

한국어판 Neck Disability Index 번역본의 다양성과 사용 현황: Scoping Review

이가행 · 김선엽^{1†}

대전대학교 대학원 물리치료학과, ¹대전대학교 보건의료과학대학 물리치료학과

Diversity of Korean Translations of the Neck Disability Index and their Current Uses: A Scoping Review

Ga-Haeng Lee, PT · Suhn-Yeop Kim, PT, PhD^{1†}

Department of Physical Therapy, Graduate School, Daejeon University

¹Department of Physical Therapy, College of Health and Medical Science, Daejeon University

Received: November 1, 2025 / Revised: November 3, 2025 / Accepted: December 2, 2025

© 2026 J Korean Soc Phys Med

| Abstract |

PURPOSE: In the absence of an official, authorized Korean version of the Neck Disability Index (NDI) and amid the coexistence and mixed use of unofficial translations, this scoping review aimed to map the types, characteristics, and uses of Korean-language NDI translations in Korea to identify key issues.

METHODS: A single reviewer searched Korean databases (RISS, DBpia, KCI, KoreaMed; 2015-2025), screened records, and charted data (with data extraction verification) according to Arksey-O'Malley, Levac et al., and PRISMA-ScR guidance. A descriptive synthesis was conducted.

RESULTS: Of 94,404 records, 515 studies were included

after screening. Four unofficial Korean translations used in 159/515 (30.9%) studies were identified. The remaining 356/515 (69.1%) used questionnaires of unknown origin or modified versions of such questionnaires. The rules for handling missing responses were reported in only 13/515 (2.5%). Unauthorized inclusion of full questionnaires as appendices occurred in 158/515 (30.7%), and arbitrary changes to the scoring or scale structure were made in 20/515 (3.9%).

CONCLUSION: In the absence of an official Korean NDI, unofficial, unknown-source, and modified versions are widely used, resulting in copyright and ethical risks and a weakened score for comparability and validity. The poor reporting of missing data handling and scoring rules exacerbated these issues. The Mapi Research TrustClick or tap here to enter text., the copyright and licensing body responsible for the NDI, oversees the official Korean version and should adopt a single standard and institutionalize core reporting requirements, including mandatory license verification and restrictions on posting full questionnaires as appendices.

†Corresponding Author : Suhn-Yeop Kim
kimsy@dju.kr, <https://orcid.org/0000-0002-0558-7125>

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.

Key Words: Copyright, Cross-cultural adaptation, Neck Disability Index, Patient Reported Outcome Measures, Scoping Review

I. 서론

환자자기평가결과(Patient-Reported Outcome Measures, PROMs)는 환자가 직접 보고한 증상, 기능, 삶의 질을 표준화된 설문을 통해 측정하는 도구로, 임상시험뿐 아니라 임상 의사결정, 의료 질 관리, 정책 평가 등 다양한 영역에서 널리 사용되고 있다. PROMs는 환자 경험을 정량적으로 반영함으로써 의료진과 환자의 의사소통을 개선하고, 개인 수준의 맞춤형 치료에서 보건의료체계 전반의 질 향상에 이르기까지 중요한 역할을 한다[1].

경부 기능장애 지수(Neck Disability Index; NDI)는 목통증 관련 기능장애를 평가하기 위한 대표적인 PROMs으로, 개발 이후 신뢰도와 타당도가 광범위하게 입증되어 임상과 연구 현장에서 가장 널리 사용되는 도구 중 하나이다. NDI의 저작권은 원저자인 Vernon이 보유하고, 현재는 Mapi Research Trust(MRT)가 공식 관리 기관으로서 도구의 배포 및 다국어 번역 절차를 관장하고 있으며, 현재 공식 한국어판은 아직 없는 상태이다[2,3].

그럼에도 불구하고 국내에서는 NDI가 발표된 이후 여러 연구자들에 의해 한국어 번역이 반복적으로 시도되었으며[4,5], 이 과정에서 공식 절차를 거치지 않은 비공식 번역본과 임의 수정본이 연구 및 임상에 사용되어왔다. 이러한 비공식적 번역본 사용의 문제점은 허리 기능장애를 평가하기 위한 도구인 오스웨스트리 기능장애 지수(Oswestry Disability Index; ODI)의 공식적으로 승인되지 않은 변형판(rogue version)의 사례에서도 볼 수 있다. ODI가 같은 영어권에서도 다양하게 변형되어 사용되는 것에 대해서 원저자인 Fairbank는 ‘공식적으로 검증되지 않은 변형판의 사용은 연구 결과 자체의 신뢰도를 훼손한다’고 명시적으로 언급한 바 있다[7].

PROMs 도구는 국제적으로 확립된 공식 가이드라인에 따라 다양한 단계를 거쳐 개념적, 의미적, 측정학적 등가성

을 확보해야 한다. 이를 따르지 않은 비공식·임의 번역본은 타당도와 국제 비교 가능성의 훼손 위험이 크다[8]. 특히 NDI의 경우 공식 한국어판이 부재한 상황에서 다양한 비공식 번역본이 병존·혼용되어 왔음에도, 각 번역본의 특성과 측정 속성, 실제 임상연구에서의 사용 양상 및 보고상의 문제점을 종합적으로 연구한 연구는 없는 실정이다.

이에 본 스코핑 리뷰는 한국에서 사용된 NDI 번역본의 유형과 특성을 파악하고, 국내 임상연구에서의 사용 현황과 보고 과정의 쟁점을 체계적으로 정리함으로써, 향후 공식 한국어판 개발 및 PROMs 관리 체계 개선에 필요한 기초자료를 제공하고자 한다.

II. 연구 방법

본 연구 방법은 스코핑 리뷰(Scoping Review, ScR)로서 체계적 문헌고찰(Systematic Review, SR)과 달리 특정 증거 효과를 통계적으로 종합하거나 근거 수준을 평가하는 것을 목적으로 하지 않는다. 스코핑 리뷰는 특정 주제에 대한 문헌의 범위와 특성을 탐색하고 연구 공백을 확인하는데 적합한 방법론이다[9]. 이 연구에서는 한국에서 사용된 NDI 번역본의 현황을 탐색하기 위해 Arksey와 O'Malley[9]가 제시한 5단계 방법론적 틀을 기반으로, Levac 등[10]의 확장된 6단계 절차와 Joanna Briggs Institute(JBI)에서 제시한 가이드라인[11]에 따라 스코핑 리뷰를 수행하였다. 또한 최종 보고는 PRISMA-ScR 지침을 따랐다[12]. 한 예로 Al-Ebrahim 등[13]은 스코핑 리뷰를 통해 아랍어로 번역된 PROMs 317개를 검토한 결과, 대부분이 순방향 번역과 역방향 번역은 보고하였으나 환자 사전검증이나 개발자 승인과 같은 국제적으로 인정받기 위한 핵심 단계가 누락된 경우를 밝힘으로써 자국의 평가도구의 번역과정의 공백을 확인하였다.

본 연구는 공개된 문헌을 대상으로 수행되는 스코핑 리뷰로서, 외부 연구비 지원 없이 수행되었으며, 대전대학교 기관생명윤리위원회(IRB)로부터 심의면제 판정을 받았다(승인번호: 1040647-202510-HR-002-03, 승인일: 2025-10-26).

1단계: 연구 질문 규정(Identifying the research question)

본 스코핑 리뷰의 연구질문은 다음과 같다.

- 1) 한국어에서 사용된 NDI 번역본의 종류와 특성(예: 번역·문화적 적합화 절차, 신뢰도와 타당도 등 측정속성 결과)은 무엇인가?
- 2) 한국 내 실제 임상연구에서 한국어판 NDI가 사용된 비율은 어떠하며, 사용과 보고 과정에서 의 문제점은 무엇인가?

위와 같은 연구질문을 JBI 가이드라인[11]에서 제시한 PCC(Population, Concept, Context) 틀을 사용하여 정의하였다.

- 대상(Population): 성인 한국어 사용자를 대상으로 한국어판 NDI를 적용한 임상연구 문헌, 또는 한국어판 NDI의 번역·문화적 적합화 검증을 다룬 문헌(학술지, 학위논문 포함)
- 개념(Concept): 한국어 NDI 번역본의 유형, 특성, 측정속성 결과, 사용 비율 및 문제점
- 맥락(Context): 한국어판 NDI를 대상으로 번역 및 한국 내 임상연구 환경에서 수행한 국내 학술문헌; 2015-2025, 국내 데이터베이스

2단계: 관련 연구의 식별(Identifying relevant studies)

국내 학술 데이터베이스(RISS, DBpia, KCI, KoreaMed)를 중심으로 순차적으로 검색을 수행하였다. 검색일은 2025년 8월 01부터 2025년 10월 23일까지이며, 사용된 주요 검색어는 다음과 같다:

“Neck Disability Index”, “NDI”, “목장애지수”, “목 장애지수”, “목 장애 지수”, “목기능장애지수”, “목 기능장애지수”, “목 기능 장애지수”, “목 기능 장애 지수”, “경부장애지수”, “경부 장애지수”, “경부 장애 지수”, “경부기능장애지수”, “경부 기능장애지수”, “경부 기능 장애지수”, “경부 기능 장애 지수”

주요 검색어를 통해 선정된 논문의 본문 및 참고문헌을 확인하는 방식으로 연구를 수행하였다.

3단계: 문헌 선정(Study selection)

문헌 선정은 본 연구의 범위·자원 제약을 고려하여 주저자에 의한 주 검색과 제3자에 의한 검토로 2인이

시행하였다. 문헌 선정의 포함 및 제외 기준은 다음과 같다.

- 1) 포함 기준: (1) NDI를 한국어로 번역한 논문(1991년-2025년), (2) NDI를 실제 연구에 사용하여 설문을 수집, 분석한 논문(2015년-2025년), (3) 한국어에서 발표된 석·박사 학위논문 및 학술지, (4) 한국인을 대상으로 한국에서 시행된 연구
- 2) 제외 기준: (1) 같은 검색 용어를 사용하지만 다른 분야의 논문, (2) 설문지를 직접 사용하지 않은 체계적 고찰 등

4단계: 자료 추출 및 차트화(Charting the data)

최종으로 선정된 논문들은 자료 기입을 위해 Excel 365(Microsoft Corporation, Redmond, USA)를 이용하여 연도별로 학위논문 및 학술지별로 사전 설계한 양식에 중복을 제거하면서 정리하여 문제점을 각 항목 별로 정리하였다.

5단계: 결과 통합·요약 및 보고(Collating, summarizing and reporting the results)

1) 수치 요약 기술 분석(Descriptive numerical summary analysis): 포함된 연구의 특성을 설명하는 분석으로 한국어로 번역된 NDI의 종류와 실제 연구에 설문지의 사용을 검토하였다.

2) 주제별 질적 분석(Qualitative thematic analysis): 각 번역본별 NDI의 특성 및 설문지의 사용빈도, 채점방식 및 변형유무를 검토하였다.

6단계: 자문(Consultation)

예비 결과의 타당성과 방법론적 정합성을 점검하기 위하여 지도교수와 내부 자문을 실시하였다. 자문은 개념 정의, 포함·제외 기준, 데이터 차트화 항목 등을 중심으로 진행하였으며, 의견은 문헌 합성 결과에 반영하였다. 본 연구의 전체 검색 과정, 배제 기준, 스크리닝 결과를 포함한 엑셀 raw data를 OSF(Open Science Framework, 연구 사전등록과 자료·코드 공유를 지원하는 오픈 플랫폼)에 사전 등록된 프로토콜과 함께 모두 공개하였다. (DOI: <https://doi.org/10.17605/OSF.IO/JA4G8>).

III. 연구 결과

1. 수치 요약 기술 분석

국내 학술 데이터베이스에서 총 94,404편의 기록(records)을 확인하였다. 검색어 간 중복으로 인한 중복 레코드 1,977편을 사전 제거한 뒤, 2차 단계로 92,427편을 선별(screening)하였다. 이 가운데 DBpia에서 약어 모호성으로 기인한 약어-유발 오류(abbreviation-induced false) 87,502편을 제외하고 4,925편의 연구를 적격성 평가에 포함시켰다. 이 중 원저자 비공개 등의 전문 확보가 불가능한 3편과 엑셀 검색에서 중복된 4,367건을 제외하였다. 그 결과, 최종 포함 연구는 555편이었다. 아울러 동일한 기초 연구가 학위논문과 학술지 논문에 중복 발표된 경우 40건

을 하나의 연구 단위로 통합하여 산정하였으며, 이에 따라 최종 연구의 수는 515건으로 도출되었다. 연구의 문헌 선정과정은 아래의 Fig. 1. PRISMA flow diagram에 제시하였다.

NDI 초기 번역본을 추가적으로 확보하기 위한 보조 탐색으로, 1991 년(원저 최초 발표)부터 2015 년까지 RISS 에서 “Neck Disability Index” 및 “NDI”를 각각 검색하였고, 각 136 건 및 219 건 총 355 건을 확인하였다. 이 결과는 추가적인 탐색에 해당하므로 PRISMA flow diagram 에 포함하지 않았다. 각 한국어 NDI 번역본에 대해 보고된 통계치(예: ICC 모형과 신뢰구간, Cronbach's α , Rasch 분리 신뢰도)와 설문지 실재존재 여부를 요약하여 Table 1 에 제시하였다. 각 한국어판 NDI 관련 선행연구

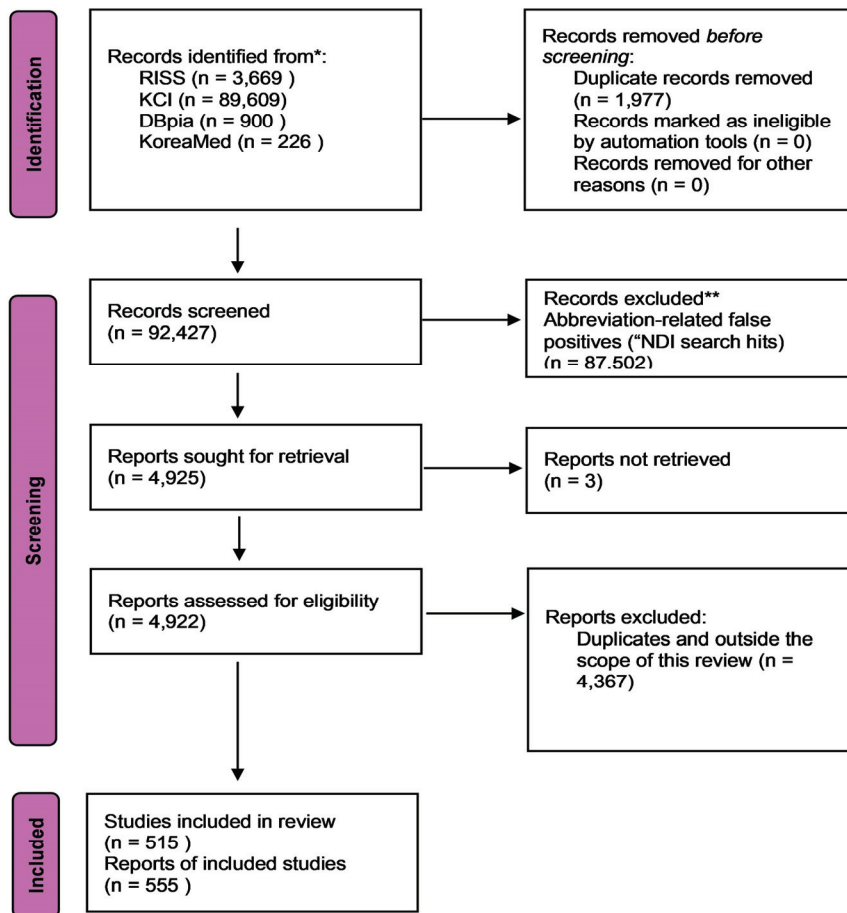


Fig. 1. PRISMA flow diagram.

에서는 일반 성인 자원자 40 명을 대상으로 한 검사·재검사 신뢰도를 평가하였으며[14,15,17], 비특이적/특이적 경부 통증 환자와 건강 대조군을 포함한 임상 연구에서 총 180 명의 환자(16~69 세)와 81 명의 건강한 성인(18~32 세)를 대상으로 내적 일관성, 판별 타당도를 평가하였다 [15,17]. 또한 병원 외래 경부 통증 대상자 71 명(평균 연령 39.7 세, 평균 유병기간 19.6 개월)을 대상으로 신뢰도와 타당도를 검증하였고[4], 대학생 경부 통증 대상자 71 명(평균 통증 기간 14.25 개월)을 대상으로 Rasch 분석을 하였다[16]. 마지막으로 경추 수술 환자 60 명과 78 명(평균 연령 약 56 세, 퇴행성 경추질환 및 신경병증/척수증 포함) 등을 통해 다양한 임상집단에서 각각 NDI 특성을 검증하였다[5, 18]. 이중 국내 논문은 Lim 등[14], Lee [15], Lee 등[4], Kim 등[16], Song 등[5]의 5 편, 국외 논문은 Lee 등[17], Song 등[18]의 2 편, 총 7 편으로 각 논문에서 공개

된 설문지를 모두 입수하여 기술된 통계치와 함께 설문지를 1:1 로 직접 비교 후 공통된 설문지를 제외하고 총 4 개의 설문지(Lee 등[17], Lee 등[4], Kim 등[16], Song 등[5])를 이후의 질적 주제별 분석의 기준으로 사용하였다.

또한 국내 발표된 Lim 등[14]과 국제 학술지에 발표된 Lee 등[17]은 동일 저자가 동일한 설문지를 사용하였으며, Lee [15]의 연구 역시 동일 연구자가 수행한 것으로 통계학적 결과 값을 고려할 때 NDI에 대한 새로운 번역 시도는 없었던 것으로 판단된다. 또한 이들 번역본들은 모두 1991년 초판(original version 1)을 기반으로 하고 있다.

한편 Song 등[5]이 국내 학술지에 발표한 한국어판 NDI는 설문 문항에서 모든 표현을 단순히 “통증(pain)”으로만 기술하고 있어, NDI 원 저자(Vernon)에 의한 초판 개정 사항(pain → neck pain)이 반영되지 않았다. 따라서 이 번역본 또한 1991년 초판을 기반으로 한 것으로 판단된다.

Table 1. Summary of the reported psychometric properties and questionnaire existence

No.	Year	Articles	Reported Reliability and Validity	Questionnaire availability
1	2005	Lim et al. [14]	ICC (2,1) = .900, (95% CI: .810 -.940)	Provided
2	2006	Lee et al. [17]	ICC (2,1) = .900, (95% CI: .810 -.940) Cronbach's α = .920 Effect size = 1.040, SRM = 1.170 r with GPE = -.040, (95% CI: -.570 to -.330) AUC = .790, (95% CI: .720 -.860)	Provided
3	2007	Lee [15]	ICC (2,1) = .900, (95% CI: .810 -.940) Cronbach's α = .920 Effect size = 1.040, SRM = 1.170 r with GPE = -.040, (95% CI: -.570 to -.330) AUC = .790, (95% CI: .720 -.860)	Not provided
4	2007	Lee et al. [4]	ICC (2,1) = .900, (95% CI: .850 -.950) Cronbach's α = .950 (p < .05) Criterion-related validity: r = .72 (p < .01)	Provided
5	2009	Kim et al. [16]	Item separation reliability: .970 Subject separation reliability: .850	Provided
6	2009	Song et al. [5]	ICC = .927 Cronbach's α = .820 NDI and VAS Correlation Coefficient r = .489 NDI and SF-36 Correlation Coefficient r = -.440	Provided
7	2010	Song et al. [18]	ICC = .930 Cronbach's α = .820 NDI and VAS Correlation Coefficient r = .490 NDI and SF-36 Correlation Coefficient r = -.440	Not provided

ICC: intraclass correlation coefficient, CI: confidence interval, SRM: standardized response mean, GPE: global perceived effect, AUC: area under the ROC curve, VAS: Visual Analog Scale, SF-36: 36-Item Short Form Health Survey

또한 같은 연구자가 환자 수를 확대하여 발표한 국제 논문 Song 등[18]은 Vernon으로부터 ‘neck pain’ 용어 변경의 권고를 받았다고 언급하였으나, 설문지를 공개하지 않아 실제 반영 여부는 확인할 수 없었다.

2. 주제별 질적 분석

1) 번역본별 NDI의 특성

본 연구는 NDI의 한국어 번역본을 ‘공식’과 ‘비공식’으로 구분하였다. 구분 기준은 저작권 및 라이선스 관리기관인 MRT의 공지 및 회신에 근거하였으며, PROMs 도구의 번역과 배포 및 연구 사용은 저작권의 사전 허가 없이는 허용되지 않으며, MRT에 등록 및 승인된 버전만을 공식 번역본으로 간주하였다. 이 분류 기준은 MRT와의 개인적 교신(2025-07-24)을 통해 재확인하였으며, 현재 한국어 NDI 공식 버전은 없는 상태이다.

대부분의 비공식 번역본은 MRT의 공식 번역 가이드라인(2012 제정, 2018 개정)[19,20]이 아닌 Beaton 등[8]의 방법론을 따랐다. 두 절차의 주요한 차이는 다음과 같으며, 번역본 간 번역 인원, 원 저자의 허가 등 상세 비교는 Table 2에 제시하였다.

MRT의 공식 절차(라이선스, 배포 일관성 중심): 로컬 코디네이터(Local Coordinators)와 2인의 순방향 번역

(Forward Translation) → 1인의 역방향 번역(Backward Translation) → 인지 면담(Cognitive Interview) → 교정 및 최종화(Proofreading & Finalization)

Beaton 방법(방법론 표준화 중심): 2인의 순방향 번역 → 2인의 역방향 번역 → 전문가 위원회 검토(Expert committee review) → 사전 검사(Prestesting)

2) 번역본별 NDI의 설문지의 사용 빈도 및 변형 검토
본 분석은 4개의 기준 번역본(Lee 등[17], Lee 등[4], Kim 등[16], Song 등[5])을 레퍼런스 삼아, 학위논문 및 학술지에서 실제 사용된 NDI 설문지를 수집, 검토하였다. 각 논문의 부록 제공 여부를 확인하고, 제공된 설문지는 4개의 기준 번역본과 1:1로 대조하여 서문, 지시문, 문항 등의 변형 여부를 평가하였다. 또한 연구의 방법과 참고문헌을 검토하여 번역본별 사용 빈도를 분석하였다. 설문지 대조 결과, 기준 번역본을 레퍼런스로 인용하면서 서문의 변경·문항의 삭제·추가·단어 변경 등으로 내용이 변형된 설문지는 수정본(modified version)으로 분류하였다. 반대로 네 개의 기준 번역본 가운데 어느 것에도 일치하지 않거나, 설문지의 출처가 명시되지 않아 사용한 번역본을 특정할 수 없는 경우에는 모두 불명본(unknown version)으로 처리하였다.

Table 2. Comparison of the translation procedures for Korean versions of the NDI

Classification	Forward translation	Backward translation	Cognitive Debriefing	Proofreading & Finalization	Notes
MRT Standard Procedure [19,20]	Two Professional translators (native Korean speakers)	One professional translator (native English speaker)	Required (patient interviews)	Bilingual proofreader (native Korean speaker)	MRT international standard
Lee et al. [17]	Six Professional translators (native Korean speakers)	Four translators (native English speakers)	Conducted along with test-retest reliability	Reviewed by language professionals with an expert committee	Translation based on the early 1991 version
Lee et al. [4]	Three Professional translators (native Korean speakers)	Two professional translators (native English speakers)	Not performed	Reviewed by language professionals with an expert committee	
Kim et al. [16]	Not conducted	Not conducted	Not conducted	Modified directly by the authors	Based on a textbook version
Song et al. [5]	Two Professional translators (native Korean speakers)	Two professional translators (native English speakers)	Conducted with 30 patients	Reviewed by language professionals with an expert committee	Verified with Author H. Vernon, Items 8 and 10 modified after patient interviews

특히 Lee [15], Song 등[18]을 번역본 레퍼런스로 인용한 연구들은 다음과 같은 이유로 불명본으로 포함하였다. 두 논문 자체에는 한국어판 NDI 설문지가 부록으로 제공되지 않으며, 이를 인용한 후속 연구들 역시 해당 연구에서 실제로 사용한 설문지를 부록으로 제공하지 않거나, 수집하여 대조한 네 개 기준 번역본과 다른 문항 구성을 보였다. 이와 같이 기준 번역본에 대한 명확한 레퍼런스 없이 Lee [15], Song 등[18]만을 인용하거나, 기준 번역본과 일치하지 않는 설문지를 사용한 경우에는 모두 불명본으로 분류하였다.

4개의 기준 번역본별 사용빈도는 Table 3에 제시하였다. 이 중 설문지의 수정본은 58건(11.26%)이었고, 불명본은 298건(57.86%)이었다. 각 버전의 연도별 사용 변화는 백분율로 정규화 하여 Fig. 2에 도식화하였다.

부록 등의 형태로 설문지를 첨부한 학술지 논문은 4편, 학위논문은 154편으로 전체 논문 30%가 저작권자의 허가 없이 설문지를 직접 논문에 첨부하였다. 저자에 의해 제공되는 매뉴얼에 포함된 미응답 시 채점 방식을 언급한 논문은 전체 논문 중 단 13편(2.5%)이었다. 이와 반대로 점수 산정기준을 임의로 변경하는 등 척도를 변경한 연구는 20편(3.9%)이었다.

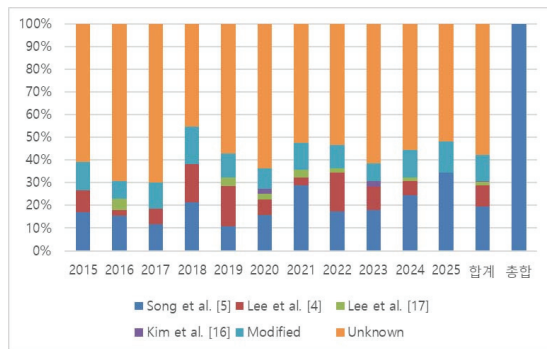


Fig. 2. Percentage of studies by the translation version.

Table 3. Counts and percentages of studies by translation version

Total	Lee et al. [17]	Lee et al. [4]	Kim et al. [16]	Song et al. [5]	Modified version	Unknown version
N = 515	9	48	2	100	58	298
(100%)	(1.75%)	(9.32%)	(.39%)	(19.42%)	(11.26%)	(57.86%)

IV. 고 찰

본 스코핑 리뷰는 국내에서 사용된 한국어판 NDI가 번역 과정, 설문지 첨부 관행, 점수 산정 방식 등에서 여러 측정학적 및 구조적 한계를 지니고 있음을 보여준다. 우선, 저작권 관리 기관 MRT는 시험용 사본(evaluation copy)의 사용 목적을 설문지의 후속 연구 또는 임상적 사용 적합성을 검토하는 제한된 범위로 한정하고, 사전 서면 허가 없이 연구 및 임상 사용, 제3자 배포, 수정, 재입력, 번역, 복제 등을 금지하고 있음을 명시한다[3]. 실제로 천식 삶의 질 설문지(Asthma Quality of Life Questionnaire; AQLQ)의 그리스어판 사례에서는 저작권자의 승인 없이 문항 수를 축소한 비공식 번역본이 공개되었다가 학문적, 임상적 사용이 금지되었고 [6], 환자의 약물 순응도를 검사하기 위한 MMAS(Morisky Medication Adherence Scale)을 사용한 여러 연구가 저작권자의 허가 없이 번역 및 응용됨에 따라 PLOS ONE 저널에서 2023년 해당 논문이 철회된 사례 [21], 그리고 이러한 저작권 이슈로 인한 다수의 철회된 사례를 보도한 Retraction Watch의 Avery Orral의 보고[22]는, 저작권자의 승인 없는 PROMs 사용이 연구의 지속 가능성 자체를 위협할 수 있음을 뒷받침한다.

둘째, NDI의 개발 모체인 ODI의 저자 Fairbank는 항목의 임의 수정 및 변형은 도구 간 점수의 비교가능성을 심각하게 훼손한다고 경고하였는데, 그 근거로 Rasch 분석 관점에서 동일 구성개념을 측정하지 못하는 것으로 드러나 점수가 유의하게 달라져 연구 간 비교를 무력화하게 될 수 있으므로, 근거 없는 변형본(rogue version) 사용은 배제되어야 한다고 주장했다[7]. NDI 역시 Rasch 분석 기반으로 한 검토에서 일차원성 부족·큰 바닥효과·원점수 간의 등간성 결여 등 측정 안정성 문제가 제기되었고, 따라서 문항·채점 방식이 조금만 달라져도 점수 해석이 달라질 여지가 큰 도구이다[23]. 또한, 검증된 Rasch 기반

수정판들(NDI-8, NDI-5)조차도 원본 NDI-10과 비교 시 동일 표본에서 평균 약 10% 내외의 점수 차이를 보이며 상호 대체 사용이 불가능한 것으로 확인되었다[24]. 이는 공식적으로 검증된 단축판 사이에서도 유의한 점수 불일치가 발생할 수 있음을 시사하며, 미검증 및 임의 변형된 번역본에서는 더 큰 편향이 발생할 개연성을 높인다.

셋째, PROMs 설문지는 단어 선택·응답 형식·회상 기간·서식 요소 등 세부 구성까지 타당도와 신뢰도를 확보하도록 설계된 정밀 측정 도구이다. 따라서 승인 및 검증 절차 없이 문항이나 형식을 변경하는 것은 측정 속성을 훼손하여 결과 해석을 왜곡할 수 있다[25]. 이러한 위험을 줄이기 위해서는 저작권이 있는 설문지를 직접 제작하여 쓰지 않고, 저작권자 및 관리기관(MRT)의 사전 허가와 공인 번역 절차 준수, 언어 및 문화 등가성에 대한 체계적 번역과정이 필수적이다.

넷째, MRT가 공식 제공하는 원저자판 NDI 매뉴얼은 미응답 발생 시 점수 산정 규칙(예: 비례 환산 등)을 명시하고 있으며, 원저자와 협업한 비공식 번역 개발 과정에서도 7번 일(Work), 8번 운전(Driving)에서의 높은 미응답 비율이 관찰되었고 이에 대한 대응 산정 방식이 공개된 바 있다[2, 5]. 그럼에도 불구하고 본 스코핑 리뷰에서 확인한 515편 중 단 13편(2.5%)만이 미응답 처리 규칙을 명시적으로 보고하거나 매뉴얼과 합치되는 방식으로 산정하였다. NDI가 10문항(총점 50점) 체계임에도 다수의 연구에서 미응답 처리 기준의 불명확한 운영이 확인되었고, Vernon의 매뉴얼에 비추어 계산할 경우 최대 약 10점의 점수 손실이 발생할 수 있다. 이는 실제로 효과가 있는 중재를 무효화로 오판하게 만들고, 누적 근거의 신뢰도를 체계적으로 약화시킨다. 이와 같이 낮은 보고 처리 수준은 연구간 점수의 동등성(equivalence)과 비교가능성(comparability)을 저해하고, 결과적으로 측정 신뢰성 및 해석 타당성을 약화시킬 수 있다.

다섯째, Table 1에 제시된 선행연구를 종합해 보면, 한국어판 NDI 번역본들은 번역 및 문화적 적응 절차의 보고 수준, 신뢰도·타당도 지표(ICC, Cronbach's α , effect size, SRM 등)의 산출·인용 방식, 설문지 제공 여부에서 공통적인 한계와 문제점을 보인다. 일반적으로 연구자는 기존 연구에서 보고된 신뢰도와 타당도 결과를 근거로 도구의

적합성을 판단하고 인용하는데, 동일한 번역본과 동일한 표본에서 산출된 결과값이 여러 편의 논문으로 반복 출판되면, 독립된 검증 연구가 많은 것처럼 보이면서 한국어판 NDI의 근거 수준과 검증 연구 수를 실제보다 과대평가하게 만들 수 있다. 설문지를 논문 부록으로 직접 제공하는 관행 역시 후속 연구자의 접근성을 높이는 장점이 있는 반면, 이를 바탕으로 연구자가 설문지를 다시 제작하는 과정에서 문항 표현, 문항 수, 채점 방식에 미세한 변형이 누적되어 장기적으로 동일 도구에 대한 비표준화된 여러 버전을 양산하는 결과를 초래할 수 있다. 이러한 문제를 줄이기 위해서는 향후 한국어판 NDI 연구에서 번역 및 검증 절차, 사용한 번역본의 버전과 출처, 통계치의 직접 산출 여부와 인용 여부, 설문지 제공 방식 등을 일관되게 보고하도록 학술지·학회 차원의 가이드라인을 마련하고, 장기적으로는 공식 한국어판 번역과 기관에 의한 일원화된 배포 체계를 구축할 필요가 있다.

한편, 본 연구에서 번역본 분류 결과 불명본이 가장 높은 비율을 차지한 것은 이러한 구조적 한계를 반영하는 결과로 해석할 수 있다. 우선 많은 국내 연구에서 기존번역본이 되는 논문을 직접확인하고 그 안에 부록으로 제공되는 설문지를 사용하는 것이 아닌, 선행연구의 참고문헌을 재인용하는 방식으로 설문지를 확보하고 사용하는 관행으로 인해 사용 번역본의 출처와 버전이 명시되지 않은 경우가 많았다. 실제로 본 분석 과정에서 특정 설문지에서 발견된 오타가 다양한 연구의 설문지에서도 반복되는 사례가 관찰되어, 설문지가 원출처가 아닌 2차 자료를 통해 복제 및 전파되었음을 시사한다. 더불어 한국어판 NDI 번역본이 여러 개 병존하는 상황에서도 연구자들이 사용 중인 설문지가 공식 번역본인지 비공식 및 임의 번역본인지 여부를 적극적으로 확인하지 않아, 결과적으로 출처가 불명확한 설문지가 광범위하게 사용된 점 역시 불명본 비율을 높인 주요 요인으로 판단된다.

종합하면, 비공식적 번역과 출판 그리고 이러한 출판물의 2차적 사용에 의한 변형본의 혼용 및 검증되지 않은 출처 불명의 평가도구의 사용은 저작권·윤리 문제와 측정학적 불확실성을 동시에 증폭시킨다. 이러한 문제는 MRT가 관리하는 공식 번역 체계가 국내에서 충분히 채택 및 준수되지 않았고, 공식 한국어판이 부재한 것에 있다.

이에 따라 MRT 절차에 부합하는 공식 한국어판의 마련과 그 단일 사용이 요구된다.

본 연구는 한국인 대상 연구의 체계적 검색을 위해 국내 자료로 범위를 제한하였으며, 이로 인해 해외 학술지에 게재되었으나 한국인 자료를 포함한 일부 연구가 누락되었을 가능성이 있다. 또한 스코핑 리뷰에서는 정식 학술 출판 경로에 포함되지 않는 비정규 자료(예: 학술 대회 발표)를 의미하는 회색 문헌(grey literature)을 함께 탐색하는 것이 권고되나, 본 연구에서는 포함하지 못하였다. 더불어 부록이 없는 경우 또는 출처 불명 연구의 경우 해당 설문 원문을 연구자에게 직접 입수하지 못했다는 한계를 가진다. 또한 스코핑 리뷰는 원칙적으로 최소 2인이 독립적으로 문헌 검색과 선정을 수행해야 하나, 본 연구는 연구 범위와 자원 제약을 고려하여 주저자가 주 검색을 수행하고 제3자가 이를 검토하는 방식으로 2인 검토 절차를 적용하였다. 이에 따라 문헌 선정과정에서 일부 선택 편향(selection bias)이 존재할 가능성이 있다.

V. 결론

본 스코핑 리뷰를 통해 국내에서 공식 한국어판 NDI가 부재한 상황에서 다양한 비공식 번역본과 이를 2차적으로 변형한 변형본, 더 나아가 출처 불명의 평가도구가 'NDI'라는 동일 명칭 아래 광범위하게 사용되어 왔음을 규명하였다. 이와 같은 혼용은 첫째, 문항·채점 규칙·미응답 처리 방식의 불일치로 인해 연구 간 점수의 비교가능성과 측정 신뢰성을 저하시킬 뿐 아니라, 둘째, 도구 사용에 관한 권리와 책임의 경계를 불명확하게 함으로써, 해당 연구의 학술 기록으로서의 지속 가능성과 후속 검증 연구에서의 활용 가능성을 불투명하게 만든다. 나아가 비공식 번역본의 지속적인 사용은 측정 결과의 안정성을 저해하며, 이를 토대로 수행되는 체계적 문헌고찰과 메타분석의 결론에도 부정적 영향을 미친다.

이에 따라 본 연구를 통해 다음과 같은 제도적 및 편집정책적 개선을 제안한다. 첫째, MRT 절차에 부합하는 공식 한국어판 NDI의 조속한 마련 및 단일 표준 채택이 선행되어야 한다. 둘째, 연구자, 의학저널, IRB 차원에서 저작권

과 라이선스 의무화 및 설문지 원문 부록 게재 제한, 혹은 라이선스 근거 제시 의무 등의 투명 보고를 핵심 항목으로 정례할 것을 제안하는 바이다.

References

- [1] Churruca K, Pomare C, Ellis LA, et al. Patient-reported outcome measures (PROMs): a review of generic and condition-specific measures and a discussion of trends and issues. *Heal Expect.* 2021;24(4):1015-24.
- [2] Vernon H. The neck disability index: state-of-art, 1991-2008. *J Manipulative Physiol Ther.* 2008;31(7):491-502.
- [3] Mapi Research Trust. Neck Disability Index - Mapi ePROVIDE page [Internet]. 2025. Available at <https://eprovide.mapi-trust.org/instruments/neck-disability-index>
- [4] Lee EW, Shin WS, Jung KS, et al. Reliability and validity of the neck disability index in neck pain patients. *Korean Res Soc Phys Ther.* 2007;14(3):97-103.
- [5] Song K, Choi B, Kim S, et al. Cross-cultural adaptation and validation of the korean version of the neck disability index. *J Korean Orthop Assoc.* 2009;44(3):350-9.
- [6] Anfray C, Arnold B, Martin M, et al. Reflection paper on copyright, patient-reported outcome instruments and their translations. *Health Qual Life Outcomes.* 2018;16(1):224.
- [7] Fairbank JCT. Why are there different versions of the oswestry disability index? a review. *J Neurosurg Spine.* 2014;20(1):83-6.
- [8] Beaton DE, Bombardier C, Guillemin F, et al. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine (Phila Pa 1976).* 2000;25(24):3186-91.
- [9] Arksey H, O'Malley L. Scoping studies: towards a

- methodological framework. *Int J Soc Res Methodol.* 2005;8(1):19–32.
- [10] Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. *Implement Sci.* 2010;5(1):69.
- [11] Peters MDJ, Godfrey CM, Khalil H, et al. Guidance for conducting systematic scoping reviews. *Int J Evid Based Healthc.* 2015;13(3):141–6.
- [12] Tricco AC, Lillie E, Zarin W, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Ann Intern Med.* 2018;169(7):467–73.
- [13] Al-Ebrahim SQ, Harrison J, Chen TF, et al. Cross-cultural adaptation and psychometric properties of patient-reported outcome measures in Arabic speaking countries: a scoping review. *Res Soc Adm Pharm.* 2023;19(7):989–1006.
- [14] Lim CH, Lee HJ, Bae SS, et al. The cross cultural adaptation for neck disability questionnaires. *J Korean Soc Phys Ther.* 2005;17(2):46–66.
- [15] Lee HJ. Cross-cultural adaptation of korean language version of neck pain and disability questionnaires and their psychometric testing. *Korean J Meridian Acupoint.* 2007;24(2):99–112.
- [16] Kim TH, Kim JH, Gong WT. Rasch analysis to neck disability index with neck pain subjects. *J Korean Soc Phys Ther.* 2009;21(3):1–8.
- [17] Lee H, Nicholson LL, Adams RD, et al. Development and psychometric testing of korean language versions of 4 neck pain and disability questionnaires. *Spine (Phila Pa 1976).* 2006;31(16):1841–5.
- [18] Song KJ, Choi BW, Choi BR, et al. Cross-cultural adaptation and validation of the korean version of the neck disability index. *Spine (Phila Pa 1976).* 2010;35(20):E1045–9.
- [19] Acquadro C, Conway K, Giroulet C, et al. Linguistic validation manual for health outcome assessments. Lyon. Mapi Institute. 2012.
- [20] Mapi Research Trust. Linguistic validation guidance of a clinical outcome assessment (COA). Lyon. Mapi Research Trust. 2018.
- [21] Yazie TS, Mengistu WE, Asmare Z, et al. Retraction: medication regimen complexity and its association with adherence and blood pressure control among hypertensive patients at selected hospitals of south gondar zone: a hospital based cross sectional study. *PLoS One.* 2023;18(6):e0282464.
- [22] Orrall A. Papers continue to face retractions for failure to license pricy tool. *Retraction Watch.* 2025. Available at: <https://retractionwatch.com/2025/09/02/papers-continue-to-face-retractions-for-failure-to-license-pricy-tool/>. Accessed March 3, 2026.
- [23] Hung M, Cheng C, Hon SD, et al. Challenging the norm: further psychometric investigation of the neck disability index. *Spine J.* 2015;15(11):2440–5.
- [24] Lu Z, MacDermid JC, Nazari G. Agreement between original and rasch-approved neck disability index. *BMC Med Res Methodol.* 2020;20(1):180.
- [25] Juniper EF. Validated questionnaires should not be modified. *Eur Respir J.* 2009;34(5):1015–7.