, 이런 결정들을 얼마나 신뢰할 수 있는가?

PROMs 체계적 고찰 연구를 하는 궁극적인 이유는 의료·보건분야 전문가들이 실무나 연구를 할 때, 어떤 측정 도구를 사용해야 하는지에 대한 고민에 도움을 주기 위함이다. 그렇기 때문에 PROMs 체계적 고찰 연구에서 결과에 대한 결정은 매우 중요하고 그 파급력이 크다고 할 수 있다. 그러므로 연구자(팀)는 체계적 고찰을 수행하기 전에 COSMIN 지침을 숙지해야 하며, 학술지 편집위원 및 심사자도 워크숍 등을 통해서 PROMs 체계적 고찰 방법을 학습해서 국내 의료·보건분야 학술지에 더 높은 학술적 가치를 지닌 논문이 지속적으로 게재될 수 있도록 노력할 필요가 있다.

결론

본 연구에서는 국내 의료·보건분야 학술지에 게재된 PROMs 체계적 고찰 연구의 현황을 살펴보았다. COSMIN지침을 적용한 8편의 체계적 고찰 연구가 7개의 국내 학술지에 게재되었다. COSMIN

지침에서 제시한 PROMs 체계적 고찰 8단계 과정을 모두 수행한 연구는 없었다. 특히 문헌검색 전략/시행 단계 및 측정평가 단계 중에서 각 PROM별로 질적/양적 요약하기, 각 PROM별로 요약된 결과에 대한 전반적인 평가와 근거의 질 평가가 매우 취약한 것으로 나타났다. 국내 연구자와 실무자에게 근거 기반으로 가장 적합한 PROM에 대한 정보를 제공하기 위해서는, PROMs 체계적 고찰 방법을 숙달하고 이를 적용하기 위한 연구자와 학술지 편집·심사위원의 지속적인 노력이 필요하다.

ORCID

Duck-Hee Chae, <https://orcid.org/0000-0003-3259-7385>

Jiyeon Lee, <https://orcid.org/0000-0001-6413-329X>

Eun-Hyun Lee, <https://orcid.org/0000-0001-7188-3857>

Conflict of interest

Duck-Hee Chae has been an editorial board member of the Research in Community and Public Health Nursing. She was not involved in the review process of this manuscript. No conflict of interest has been declared by all authors.

Funding

None.

Authors' contributions

Duck-Hee Chae and Jiyeon Lee contributed to data curation, formal analysis, investigation, writing-review & editing, and validation. Eun-Hyun Lee contributed conceptualization, formal analysis, investigation, writing-original draft, supervision, and validation.

Data availability

Please contact the corresponding author for data availability.

Acknowledgments

None.

References

1. U.S. Food and Drug Administration. Food and Drug Administration. Guidance for industry on patient-reported outcome

measures: Use in medical product development to support labeling claims. Federal Register. 2009;74(235):65132–65133. <https://www.fda.gov/media/77832/download>

2. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, et al. Protocol of the COSMIN study: Consensus-based Standards for the selection of health measurement instruments. BMC Medical Research Methodology. 2006;6:1–7. <https://doi.org/10.1186/1471-2288-6-2>
3. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of Life Research. 2018;27(5):1147–1157. <https://doi.org/10.1007/s11136-018-1798-3>
4. Mokkink LB, Elsman EBM, Terwee CB. COSMIN guideline for systematic reviews of patient-reported outcome measures version 2.0. Quality of Life Research. 2024;33(11):2929–2939. <https://doi.org/10.1007/s11136-024-03761-6>
5. Terwee CB, Jansma EP, Riphagen II, de Vet HC. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Quality of Life Research. 2009;18(8):1115–1123. <https://doi.org/10.1007/s11136-009-9528-5>
6. Elsman EBM, Mokkink LB, Terwee CB, Beaton D, Gagnier JJ, Tricco AC, et al. Guideline for reporting systematic reviews of outcome measurement instruments (OMIs): PRISMA-COSMIN for OMIs 2024. Quality of Life Research. 2024;33(8):2029–2046. <https://doi.org/10.1007/s11136-024-03634-y>
7. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine-A practical guide. New York: Cambridge University Press. 2011. <https://doi.org/10.1017/CBO9780511996214>
8. Kim SH, Lee JH. Posttraumatic growth measures in breast cancer survivors: A systematic review. Korean Journal of Adult Nursing. 2017;29(4):343–362. <https://doi.org/10.7475/kjan.2017.29.4.343>
9. Oh SM, Park SY, Lee HJ, Lee JH. Measurement properties of self-reported questionnaires measuring the social adjustment for youth after treatment of childhood cancer: Systematic review. Child Health Nursing Research. 2018;24(1):78–90. <https://doi.org/10.4094/chnr.2018.24.1.78>
10. Kang JY, Lee MJ, Jeong YJ, Kim SK, Cho YS, Park JH, et al. Instruments to assess physical impairments in post-interventive care syndrome: A systematic review. Journal of Korean Critical Care Nursing. 2018;11(1):46–66.
11. Kim GU, Lee JH. Systematic review of the Pittsburgh Sleep Quality Index used for measuring sleep quality among adults

- with traumatic experience. *Korean Journal of Adult Nursing*. 2019;31(4):337–350. <https://doi.org/10.7475/kjan.2019.31.4.337>
12. Lee VV, Lau NY, Xi DJY, Truong ATL, Blasiak A, Siah KTH, et al. A systematic review of the development and psychometric properties of constipation-related patient-reported outcome measures: Opportunities for digital health. *Journal of Neurogastroenterology and Motility*. 2022;28(3):376–389. <https://doi.org/10.5056/jnm22004>
 13. Kim KE, Lee HJ, Jeon NY. Discharge readiness scale for parents of high-risk infants: A systematic review. *Journal of the Korean Society of Maternal and Child Health*. 2022;26(4):205–224. <https://doi.org/10.21896/jksmch.2022.26.4.205>
 14. Kim OS, Kim KA, Chu SH. Properties of patient-reported outcome measures for post-traumatic stress disorder in North Korean defectors: A scoping review. *Psychiatry Investigation*. 2023;20(7):593–608. <https://doi.org/10.30773/pi.2022.0363>
 15. Ahn JW, Kim MY. A systematic review of questionnaire measuring eHealth literacy. *Research in Community Public Health Nursing*. 2024;25(3):297–312. <https://doi.org/10.12799/rcphn.2024.00752>
 16. Lee JY, Lee EH, Chae DH. eHealth literacy instruments: Systematic review of measurement properties. *Journal of Medical Internet Research*. 2021;23(11):e30644. <https://doi.org/10.2196/30644>
 17. Chae DH, Lee JY, Lee EH. Internal structure of the patient health questionnaire-9: A systematic review and meta-analysis. *Asian Nursing Research*. 2025;19(1):1–12. <https://doi.org/10.1016/j.anr.2024.12.005>
 18. Lee JY, Lee EH, Moon SH. Systematic review of the measurement properties of the Depression Anxiety Stress Scales–21 by applying updated COSMIN methodology. *Quality of Life Research*. 2019;28(9):2325–2339. <https://doi.org/10.1007/s11136-019-02177-x>
 19. Lee JY, Lee EH, Chae DH, Kim CJ. Patient-reported outcome measures for diabetes self-care: A systematic review of measurement properties. *International Journal of Nursing Studies*. 2020;105:103498. <https://doi.org/10.1016/j.ijnurstu.2019.103498>
 20. Lee JY, Lee EH, Chae DH. Self-efficacy instruments for type 2 diabetes self-care: A systematic review of measurement properties. *Journal of Advance Nursing*. 2020;76(8):2046–2059. <https://doi.org/10.1111/jan.14411>