



대한물리치료과학회지

Journal of Korean Physical Therapy Science
2025. 09. Vol. 32, No.3 pp. 1-17

물리치료 평가도구의 분류체계와 임상 적용성 분석 : 국가참조표준 수립을 위한 기초연구

김정현¹ · 이병희²

신한대학교 통합대체의학과¹ · 삼육대학교 물리치료학과²

Classification and Clinical Applicability of Physical Therapy Assessment Tools : A Foundational Study for Establishing National Reference Standards

Jung-Hyun Kim¹, Ph.D., P.T.. · Byoung-hee Lee², Ph.D., P.T.

¹Department of Integrated Alternative Medicine, Shinhan University, Korea

²Department of Physical Therapy, Sahmyook University, Korea

Abstract

Background: Physical therapy assessment tools play a critical role in quantifying function and shaping rehabilitation strategies. However, their application in Korea has been inconsistent, primarily due to fragmented classification schemes and underdeveloped links to electronic medical records (EMRs) and reimbursement systems. This inconsistency reduces both clinical utility and policy relevance.

Design: Integrative systematic review.

Methods: Nine major tools commonly used in Korean rehabilitation settings were reviewed: range of motion (ROM), manual muscle testing (MMT), the Berg Balance Scale (BBS), the 10-Meter Walk Test (10MWT), the 6-Minute Walk Test (6MWT), activities of daily living (ADLs), instrumental activities of daily living (IADLs), the Functional Independence Measure (FIM) and the Barthel Index/Modified Barthel Index (BI/MBI). The functional scope of each tool, the patient populations it covers, its alignment with the International Classification of Functioning (ICF) domains and SNOMED CT terminology, and its psychometric indices (e.g. Standard Error of Measurement (SEM), Minimal Clinically Important Difference (MCID) and Minimal Detectable Change (MDC)) were examined.

Results: Distinctive strengths were observed. The BBS and 10MWT yielded sensitive thresholds that were particularly relevant to stroke survivors and older adults. By contrast, BI and

FIM offered broader insights into independence across daily living activities. The IADL demonstrated added value in detecting functional decline linked to cognitive impairment and dementia. However, reported SEM, MDC and MCID values varied across studies and patient cohorts, indicating that interpretation must be contextualised. From an administrative standpoint, the BBS, 10MWT, IADL and BI had reimbursement codes in place, whereas the FIM and 6MWT lacked such linkage – an omission that restricts their integration into EMR-driven systems and insurance frameworks.

Conclusion: These findings emphasise the urgent need to develop a standardised framework for physical therapy assessment in Korea. The development of the Korean Physical Therapy Toolkit (KPT) would harmonise clinical decision-making, strengthen reimbursement alignment, and promote interoperability with international standards such as the International Classification of Functioning, Disability and Health (ICF) and the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED CT).

Key words: ICF, Physical Therapy, SNOMED CT, Standardization, Tools

교신저자

이병희

서울시 노원구 화랑로 815, 삼육대학교 제3과학관 112호

T: 02-3399-1634, E: 3679@syu.ac.kr

I. 서론

재활은 단순히 손상된 신체 기능의 회복에 머무르지 않고, 환자의 삶의 질을 높이고 사회적 참여를 촉진하는 다학제적 접근을 필요로 한다. 이러한 과정에서 물리치료는 치료 효과를 객관적으로 평가하고 경과를 추적하는 핵심적 역할을 수행하며, 이를 위해 다양한 임상 평가도구가 임상 현장에서 활용되고 있다 (Üstün 등, 2003; Leonardi 등, 2022).

그러나 국내 임상에서는 평가도구 선택과 활용이 일관성을 확보하지 못한 채 이루어지고 있으며, 표준화된 분류체계와 적용 지침의 부재가 지속적으로 지적되고 있다. 이로 인해 전자의무기록(Electronic Medical Record, EMR) 시스템이나 건강보험 청구, 치료 효과의 객관적 분석 과정에서 정보가 분절적으로 축적되며, 장기적으로는 의료데이터 기반의 인공지능 분석과 정책 설계에도 제약이 발생한다 (Maritz 등, 2017; Reed 등, 2005).

세계보건기구(World Health Organization: WHO)가 제시한 ICF(International Classification of Functioning, Disability and Health)는 기능(Function), 활동(Activity), 참여(Participation)을 연속적인 개념으로 파악하고 있으며, 이를 기반으로 한 평가도구의 표준화와 국제 연계의 중요성이 강조되고 있다 (Üstün 등, 2003; Cieza 등, 2009). 이러한 흐름에 따라 국내에서도 물리치료 평가도구의 분류체계를 정립하고, 이를 토대로 국가참조표준(National Reference Standard)을 수립하는 작업이 시급하다.

또한 건강보험 수가체계는 단순한 의료비 보상 수단을 넘어 평가도구의 활용도를 높이는 중요한 장치로 작동한다. 일부 도구는 이미 수가코드가 부여되어 있으며, 이는 임상에서의 공식적 기록과 활용을 촉진한다. 예컨대 버그 균형 척도(Berg Balance Scale: BBS)는 ‘외과·근골 기능 검사(E6613)’, 10미터 보행 검사(10-Meter Walk Test, 10MWT)는 ‘치료 및 검사(52050)’ 등의 코드로 적용되고 있으며 (Leonardi 등, 2022), 이러한 수가 부여는 도구의 임상적 유용성을 인정받았음을 의미한다. 더 나아가 수가코드와 EMR의 연계는 평가도구의 사용 현황을 추적·분석하는 기반이 되어, 데이터 기반 정책 수립에도 기여할 수 있다 (Maritz 등, 2017).

평가도구는 특정 기능 영역을 측정하는 목적을 가지며, 기능 중심의 분류는 치료계획 수립, 환자 상태 모니터링, 다학제 간 소통을 위한 지표로 활용된다. 예를 들어 BBS는 균형 능력을, 10MWT는 보행 속도를, 도수근력검사(Manual Muscle Test, MMT)는 근력을, 관절가동범위검사(Goniometry)는 관절 움직임 범위를 측정한다. 또한 기능적 독립 척도(Functional Independence Measure: FIM)와 수정 바텔 지수(Modified Barthel Index: MBI)는 일상생활기능(ADL)과 도구적 일상생활기능(IADL)을 평가하는 대표적 도구이다 (Üstün 등, 2003; Lawton과 Brody, 1969; Keith 등, 1987). 그러나 국내에서는 이러한 기능적 분류체계가 아직 정립되지 않아 도구 선택과 활용의 혼선이 발생하고 있다 (Cieza 등, 2009). 더 나아가 최근 연구에서는 공공 물리치료 인프라 구축이 지역 사회 건강 증진과 비용 효율성에 유의한 효과를 미치는 것으로 보고되어, 물리치료 평가도구의 표준화가 임상적 차원을 넘어 정책적·재정적 파급효과와도 직결됨을 보여준다(김종엽과 김정현, 2025). 또한 디지털 치료제가 뇌성마비 아동의 균형 및 보행 능력 향상에 긍정적 효과를 보였다는 결과가 제시되면서, 최신 디지털 중재 효과를 평가도구 적용성 검토에 반영해야 할 필요성이 강조되고 있다(김정현, 2023).

ICF 기반의 구조화는 평가도구 적용의 일관성을 높이는 중요한 방법론이다. 평가도구를 ICF의 세 가지 도메인(Activity, Participation, Body Function)에 따라 분류하면 도구 간 비교 가능성이 강화되고, 중복성 여부를 확인할 수 있으며, 선택 과정의 일관성이 확보된다 (Üstün 등, 2003). 이는 다학제 팀 간의 협력을 촉진하고 환자의 전반적 기능 상태를 통합적으로 파악하는 데에도 유용하다. 하지만 국내에서는 ICF 도메인을 기준으로 한 체계적 구조화가 부족하여, 임상 현장에서의 활용이 제한되고 있다.

더불어 SNOMED CT(Systematized Nomenclature of Medicine - Clinical Terms)는 국제적으로 통용되는 임상 표준 용어 체계로, EMR 상호운용성과 데이터 표준화를 위한 핵심 기반이 된다 (Maritz et al., 2017). 예를 들어 BBS는 SNOMED CT에서 ‘Balance assessment (procedure)’, MBI는 ‘Assessment of activities of daily living (procedure)’로 분류되어 있으며, 이는 표준화된 정의를 통해 시스템 간 데이터 교환을 용이하게 한다. 그러나 국내에서는 SNOMED CT 기반 코드화가 미흡하여 국제적 호환성과 상호운용성 확보에 제약이 따른다. 따라서 물리치료 평가도구의 SNOMED CT 매핑을 강화하고, 이를 기반으로 한 국가 차원의 표준화 작업이 필요하다.

본 연구는 이러한 문제의식을 바탕으로, 국내 건강보험 수가가 부여된 7개 평가도구(관절가동범위검사, 도수근력검사, 버그 균형 검사, 10미터 보행 검사, 일상생활동작검사, 도구적 일상생활동작검사, 바텔 지수)를 중심으로 분석하였다. 아울러 수가코드가 부재한 2개 도구(기능적 독립 측정도구, 6분 보행 검사)와 SNOMED CT 코드가 존재하지 않는 10MWT를 포함하여, 기능 영역, 적용 환자군, 평가 목적, ICF 도메인 연계, SNOMED CT 코드 유무, 신뢰도 및 타당도 지표를 종합적으로 검토하였다. 이를 통해 물리치료 평가도구의 체계적 분류 기준을 제시하고, 국가참조표준 (Korean Physical Therapy Toolkit: KPT) 수립을 위한 기초자료를 제공하고자 한다.

II. 연구방법

1. 연구 설계 및 문헌 수집 절차

본 연구는 물리치료 영역에서 사용되는 주요 임상 평가도구의 구조적 특성과 임상 적용성을 종합적으로 분석하고, 이를 바탕으로 국가 차원의 표준 도구체계(Korean Physical Therapy Toolkit: KPT) 수립을 위한 기초자료를 마련하고자 수행되었다. 연구 설계는 체계적 문헌고찰(Systematic Review)의 원칙을 따르되, 고전적 근거 논문과 최신 연구를 모두 포함하는 통합적 문헌고찰(Integrative Review) 접근법을 적용하였다. 전 과정은 체계적 문헌고찰 및 메타분석을 위한 권장지침(Preferred Reporting Items for Systematic Reviews and Meta-Analyses: PRISMA) 2020을 참조하여 진행하였다.

문헌 검색은 1965년 Barthel Index 원저 발표 시점부터 2024년 12월까지 발표된 국내외 문헌을 대상으로 하였다. PubMed, Scopus, Web of Science, CINAHL, Google Scholar 등 주요 국제 학술 데이터베이스를 활용하였으며, 국내 문헌 확보를 위해 학술연구정보서비스(Research Information Sharing Service: RISS), 한국학술지인용색인(Korea Citation Index: KCI), DBpia를 병행하였다. 검색어는 영문으로는 ‘physical therapy’, ‘assessment tool’, ‘scale’, ‘index’, ‘validity’, ‘reliability’, ‘ICF’, ‘SNOMED CT’, ‘standardization’을 사용하였으며, 국문으로는 “물리치료 평가도구(physical therapy assessment tool)”, “신뢰도(reliability)”, “타당도(validity)”, “ICF(국제 기능·장애·건강 분류)”, “수가코드(reimbursement code)” 등을 활용하였다.

문헌 선정 기준은 다음과 같다. 첫째, 임상 현장에서 실제 사용되는 물리치료 평가도구를 대상으로 한 연구일 것. 둘째, 평가도구의 구조, 기능 영역, 적용 대상군, 신뢰도 및 타당도 관련 지표 또는 비교자료가 보고된 문헌일 것. 셋째, 국제 표준 분류체계인 ICF(International Classification of Functioning, Disability and Health) 혹은 SNOMED CT(Systematized Nomenclature of Medicine - Clinical Terms)와의 연계 가능성이 언급되었거나 이를 활용한 연구일 것. 또한 한국형 평가도구의 개발과 검증을 다룬 논문을 적극적으로 포함하였다.

반면, 평가도구가 명확히 정의되지 않았거나 비정형적 측정 기준을 적용한 경우, 단일 증례연구(case report), 서신(letter), 개인 의견서(opinion paper) 등은 제외하였다. 또한 물리치료와 직접적인 관련성이 낮은 순수 인지기능 검사나 간호중심 지표를 다룬 문헌 역시 분석 대상에서 배제하였다.

문헌 검색과 선별 과정은 두 명의 연구자가 독립적으로 수행하였다. 제목과 초록을 기반으로 1차 선별을 진행하고,

원문 검토를 통해 선정 기준에 부합하지 않는 문헌을 2차로 배제하였다. 의견 불일치가 발생한 경우에는 제3의 연구자가 개입하여 최종 합의를 도출하였다. 최종적으로 포함된 문헌은 평가도구의 기능 영역, 적용 대상, 신뢰도 및 타당도, 수가 및 코드 매핑 여부 등을 중심으로 체계적으로 정리하였으며, 결과는 서술적 합성(narrative synthesis)과 표 기반 비교분석을 통해 제시하였다. 더불어 본 연구에서 참고한 주요 임상 평가도구의 특성과 지표는 Shirley Ryan AbilityLab Rehabilitation Measures Database (RMD) 에서 확인하였다(Shirley Ryan AbilityLab, 2021).

2. 분석틀 구성 및 임상 적용성 분석

선정된 평가도구는 구조화된 분석틀에 따라 항목별로 정리되었으며, 분석은 도구의 기능적 속성과 제도적 활용 가능성을 아우르는 다차원적 접근으로 이루어졌다. 기능적 분류는 평가 영역(Function Domain), 평가 목적(Evaluation Purpose), ICF 도메인 연계(ICF Mapping), SNOMED CT 코드 매핑(SCTID Mapping)이라는 네 가지 기준에 따라 수행되었다. 기능 영역은 근력, 균형, 이동성, 관절가동범위, 일상생활동작(ADL), 도구적 일상생활동작(IADL) 등으로 범주화하였고, 평가 목적은 기능 수준 측정, 치료 목표 설정, 경과 모니터링, 치료 반응성 평가 등으로 구분하였다. 또한 ICF 도메인 연계는 WHO에서 제시한 활동(Activity), 참여(Participation), 신체기능(Body Function)의 기준에 따라 도구의 위치를 지정하였으며, SNOMED CT 매핑은 각 평가도구가 국제 임상용어체계 내 procedure 또는 assessment scale로 등재되어 있는지 여부 및 SCTID 코드 존재 여부를 검토하였다.

이와 함께 임상 적용성 분석(Clinical Applicability Analysis)은 도구가 실제 진료 및 행정체계에서 갖는 활용 가치를 평가하는 데 중점을 두었다. 이를 위해 네 가지 항목을 기준으로 정성적 비교분석을 수행하였다. 첫째, 임상 적용 가능성(Clinical Utility)은 물리치료 현장에서의 사용 빈도, 환자군별 적합성, 국내 임상지침 및 국가재활기관 적용 사례를 중심으로 검토하였다. 둘째, 행정적 수용성(Administrative Acceptability)은 건강보험 수가코드의 존재 여부, 전자의무기록(Electronic Medical Record, EMR) 내 등록 가능성, 정책적 활용 잠재성을 기준으로 분석하였다. 셋째, 표준 연계성(Standard Linkage)은 ICF 및 SNOMED CT와의 매핑 가능성, 국제 분류 도메인과의 정합성, WHO 또는 국제 물리치료학회 지침 포함 여부 등을 확인하였다. 넷째, 보편성(Generality)은 신경계, 근골격계, 노인성 질환 등 다양한 환자군과 입원·외래·지역사회 재활 등 다양한 환경에서의 적용 유연성과 범용성을 중심으로 평가하였다.

이러한 분석틀은 각 평가도구의 기능적 특성, 임상적 유효성, 제도적 활용성, 국제 표준화 연계 가능성을 종합적으로 검토하기 위한 기준으로 활용되었다. 결과는 표 기반 비교분석과 서술적 합성을 병행하여 제시되었으며, 도구 간의 유사성과 차별성, 적용 환경의 확장성, 코드화 가능성 등을 함께 논의하였다. 이러한 접근은 단순한 도구별 특성의 나열을 넘어, 한국형 국가참조표준 도구 후보군(KPT Candidate Set)을 도출하기 위한 기초 근거를 마련하는 데 기여하도록 설계되었으며, 국가참조표준 도구 후보군으로의 적합성을 검토하는 데 활용될 수 있도록 구성하였다.

3. 평가도구의 정성적 분석

선정된 9개 평가도구에 대해서는 단순한 계량적 수치 분석을 넘어, 도구의 구조적 특성과 임상적 활용성을 종합적으로 고찰하였다. 이를 위해 각 도구의 기본 정보(개발 배경, 평가 항목 구성, 사용 환경), 신뢰도 및 타당도 지표, 임상 적용 환경(입원·외래·지역사회 재활), 평가 수행자의 전문성 요건 및 활용의 용이성, 무료 사용 가능 여부 등을 포함하여 정리하였다.

특히, 임상적 의사결정에 활용 가능한 민감도 지표로서 최소 임상 중요 변화 (Minimal Clinically Important Difference: MCID), 최소 감지 변화량 (Minimal Detectable Change: MDC), 표준측정오차 (Standard Error of Measurement: SEM)를 핵심적으로 분석하였다. 이러한 지표들은 환자 기능 변화가 단순한 통계적 차이를 넘어 임상적으로 의미 있는 개선인지 판단하는 기준으로 활용된다.

정성적 고찰 과정은 도구별 특징을 일관된 기준에 따라 비교할 수 있도록 구조화하여 수술하였다. 각 도구는 임상적 강점과 한계, 표준화 가능성, 제도적 수용성 등을 포함한 다차원적 기준으로 검토되었으며, 이를 통해 국가참조표준 도구 후보군(KPT Candidate Set)으로의 적합성을 평가하기 위한 방법론적 틀을 마련하였다.

Ⅲ. 연구결과

1. 물리치료 평가도구의 임상적 적용성 및 수가·국제 표준 연계 분석

1) 관절가동범위검사(Range of Motion: ROM)

관절가동범위검사(Range of Motion: ROM)는 환자의 관절 움직임 범위를 정량적으로 평가하는 핵심 도구로, 재활 및 물리치료 영역에서 기능 상태와 운동 제한 정도를 파악하는 데 활용된다. 이 검사는 신체기능(수행기반) 영역에 속하며, 근골격계 환자를 포함한 전 연령층을 대상으로 한다. 주된 목적은 ROM 기능 평가이며, 병원 및 지역사회 재활 환경 전반에서 사용된다.

신뢰도 측면에서 ROM 검사는 매우 우수한 수준으로 보고되었다. 고관절 가동범위를 측정된 연구에서 굴곡, 외회전, 내회전의 측정자간 신뢰도(ICC) 값은 0.97~0.98로 나타났으며 (Lee와 Kim, 2021), 고관절 전치환술 환자를 대상으로 한 또 다른 연구에서는 대부분의 측정 항목에서 ICC가 0.96 이상으로 확인되었다 (Kim 등, 2023). 타당도 역시 기존의 관절각도계 측정과 스마트폰 기반 측정 간 비교에서 $r > 0.90$ 의 높은 상관계수가 일관되게 보고되어, 동시 타당도(concurrent validity)가 확보되었다 (Lee와 Kim, 2021; Kim 등, 2023).

국제 표준 연계 측면에서 본 검사는 ICF의 신체구조 및 신체기능(운동) 영역과 직접적으로 매핑된다. 또한 SNOMED CT에서는 SCTID 364564000으로 등록되어 있으며, 국내 행정체계에서는 수가코드 EX773으로 분류된다 <Table 1>. 이러한 점은 ROM 검사가 임상적 신뢰성과 타당성을 확보했을 뿐 아니라, 국내·국제적 표준화 기반에서도 활용 가능성이 높은 평가도구임을 보여준다.

2) 도수근력검사(Manual Muscle Test: MMT)

도수근력검사(Manual Muscle Test: MMT)는 근력 수준을 평가하는 대표적인 임상 도구로, 근골격계 및 신경계 환자를 포함한 전 연령층을 대상으로 사용된다. 본 검사는 신체기능(수행기반) 영역에 속하며, 병원 및 지역사회 재활 환경에서 널리 활용된다. 주된 목적은 근력 평가이며, 임상적 진단과 치료 목표 설정, 경과 추적에 중요한 역할을 한다 (Shirley Ryan AbilityLab, 2021).

신뢰도 측면에서 MMT는 매우 우수한 결과가 보고되었다. 골관절염 환자를 대상으로 한 연구에서는 검사-재검사 신뢰도(Test-retest reliability)가 오른쪽 ICC = 0.98, 왼쪽 ICC = 0.97로 나타나 높은 일관성을 입증하였다 (Youdas 등, 2010). 타당도 또한 기계적 측정 도구와 비교해 확보되었다. 핸드 다이내모미터(Hand-held dynamometer, HHD)와의 수렴 타당도(convergent validity) 분석에서 어깨 외전, 목 굴곡, 고관절 외전 근육에서 상관계수 $r = 0.61\sim 0.85$ 로 나타났으며, 이는 MMT가 임상적으로 충분히 활용 가능한 평가도구임을 시사한다 (Schwartz 등, 1992).

국제 표준 연계 측면에서 MMT는 ICF의 신체구조 및 신체기능(운동) 영역과 매핑된다. 또한 SNOMED CT에서는 SCTID 350681005로 등재되어 있으며, 국내 행정체계에서는 건강보험 수가코드 E6612로 분류된다 <Table 1>. 이와 같이 MMT는 신뢰도와 타당도가 확보된 검증된 도구일 뿐 아니라, 국내·국제적 표준 체계에서도 활용 가능성이 높은 평가 도구이다.

3) 버그 균형 검사(Berg Balance Scale: BBS)

버그 균형 검사(Berg Balance Scale: BBS)는 성인과 노인, 특히 뇌졸중을 포함한 신경계 질환 환자의 균형 능력과 낙상 위험을 평가하기 위해 널리 사용되는 임상 도구이다. 이 검사는 신체기능과 정신기능을 포함하는 수행기반 평가에 속하며, 적용 대상은 신경계 질환자 및 고령 환자군이다. 주된 목적은 균형 기능의 정량적 평가와 낙상 위험 예측에 있으며, 적용 환경은 병원 입원 및 외래 진료 상황에서 이루어진다(Shirley Ryan AbilityLab, 2021).

BBS의 신뢰도는 매우 높은 내적 일관성을 보인다. 초기 연구에서 Cronbach's α 값은 0.96으로 보고되어 항목 간 일관성이 우수하고, 반복 적용 시 안정적 측정이 가능함이 입증되었다 (Berg 등, 1992). 타당도 역시 다양한 연구에서 입증되었다. 대표적으로 이동성 평가도구인 TUG (Timed Up and Go)와는 $r = -0.76$ 의 음의 상관관계를 보여 균형 능력이 높을수록 이동 시간이 단축됨을 시사한다. 또한, FRT (Functional Reach Test)와는 $r = 0.84$ 의 높은 양의 상관관계가 확인되어, 구조적 타당도가 확보되었다 (La Porta 등, 2012; Muir 등, 2008).

마지막으로, 본 도구는 ICF 기준상 본 도구는 활동(Activity, 운동) 영역에 매핑되며, 국내 행정체계에서 수가코드 E6613으로 분류되고, 국제적으로는 SNOMED CT 코드 428203000으로 등재되어 있다<Table 1>. 이는 BBS가 임상적 신뢰성과 타당성을 갖출 뿐 아니라 국내·국제 표준화 기반에서의 활용 가능성 또한 높음을 보여준다.

4) 10미터 보행 검사(10-Meter Walk Test: 10MWT)

10미터 보행 검사(10-Meter Walk Test: 10MWT)는 환자의 보행 속도와 이동성, 균형, 기능 회복 정도를 평가하기 위해 널리 사용되는 도구이다. 이 검사는 신체기능(수행기반) 영역에 속하며, 전연령(2세 이상)을 포함한 신경계 질환 환자를 대상으로 적용된다. 주된 목적은 보행속도, 이동성, 균형 및 기능회복 평가이며, 병원 및 지역사회 재활 환경에서 활용된다(Shirley Ryan AbilityLab, 2021).

신뢰도 측면에서 10MWT는 매우 우수한 결과가 보고되었다. 여러 연구에서 ICC = 0.95~0.99의 높은 측정자 간 신뢰도가 확인되었으며, 이는 급성기 뇌졸중 환자군을 포함한 다양한 집단에서 일관된 평가 도구로 활용될 수 있음을 보여준다 (Wienecke 등, 2023). 타당도 역시 충분히 입증되었다. 10MWT는 대표적 기능 평가도구인 Timed Up and Go (TUG) 검사와 $r = -0.91$ 의 높은 음의 상관관계를 보였으며, 이는 보행 속도가 증가할수록 TUG 수행 시간이 단축됨을 의미한다. 또한 6분 보행 검사(6MWT)와도 $r = 0.95$ 의 강한 양의 상관관계를 나타내어, 보행 기능과 지구력을 반영하는 구조적 타당도가 확보되었다 (Salbach 등, 2020).

국제 표준 연계 측면에서 10MWT는 ICF 기준의 활동(Activity) - 운동 도메인과 매핑되며, 국내 행정체계에서는 수가코드 52050으로 분류된다<Table 1>. 그러나 SNOMED CT 코드가 현재 부재하여, 국제 표준화 연계성 측면에서 추가적인 논의가 필요한 상태이다.

5) 6분 보행 검사(Six-Minute Walk Test: 6MWT)

6분 보행 검사(Six-Minute Walk Test: 6MWT)는 환자의 기능적 보행 능력, 심폐 지구력, 하지 근지구력을 종합적으로 평가하는 도구로, 재활 및 물리치료 임상에서 폭넓게 사용된다. 이 검사는 신체기능 및 활동-참여(관찰 기반) 영역에 속하며, 전연령(2세 이상)과 신경계 질환 환자를 대상으로 적용된다. 주요 목적은 기능적 보행과 지구력 평가이며, 병원과 지역사회 재활 환경에서 모두 활용된다(Shirley Ryan AbilityLab, 2021).

신뢰도 측면에서 6MWT는 높은 수준으로 보고되었다. 여러 연구에서 검사-재검사 신뢰도가 ICC = 0.86 ~ 0.99 범위로 확인되었으며(Eng, Dawson과 Chu, 2004; Flansbjerg 등, 2005), 이는 급성기부터 만성기에 이르는 다양한 환자군에서 일관적이고 안정적인 측정을 보장함을 시사한다. 타당도 또한 충분히 입증되었다. Timed Up and Go

(TUG) 검사와는 $r = -0.88 \sim -0.89$ 의 강한 음의 상관관계가 확인되어, 보행 기능이 향상될수록 TUG 수행 시간이 단축되는 경향을 설명한다. 또한 10미터 보행 검사(10MWT)와는 $r = 0.84 \sim 0.94$ 의 높은 양의 상관관계를 보여, 실제 보행 능력과 지구력을 반영하는 구조적 타당도를 입증하였다 (Flansbjer 등, 2005).

국제 표준 연계 측면에서 본 검사는 ICF 기준의 활동(Activity) - 운동 도메인과 매핑되며, SNOMED CT 코드 252478000으로 등록되어 있다<Table 1>. 그러나 국내에서는 아직 별도의 수가코드가 부재하여, 향후 제도적 수용성을 위한 논의가 필요한 상태이다.

6) 일상생활활동검사(Activities of Daily Living: ADL)

일상생활활동검사(Activities of Daily Living: ADL)는 성인 및 노인을 포함한 신경계 환자의 기능적 독립성을 평가하기 위해 널리 사용되는 임상 도구이다. 이 검사는 활동 및 참여(관찰) 영역에 속하며, 병원과 지역사회 재활 환경 모두에서 활용 가능하다. 주된 목적은 환자의 일상생활 수행 능력을 객관적으로 측정하여 재활 목표 설정 및 치료 효과 판정에 기초자료를 제공하는 것이다.

신뢰도 측면에서, ADL 검사는 매우 높은 수준의 내적 일치도(internal consistency)가 보고되었다. 국내 연구에서는 Cronbach's $\alpha = 0.937$ 로 나타나(김상규 등, 1999), 평가 문항 간의 일관성과 안정성이 충분히 확보되었음을 보여준다. 타당도 또한 인지기능 검사와의 비교를 통해 입증되었다. 특히 한국형 간이정신상태검사(K-MMSE)와의 상관관계 분석에서 $r = 0.54$ 가 보고되었으며, 이는 일상생활활동 수행 능력이 인지 기능과 유의한 연관성을 지니고 있음을 시사한다(김상규 등, 1999). 이러한 결과는 ADL 검사가 신체적 기능뿐 아니라 인지적 요인을 반영하는 포괄적 지표임을 의미한다.

국제 표준 연계 측면에서, 본 도구는 ICF의 활동·참여(일상생활활동) 영역과 매핑된다. 또한 SNOMED CT에서는 SCTID 129025006으로 등록되어 있으며, 국내 행정체계에서는 수가코드 EY773으로 분류된다<Table 1>. 따라서 ADL 검사는 임상적 신뢰성과 타당도를 갖춘 동시에, 행정적 수용성과 국제적 표준화 기반에서도 중요한 의미를 갖는 평가 도구라 할 수 있다.

7) 도구적 일상생활활동검사(Instrumental Activities of Daily Living, IADL)

도구적 일상생활활동검사(IADL)는 노인과 인지장애 환자를 대상으로 사회적 독립성 평가를 목적으로 활용되는 임상 도구이다. 본 검사는 활동 및 참여(관찰) 영역에 속하며, 기본적 자기관리 능력을 평가하는 ADL과 달리, 인지적·사회적 기능이 요구되는 도구적 활동(예: 금전 관리, 전화 사용, 약물 관리 등)을 중심으로 환자의 실제 생활 수행 능력을 평가한다는 점에서 차별성을 가진다. 따라서 IADL은 인지기능 저하와 관련된 기능적 독립성 평가에 특히 유용하다.

신뢰도 측면에서, IADL은 매우 높은 내적 일치도를 보이는 것으로 보고되었다. 국내 K-IADL 검증 연구에서 Cronbach's $\alpha = 0.96$ 으로 나타나 문항 간의 일관성이 탁월하게 확보되었음을 확인할 수 있다(원장원 등, 2002). 타당도 또한 인지기능과의 상관성을 통해 검증되었다. 인지기능 평가도구인 한국형 간이정신상태검사(Korean Mini-Mental Examination: K-MMSE)와의 상관분석에서 $r = 0.62$ 의 유의한 양의 상관관계가 확인되었으며(김상규 등, 1999), 이는 IADL이 인지적 기능 수준과 일상생활활동 수행 능력을 동시에 반영할 수 있는 수렴타당도를 갖추고 있음을 보여준다.

국제 표준 연계 측면에서 본 검사는 ICF의 활동(도구적 일상생활활동/인지/운동) 영역과 매핑된다. 또한 SNOMED CT에서는 SCTID 273534002로 등재되어 있으며, 국내 행정체계에서는 수가코드 EY774로 분류된다<Table 1>. 이러한 점은 IADL이 임상적 신뢰성과 타당성을 확보했을 뿐만 아니라, 국내·국제적 표준화 기반에서도 활용 가능성이 높은 평가도구임을 뒷받침한다. 더불어 ADL은 기본적 자기관리 능력에 초점을 두는 반면, IADL은 사회적·인지적 기능

을 포함한 복합적 독립성을 평가한다. 따라서 두 도구는 상호보완적으로 활용될 수 있으며, 환자의 일상생활 수행 능력을 다층적으로 이해하는 데 기여한다.

8) 기능적 독립성 측정(Functional Independence Measure: FIM)

기능적 독립성 측정(Functional Independence Measure: FIM)은 신경계 환자의 일상생활 기능을 종합적으로 평가하기 위해 고안된 대표적 도구로, 성인 및 노인을 포함한 다양한 환자군에서 활용된다. 본 검사는 활동 및 참여(관찰) 영역에 속하며, 주요 목적은 기능적 독립성 평가이다. 적용 환경은 입원 및 외래 재활 환경 전반에 걸쳐 이루어진다(Shirley Ryan Ability Lab, 2021).

신뢰도 측면에서 FIM은 높은 평가자 간 일치도를 보인다. 훈련된 평가자들 간의 신뢰도는 ICC = 0.95로 보고되어, 임상 현장에서 반복 적용 시 안정적인 측정이 가능함이 입증되었다(Hamilton 등, 1994). 타당도 또한 우수하게 입증되었다. Barthel Index(BI)와의 비교를 통해 동시 타당도를 검증한 결과, 상관관계수 $r = 0.91$ 로 강한 양의 상관성이 확인되었다(Granger 등, 1993). 이는 FIM이 실제 환자의 기능적 상태를 정확히 반영하는 도구임을 뒷받침한다.

국제 표준 연계 측면에서 FIM은 ICF의 활동(Activity) 및 참여(Participation) 영역과 직접적으로 매핑되며, SNOMED CT에서는 SCTID 718705001로 등재되어 있다<Table 1>. 국내 행정체계에서는 현재 수가코드가 부여되어 있지 않으나, 국제적 표준화 기반에서는 널리 인정받는 평가도구이다.

9) 바텔 지수(Barthel Index: BI)

바텔 지수(Barthel Index: BI)는 환자의 일상생활활동(Activities of Daily Living: ADL) 수행 능력을 평가하기 위한 대표적인 도구로, 성인 및 노인, 특히 신경계 질환 환자의 기능적 독립성을 정량화하는 데 널리 사용된다(Shirley Ryan Ability Lab, 2021). 평가 영역은 일상생활, 이동성 및 신체기능(수행기반)으로 구성되며, 병원과 지역사회 재활 환경 모두에서 활용된다(Shirley Ryan Ability Lab, 2021). 국내에서는 원형 BI보다 수정 바텔 지수(Modified Barthel Index: MBI)가 더 자주 사용되며, 실제 행정 체계에서도 수가코드 EY772로 분류되어 있다.

신뢰도 측면에서 BI는 높은 수준의 내적 일관성을 보여주며, Cronbach's α 값은 0.90으로 보고되었다(Shah 등, 1989). 이는 항목 간의 평가 결과가 일관되게 유지됨을 의미한다. 타당도 역시 입증되었는데, 기능적 독립성 측정 도구인 FIM과의 상관관계가 $r = 0.91$ 로 나타나(BI와 FIM 간 동시 타당도 확보), 두 도구 간 결과가 상호 보완적으로 활용될 수 있음을 시사한다(Hsueh 등, 2002).

국제 표준 연계성 측면에서, BI/MBI는 ICF의 활동(일상생활동작/운동) 영역과 매핑되며, SNOMED CT에서는 코드 273302005로 등재되어 있다<Table 1>. 이러한 점은 본 도구가 임상적 신뢰도와 타당도를 확보했을 뿐 아니라, 국내 수가체계와 국제 표준 분류체계 모두에서 활용될 수 있는 기반을 갖추고 있음을 보여준다.

Table 1. Analysis of Physical Therapy Evaluation Instruments

한글 평가도구명	수가 코드	평가 영역 (평가 방식)	질환군	평가 목적	적용 환경	ICF 도메인 (측정 영역)	SNOMED CT 코드 (SCTID)	신뢰도	타당도
관절가동범위 검사	EX773	신체기능 (수행기반)	전연령, 근골격계	ROM 평가	병원, 지역사회	신체구조, 신체기능 (운동)	364564000	ICC = 0.97~0.98	스마트폰 $r > 0.9$
도수근력검사 (MMT)	E6612	신체기능 (수행기반)	전연령, 근골격계, 신경계	근력	병원, 지역사회	신체구조, 신체기능 (운동)	350681005	ICC = 0.97~0.98	HHD $r > 0.60$
버그 균형 검사 (BBS)	E6613	신체기능 정신기능 (수행기반)	질환, 성인, 노인	기능 평가, 낙상 위험	입원/외래	활동 (운동)	428203000	Cronbach's $\alpha = 0.96$	TUG $r = -0.76$, FRT $r = 0.84$
10미터 보행 검사 (10MWT)	52050	신체기능 균형 움직임 (수행기반)	전연령 (2세 이상), 신경계 질환	보행속도, 이동성, 균형, 기능회복	병원, 지역사회	활동 (운동)	없음	ICC = 0.95~0.99	TUG $r = -0.91$, 6MWT $r = 0.95$
6분 보행 검사 (6MWT)	없음	신체 기능, 활동 및 참여(관찰)	전연령 (2세 이상), 신경계 질환	기능적 보행, 심폐지구 력, 하지 근지구력	병원, 지역사회	활동 (운동)	252478000	ICC=0.86 ~ 0.99	TUG $r = -0.88$ ~ -0.89 , 10MWT $r = 0.84$ ~ 0.94
일상생활활동 검사(ADL)	EY773	활동 및 참여 (관찰)	성인, 노인, 신경계	기능적 독립성 평가	병원, 지역사회	활동/참여 (일상생활 동작) 활동 (도구적 일상생활 동작 /인지 /운동) 활동 (일상생활 동작 /인지 /운동)	129025006	Cronbach's $\alpha = 0.937$	K-MMSE $r = 0.54$
도구적 일상생활활동 검사 (IADL)	EY774	활동 및 참여 (관찰)	노인, 인지장애	사회적 독립성 평가	병원, 지역사회	활동 (도구적 일상생활 동작 /인지 /운동) 활동 (일상생활 동작 /인지 /운동)	273534002	Cronbach's $\alpha = 0.96$	K-MMSE $r = 0.62$
기능적 독립성 측정 (FIM)	없음	활동 및 참여 (관찰)	신경계 질환, 성인, 노인	기능적 독립성 평가	입원/외래	활동 (일상생활 동작 /인지 /운동) 활동 (일상생활 동작 /인지 /운동)	718705001	ICC = 0.95	BI $r = 0.91$
바텔 지수 (BI)	EY772 (MBI)	일상생활, 이동성, 신체기능 (수행기반)	성인, 노인, 신경계 질환	ADL 능력 평가	병원, 지역사회	활동 (일상생활 동작/ 운동)	273302005	Cronbach's $\alpha = 0.90$	FIM $r = 0.91$

ICF = International Classification of Functioning, Disability and Health; SCTID = SNOMED CT Identifier; BBS = Berg Balance Scale; TUG = Timed Up and Go Test; FRT = Functional Reach Test; 10MWT = 10-Meter Walk Test; 6MWT = 6-Minute Walk Test; MMT = Manual Muscle Testing; HHD = Hand-Held Dynamometer; ROM = Range of Motion; IADL = Instrumental Activities of Daily Living; ADL = Activities of Daily Living; K-MMSE = Korean Mini-Mental State Examination; FIM = Functional Independence Measure; BI = Barthel Index; MBI = Modified Barthel Index

2. 평가도구의 임상적 민감도 및 정성적 분석

1) 관절가동범위검사(Range of Motion: ROM)

관절가동범위(ROM) 검사는 단순한 신뢰도와 타당도 검증을 넘어, 측정 민감도 지표를 통해 실제 환자의 기능 변화를 판별하는 데 중요한 역할을 한다. 인공 고관절 전치환술 환자를 대상으로 한 연구에서, 표준측정오차(Standard Error of Measurement: SEM)는 외전 2.71도, 내전 1.88도, 외회전 1.94도, 굴곡 1.97도, 신전 2.00도로 보고되었다. 이는 반복 측정 시 발생할 수 있는 오차 범위를 의미하며, 임상 해석 시 변화량의 신뢰도를 보완하는 기준으로 활용된다 (Huang et al., 2023). 또한, 최소 감지 변화량(Minimal Detectable Change: MDC)은 다음의 공식에 의해 산출되며,

$$MDC = 1.96 \times \sqrt{2} \times SEM$$

동일 환자군에서 외전 7.51도, 내전 5.21도, 외회전 5.37도, 굴곡 5.46도, 신전 5.54도로 나타났다. 이는 ROM 값의 변화가 단순 오차가 아닌 실제 기능적 변화임을 입증하는 기준치로, 치료 효과 판별 및 재활 경과 추적에 활용될 수 있다 (Huang 등, 2023).

2) 도수근력검사(Manual Muscle Test: MMT)

도수근력검사(Manual Muscle Test: MMT)의 임상적 민감도 지표에서 MMT는 신뢰성과 타당성을 뒷받침하는 구체적인 수치가 보고되었다. 골관절염 환자를 대상으로 한 연구에서 표준측정오차(SEM)는 오른쪽 1%, 왼쪽 2%로 나타났으며, 최소 감지 변화량(Minimal Detectable Change: MDC)은 양측 모두 4%로 보고되었다 (Youdas et al., 2010). 또한 염증성 근병증 환자에서 MMT-8 총점 기준 평가자 내 MDC는 4.9점(6.9%), 평가자 간 MDC는 6.2점(8.6%)로 보고되었다 (Pfister 등, 2018). 다만, 근력 점수가 3.5 이상인 경우 변화 감지 민감도가 낮아질 수 있다는 한계가 보고되었으며, 이러한 구간에서는 수치적 근력 측정을 위한 Myometry 등의 정량적 도구와 병행하는 것이 권장된다(Schwartz 등, 1992).

3) 버그 균형 검사(Berg Balance Scale: BBS)

BBS는 균형 능력과 낙상 위험을 정량화하는 도구로, 다양한 환자군에서 임상적으로 의미 있는 변화량이 보고되어 왔다. 표준측정오차(SEM)는 BBS 점수 해석에서 측정 안정성을 나타내는 지표로, 노인 집단과 뇌졸중 환자군 모두에서 약 2.5점으로 보고되었다 (Donoghue 과 Stokes, 2009). 이는 동일 환자에 대한 반복 측정에서 2.5점 이내의 변화는 측정 오차 범위에 속할 수 있음을 의미한다.

최소 감지 변화량(MDC)은 실제 변화를 판별하는 데 핵심적인 지표로, 환자군에 따라 다소 차이를 보인다. 노인 집단에서는 약 4~6점, 뇌졸중 환자에서는 6.3점으로 보고되었다 (Stevenson, 2001). 따라서 일반 노인의 경우 BBS 점수가 4점 이상 변화해야 실제 기능 개선으로 볼 수 있으며, 뇌졸중 환자의 경우 최소 6점 이상의 변화가 있어야 치료 효과로 해석될 수 있다.

마지막으로 최소 임상 중요 변화(Minimal Clinically Important Difference: MCID)은 환자와 치료자가 모두 '실질적인 기능적 개선'으로 체감할 수 있는 최소 수준의 변화로 정의된다. 뇌졸중 환자를 대상으로 한 연구에서 BBS의 MCID는 약 6~7점으로 보고되었다 (Stevenson, 2001). 이는 치료 개입 이후 환자의 균형 기능이 단순히 통계적으로 유의한 수준을 넘어, 실제로 보행 안정성 향상 및 낙상 위험 감소로 이어졌음을 판별하는 기준이 된다. 이와같이 BBS는 노인 집단과 뇌졸중 환자군 모두에서 SEM, MDC, MCID가 구체적으로 보고된 균형 평가 도구로 환자군별 차이를

고려한 해석이 가능하다. 특히 뇌졸중 환자의 경우 6점 이상의 변화가 실제 임상적으로 유의미한 개선으로 받아들여질 수 있으며, 이는 치료 목표 설정과 효과 판정에서 중요한 기준이 된다.

4) 10미터 보행 검사(10-Meter Walk Test, 10MWT)

10MWT는 뇌졸중 환자, 노인, 신경계 질환자를 대상으로 보행 속도를 정량화하여 기능적 이동성과 낙상 위험을 평가하는 대표적 도구로 측정 민감도 지표가 임상 해석의 핵심으로 제시된다. Perera 등(2006)의 연구에 따르면, 뇌졸중 환자 100명을 대상으로 한 분석에서 표준측정오차(SEM)는 0.04 m/s로 보고되어 반복 측정 시 높은 안정성을 보였으며, 최소 감지 변화량(MDC)은 약 0.10 m/s 전후로 제시되어 측정된 변화가 단순한 오차가 아닌 실제적인 기능 변화를 의미함을 뒷받침한다. 또한 최소 임상 중요 변화(MCID)은 작은 변화 0.06 m/s, 큰 변화 0.14 m/s로 제시되어, 이는 뇌졸중 환자가 기능적으로 제한적 지역사회 보행에서 완전한 지역사회 보행으로 전환되는 등 임상적으로 의미 있는 향상을 반영하는 수치로 활용될 수 있다. 이러한 결과는 10MWT가 뇌졸중 환자를 비롯한 다양한 임상 집단에서 치료 효과 판별과 재활 목표 설정에 있어 신뢰성과 타당성을 갖춘 평가 도구임을 보여준다 (Collen 등, 1990; Flansbjer 등, 2005; Perera 등, 2006).

5) 6분 보행 검사(Six-Minute Walk Test, 6MWT)

6분 보행 검사(6MWT)는 뇌졸중 환자의 심폐지구력과 기능적 보행 능력을 평가하는 객관적 도구로, 저하된 이동성과 일상생활 수행능력에 대한 간접 지표로 활용된다. 검사는 30미터 직선 보행 공간과 스톱워치만으로 시행할 수 있으며, 보행 보조기구 사용 여부와 관계없이 적용 가능하다.

임상적 민감도 지표로 표준측정오차(SEM)는 뇌졸중 환자를 대상으로 한 연구에서 약 19.9m로 보고되었으며, 이는 반복 측정 시 발생할 수 있는 오차 범위를 의미한다 (Tang 등, 2006). 최소 감지 변화량(MDC)은 동일 집단에서 약 34.4m로 산출되어, 측정된 변화가 단순한 오차가 아닌 실제적인 기능 변화를 의미하는 기준으로 제시되었다 (Flansbjer 등, 2005). 또한 최소 임상 중요 변화(MCID)는 뇌졸중 환자에서 약 34.4m로 보고된 바 있으며 (Flansbjer 등, 2005), 노인 집단을 대상으로 한 연구에서는 작은 변화 20m, 큰 변화 50m가 임상적으로 의미 있는 향상의 기준으로 제시되었다 (Perera 등, 2006). 이러한 수치는 환자군의 특성에 따라 기능적 회복의 의미를 구체적으로 제시하며, 6MWT가 재활 목표 설정과 치료 효과 판별에 활용될 수 있는 근거를 제공한다.

6) 도구적 일상생활활동검사(Instrumental Activities of Daily Living, IADL)

도구적 일상생활활동검사(IADL: Instrumental Activities of Daily Living)는 노인의 사회적 독립성과 자립 능력을 평가하기 위해 개발된 대표적 기능 평가 도구로, ADL보다 복잡한 인지·집행 기능 요소를 반영한다(Won 등, 2002; Chin 등, 2018). 한국형 IADL(K-IADL)은 문화적 특성을 반영하여 10개 항목으로 구성되며, 치매 및 경도인지장애(MCI) 환자의 기능 저하 선별에 유효성이 입증되었다(Chin et al., 2018).

임상적 민감도 지표로, 고령자를 대상으로 한 연구에서 표준측정오차(SEM)는 약 0.4-1.2점 범위로 보고되었으며, 반복 측정 시 발생할 수 있는 오차의 한계를 제시한다(Suijker 등, 2017). 최소 감지 변화량(MDC)은 SEM을 바탕으로 계산되며, 약 1.52점의 변화가 실제 임상적으로 신뢰할 수 있는 차이로 확인되었다(Suijker 등, 2017). 또한, 최소 임상 중요 변화(MCID)는 환자나 임상가가 '의미 있는 변화'로 인식할 수 있는 기준으로, 일반적으로 총점의 약 10-15%가 적용된다. 이는 IADL 총점 8점 기준 약 0.8-1.2점의 변화에 해당하며(Chuang 등, 2011; Revicki 등, 2008), 치매 및 경도인지장애 환자의 기능적 변화 해석에 활용된다.

7) 기능적 독립성 측정(Functional Independence Measure: FIM)

FIM은 일상생활동작(ADL)을 포함한 전반적인 기능적 독립성을 평가하기 위해 개발된 도구로, 18개 항목(운동 13개, 인지 5개)을 7점 척도로 평가하여 총점 18~126점으로 산출된다(Granger 등, 1993; Hamilton 등, 1994). 주로 신경계 질환자, 특히 뇌졸중 환자의 기능적 독립성을 평가하기 위해 훈련된 치료사나 의료 전문가가 관찰을 통해 시행한다(Granger 등, 1993).

Beninato 등(2006)의 연구에서는 급성 뇌졸중 환자를 대상으로 최소 임상 중요 변화(MCID)를 산출하였으며, FIM 총점(Motor + Cognitive)은 22점, 운동 영역은 17점, 인지 영역은 3점으로 보고되었다(Beninato 등, 2006). Coster 등(2006)은 다양한 재활 환자군을 분석하여 반응성을 제시하였는데, 운동 영역의 표준화 반응 평균(Standardized Response Mean: SRM)은 0.73~1.05, 인지 영역은 0.34~0.35로 나타났다. 반면, 일부 조건에서는 운동 영역이 1.30~1.31, 인지 영역이 1.34~2.24로 부정적 변화가 보고되었다(Coster 등, 2006). 또한, 최소 감지 변화량(MDC)을 초과한 변화 비율은 운동 영역에서 15~36%가 긍정적, 15~25%가 부정적이었으며, 인지 영역에서는 89%가 긍정적, 20~24%가 부정적으로 나타났다(Coster 등, 2006). 이러한 수치는 FIM이 뇌졸중 및 재활 환자군에서 치료 전후의 변화를 통계적·임상적으로 의미 있게 판별할 수 있는 민감한 평가도구임을 보여준다.

8) 바텔 지수(Barthel Index: BI)

Barthel Index(BI)는 10개 항목(식사, 세면, 목욕, 옷 입기, 대소변 조절, 화장실 사용, 침대-의자 이동, 보행, 계단 오르내리기 등)을 통해 일상생활활동(ADL) 수행 능력을 정량적으로 평가하는 대표적 도구로, 총점은 0~100점으로 산출된다(Mahoney와 Barthel, 1965; Shah 등, 1989).

임상적 민감도 지표에서, 최소 임상 중요 변화(MCID)는 환자군별로 상이하게 보고되었다. 뇌졸중 환자의 경우 1.85점(Hsieh 등, 2007), 고령자에서는 9.8점(Unnanuntana 등, 2018), 재활 병동 입원 환자에서는 35점(Castiglia 등, 2017)이 MCID로 제시되어, 기능 향상이 실제로 임상적 의미를 가지는지 판별할 수 있는 기준으로 사용된다. 최소 감지 변화량(MDC)은 만성 뇌졸중 환자에서 4.02점, 전체 뇌졸중 환자군에서 6.84점으로 보고되어, 측정된 변화가 단순 오차를 넘어 실제 변화임을 검증하는 지표로 활용된다(Hsieh 등, 2007; Ghandehari 등, 2012).

IV. 고 찰

본 연구는 국내 물리치료 임상에서 널리 사용되는 9개 평가도구(ROM, MMT, BBS, 10MWT, 6MWT, ADL, IADL, FIM, BI/MBI)를 대상으로 기능 영역, 적용 환자군, 국제 표준 연계성, 민감도 지표 등을 종합적으로 고찰하였다(Granger 등, 1993; Mahoney와 Barthel, 1965; Won 등, 2002). 특히 표준측정오차(Standard Error of Measurement, SEM), 최소 감지 변화량(Minimal Detectable Change, MDC), 최소 임상 중요 변화(Minimal Clinically Important Difference, MCID)와 같은 통계·심리측정 지표를 통해 단순한 수치 차이를 넘어, 임상적으로 체감 가능한 변화를 판별할 수 있음을 확인하였다(Beninato 등, 2006; Donoghue와 Stokes, 2009; Hsieh 등, 2007).

ROM과 MMT는 가장 기본적인 기능평가임에도 불구하고 여전히 임상에서 필수적인 도구로 자리잡고 있다. ROM은 스마트폰 기반 측정 연구에서 신뢰도가 확보되었으나, SEM·MDC·MCID 값은 환자군마다 일관되지 않아 보편적 기준 제시에는 한계가 있었다(Shin 등, 2020; Kim 등, 2022). MMT의 경우, 골관절염 환자에서 intra-rater 기준 4.9%,

inter-rater 기준 6.2%의 MDC가 보고되어, 근력 변화가 단순한 오차를 넘어 실제적 향상을 의미하는지 해석하는 지표로 활용될 수 있다 (Youdas 등, 2010; Noreau와 Vachon, 1998).

균형 능력 평가도구인 BBS는 특히 뇌졸중과 고령 환자군에서 임상적 민감도가 높게 나타났다. 기존 연구에서는 SEM이 약 2-3점, MDC가 6-7점으로 보고되었고, 뇌졸중 환자의 경우 6점 이상의 변화가 치료 효과를 판정하는 핵심 기준으로 제시되었다 (Berg 등, 1992; Donoghue 와 Stokes, 2009).

보행 관련 도구인 10MWT와 6MWT는 환자의 이동성과 심폐 지구력을 평가하는 데 중요한 지표로 기능하였다. 10MWT는 뇌졸중 환자에서 SEM이 0.04 m/s, MDC가 약 0.10 m/s, MCID가 0.06-0.14 m/s로 확인되어, 제한적 지역사회 보행자에서 독립 보행자로의 전환 여부를 해석하는 기준으로 유용하였다 (Perera 등, 2006; Salbach 등, 2020). 6MWT에서는 뇌졸중 환자군에서 SEM이 19.9m, MDC가 34.4m, MCID는 34m 이상으로 제시되어, 심폐지구력과 장거리 보행 능력 평가에 적절하였다 (Flansbjerg 등, 2005; Wienecke 등, 2023).

일상생활 수행능력을 평가하는 도구 중 한국형 IADL(K-IADL)은 치매 및 경도인지장애 환자에서 SEM 0.4-1.2점, MDC 1.5-2점, MCID는 총점의 10-15% 수준으로 보고되어, 고위험군 조기 선별에 효과적이었다 (Won 등, 2002; Chin 등, 2018). FIM은 급성기 뇌졸중 환자에서 총점 MCID가 22점, 운동 영역 17점, 인지 영역 3점으로 확인되어, 기능 회복 과정을 임상적으로 의미 있게 판별할 수 있는 기준을 제공하였다 (Beninato 등, 2006). BI는 환자군별로 다른 민감도 값이 보고되었는데, 뇌졸중 환자에서 MCID는 1.85점, 고령자군에서는 9.8점, 재활 병동 환자에서는 35점으로 나타났으며, 만성 뇌졸중 환자의 경우 MDC가 4.02점, 전체 뇌졸중 환자군에서는 6.84점으로 제시되었다 (Hsieh 등, 2007; Unnanuntana 등, 2018; Castiglia 등, 2017). 국내 임상에서는 BI보다 MBI가 더 널리 활용되고 있으며, 장기요양과 재활 판정의 필수 기준으로 자리매김하고 있다 (Kwon 등, 2004).

행정적·제도적 관점에서, 일부 도구는 보험 수가코드와 연계되어 전자의무기록(Electronic Medical Record, EMR) 기반 청구가 가능했으나, FIM과 6MWT는 여전히 수가코드가 부재하여 제도적 활용에 제약이 있었다 (Granger 등, 1993; Hamilton 등, 1994). 국제 표준 연계성 측면에서는 다수의 도구가 ICF 및 SNOMED CT와 매핑되어 있었으나, 10MWT는 SNOMED CT 코드가 부재해 국제적 상호운용성을 제한하는 한계가 있었다 (Salbach 등, 2020; Maritz 등, 2017).

이러한 결과는 동일 도구라도 환자군·연구 환경에 따라 SEM, MDC, MCID 값이 달라져 임상 적용에 주의가 필요함을 보여준다. BI의 경우, 뇌졸중 환자에서 1.85점, 고령자군에서는 9.8점, 재활 병동 환자에서는 35점으로 제시되어 환자 특성에 따른 기준 재설정 필요했다. 또한 국내 연구가 뇌졸중과 고령자 중심으로 편중되어, 근골격계·호흡기 환자군에 대한 근거는 부족한 실정이다 (Park 등, 2022; Eng 등, 2004).

본 연구의 가장 큰 시사점은, 국내에서 광범위하게 활용되고 있음에도 불구하고 일부 평가도구가 여전히 보험 수가 체계에 충분히 반영되지 못한다는 구조적 제약이다. 이는 임상 현장에 적합한 도구가 있음에도 불구하고, 제도적 지원이 미비하여 활용도가 제한되는 전형적 사례라 할 수 있다 (Granger 등, 1993; Hamilton 등, 1994). 따라서 향후 국가 물리치료 참조표준(Korean Physical Therapy Toolkit, KPT)을 구축할 때는 단순한 임상적 타당성 검증을 넘어 수가체계와의 연계성 확보가 핵심 과제로 포함되어야 한다. 이를 위해 국제 표준(ICF, SNOMED CT)과의 매핑 강화, 임상 빈도와 활용도에 따른 수가 반영 우선순위 재정립, 그리고 비용-효과 분석을 통한 급여 체계 검토가 필요하다 (Üstün 등, 2003; Leonardi 등, 2022; Maritz 등, 2017).

특히 FIM과 6MWT의 수가 반영은 단순한 행정적 조치가 아니라, 환자의 예후 예측과 재활 성과 표준화, 더 나아가 국가 차원의 의료 빅데이터 활용 기반을 마련하는 중요한 과제로, 향후 학문적·제도적 검토가 반드시 병행되어야 할 것이다.

V. 결론

본 연구는 국내 물리치료 임상에서 활용되는 9개 주요 평가도구(ROM, MMT, BBS, 10MWT, 6MWT, ADL, IADL, FIM, BI/MBI)를 종합적으로 분석하여 기능 영역, 적용 환자군, 국제 표준 연계성, 그리고 민감도 지표의 활용 가능성을 검토하였다. 분석 결과, 다수의 도구는 임상적 신뢰도와 타당성을 확보하였으며, 표준측정오차(Standard Error of Measurement, SEM), 최소감지변화량(Minimal Detectable Change, MDC), 최소 임상 중요 변화(Minimal Clinically Important Difference, MCID)와 같은 민감도 지표를 통해 단순한 수치 차이를 넘어 실제 임상적으로 의미 있는 변화를 판별할 수 있음을 확인하였다.

특히 BBS, FIM, BI/MBI와 같은 기능 중심 도구들은 균형, 독립성, 일상생활 수행능력의 변화를 민감하게 반영하여 치료 효과 판정과 목표 설정 과정에서 중요한 기준으로 작용하였다. 그러나 동일한 도구라 하더라도 환자군 및 재활 환경에 따라 SEM, MDC, MCID 값이 다르게 보고되어, 임상적 해석에는 주의가 필요함이 드러났다.

행정적·제도적 측면에서 ROM, MMT, BBS, 10MWT, IADL, BI 등 일부 도구는 이미 수가코드와 연계되어 전자의무기록(Electronic Medical Record, EMR) 기반 청구가 가능했으나, FIM과 6MWT는 수가코드가 부재하여 실제 활용 범위가 제한되는 구조적 한계가 있었다. 이는 도구의 임상적 유용성이 충분히 입증되었음에도 불구하고 제도적 기반이 이를 뒷받침하지 못하고 있음을 의미한다.

따라서 향후 국가 물리치료 참조표준(Korean Physical Therapy Toolkit, KPT)을 마련하는 과정에서는 단순한 도구 분류와 타당성 검증을 넘어, 건강보험 수가 반영과 국제 표준(ICF, SNOMED CT)과의 매핑을 포함한 제도적 연계가 핵심 과제로 다루어져야 한다. 특히 FIM과 6MWT의 제도적 반영은 환자 예후 예측과 재활 성과의 표준화, 나아가 국가 차원의 의료 빅데이터 활용 기반 마련이라는 측면에서 학문적·정책적으로 시급히 검토될 필요가 있다.

참고문헌

- 김정현. 디지털 치료제가 뇌성마비 아동의 균형 및 보행에 미치는 효과: 체계적 고찰. 대한물리치료과학회지. 2023;30(4):92-110.
- 김종엽, 김정현. 공공 물리치료 인프라 구축이 지역사회 보건 효과 비용에 미치는 영향-수도권과 지방 도시의 비교. 대한물리치료과학회지. 2025;32(2):75-91.
- 김상규, 김석범, 강복수. 농촌 노인들의 인지기능장애와 일상생활활동도와의 관련성. 예방의학회지. 1999;32(1):65-71.
- 원장원, 노용균, 김수영 등. 한국형 일상생활활동 측정도구(K-ADL)의 타당도 및 신뢰도. 노인복지연구. 2002;6(2):98-106.
- 원장원, 양금열, 노용균 등. 한국형 일상생활활동 측정도구(K-ADL)와 한국형 도구적 일상생활활동 측정도구(L-IADL)의 개발. 노인복지연구. 2002;6(2):107-120.
- 원장원, 노용균, 선우덕 등. 한국형 도구적 일상생활활동 측정도구의 타당도 및 신뢰도. 노인복지연구. 2002;6(4):45-63.
- Baschung Pfister P, de Bruin ED, Sterkele I, et al. Manual muscle testing and hand-held dynamometry in people with inflammatory myopathy: An intra- and interrater reliability and validity study.

- PLoS One. 2018;13(3):e0194531.
- Beninato M, Gill-Body KM, Salles S, et al. Determination of the MCID for the FIM in stroke patients. *Arch Phys Med Rehabil.* 2006;87(1):32-9.
- Berg K, Wood-Dauphinee S, Williams JI, et al. Measuring balance in the elderly: Validation of an instrument. *Can J Public Health.* 1992;83(Suppl 2):S7-11.
- Castiglia SF, Galeoto G, Lauta A, et al. The culturally adapted Italian version of the Barthel Index (IcaBI). *Funct Neurol.* 2017;32(4):221-7.
- Chin J, Park J, Yang SJ, et al. Re-standardization of the Korean-Instrumental Activities of Daily Living (K-IADL): Clinical usefulness for various neurodegenerative diseases. *Dement Neurocogn Disord.* 2018;17(1):11-22.
- Chuang LL, Chien CW, Wu CY, et al. Responsiveness, MDC, and MCID of the Nottingham EADL Scale in stroke rehabilitation. *Arch Phys Med Rehabil.* 2011;92(8):1281-7.
- Cieza A, Hilfiker R, Chatterji S, et al. The ICF could be used to measure functioning. *J Clin Epidemiol.* 2009;62(9):899-911.
- Collen F, Wade D, et al. Mobility after stroke: Reliability of measures of impairment and disability. *Disabil Rehabil.* 1990;12(1):6-9.
- Coster WJ, Haley SM, Jette AM. Measuring patient-reported outcomes after discharge from inpatient rehabilitation. *Arch Phys Med Rehabil.* 2006;87(8):1033-42.
- Donoghue D, Stokes EK. How much change is true change? The minimum detectable change of the Berg Balance Scale in elderly people. *J Rehabil Med.* 2009;41(5):343-6.
- Eng JJ, Dawson AS, Chu KS. Submaximal exercise in persons with stroke: Reliability and validity. *Arch Phys Med Rehabil.* 2004;85(1):113-8.
- Flansbjer UB, Holmbäck AM, Downham D, et al. Reliability of gait performance tests in men and women with hemiparesis after stroke. *J Rehabil Med.* 2005;37(2):75-82.
- Ghandehari K, Saffarian-Toosi G, Masoudinezhad S, et al. Comparative interrater reliability of Asian stroke disability scale, modified Rankin scale and Barthel Index in patients with brain infarction. *ARYA Atheroscler.* 2012;8(3):153-7.
- Granger CV, Cotter AC, Hamilton BB, et al. Functional assessment scales: A study of persons after stroke. *Arch Phys Med Rehabil.* 1993;74(2):133-8.
- Hamilton BB, Laughlin JA, Fiedler RC, et al. Interrater reliability of the 7-level FIM. *Scand J Rehabil Med.* 1994;26(3):115-9.
- Hsieh YW, Wang CH, et al. Establishing the minimal clinically important difference of the Barthel Index in stroke patients. *Neurorehabil Neural Repair.* 2007;21(3):233-8.
- Hsueh IP, Lin JH, Jeng JS, et al. Comparison of the FIM, 5-item BI, and 10-item BI in patients with stroke. *J Neurol Neurosurg Psychiatry.* 2002;73(2):188-90.
- Keith RA, Granger CV, Hamilton BB, et al. The Functional Independence Measure: A new tool for rehabilitation. *Adv Clin Rehabil.* 1987;1:6-18.
- Kim K, Lee SM, Lee JH, et al. The reliability and validity of hip ROM measurement using a smartphone

- for THA patients. *J Korean Phys Ther.* 2023;35(2):49-56.
- Kim KM, Lee SH, Park SD. The reliability and validity of hip ROM measurement using a smartphone in operative patients. *Healthcare (Basel).* 2022;10(2):338.
- Kristensen MT, Bandholm T, Foss NB. Performance stability and interrater reliability of culturally adapted 10-Meter Walking Test for Danes with neurological disorders. *J Stroke Cerebrovasc Dis.* 2019;28(9):2459-67.
- Kwon S, Hartzema AG, Duncan PW, et al. Disability measures in stroke: Relationship among the Barthel Index, the Functional Independence Measure, and the Modified Rankin Scale. *Stroke.* 2004;35(4):918-23.
- La Porta F, Caselli S, Susassi S, et al. Is the Berg Balance Scale an internally valid and reliable measure of balance? A Rasch analysis. *Arch Phys Med Rehabil.* 2012;93(7):1209-16.
- Lawton MP, Brody EM. Assessment of older people: Self-maintaining and IADL. *Gerontologist.* 1969;9(3):179-86.
- Lee JM, Kim K. Reliability and validity of hip ROM measurement using a smartphone. *Phys Ther Rehabil Sci.* 2021;10(1):18-25.
- Leonardi M, Lee H, Kostanjsek N, et al. 20 years of ICF: Uses and applications worldwide. *Int J Environ Res Public Health.* 2022;19(18):11321.
- Mahoney FI, Barthel DW. Functional evaluation: The Barthel Index. *Md State Med J.* 1965;14:61-5.
- Maritz R, Aronsky D, Prodinger B. The ICF in electronic health records. *Appl Clin Inform.* 2017;8(3):964-75.
- Muir SW, Berg K, Chesworth B, et al. Use of the Berg Balance Scale for predicting multiple falls in community-dwelling elderly people. *Phys Ther.* 2008;88(4):449-59.
- Noreau L, Vachon J. Comparison of three methods to assess muscular strength in SCI. *Spinal Cord.* 1998;36(10):716-23.
- Park DS, et al. The factors affecting frailty among the elderly in Korea. *Int J Environ Res Public Health.* 2022;19(24):16553.
- Perera S, Mody SH, Woodman RC, et al. Meaningful change and responsiveness in common physical performance measures in older adults. *J Am Geriatr Soc.* 2006;54(5):743-9.
- Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and MCID in PROs. *J Clin Epidemiol.* 2008;61(2):102-9.
- Salbach NM, Mayo NE, Wood-Dauphinee S, et al. Validation of stroke-specific protocols for the 10-meter walk test and 6-minute walk test. *Top Stroke Rehabil.* 2020;27(4):251-61.
- Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. *J Clin Epidemiol.* 1989;42(8):703-9.
- Shin SH, Kim YR, Lee HJ. Reliability and validity of ROM measurement using a smartphone application. *Ann Rehabil Med.* 2020;44(2):131-8.
- Shirley Ryan AbilityLab. Rehab Measures Database. 2021. Available at: <https://www.sralab.org/rehabilitation-measures>
-

- Stevenson TJ. Detecting change in patients with stroke using the Berg Balance Scale. *Aust J Physiother.* 2001;47(1):29-38.
- Suijker JJ, van Rijn M, ter Riet G, et al. Minimal important change and MDC in ADL in older people. *Eur Geriatr Med.* 2017;8(5):415-21.
- Tang A, Sibley KM, Bayley M, et al. Do functional walk tests reflect cardiorespiratory fitness in sub-acute stroke? *J Neuroeng Rehabil.* 2006;3:23.
- Unnanuntana A, Jarusriwanna A, Nepal S. Validity and responsiveness of Barthel Index for measuring functional recovery after hemiarthroplasty for femoral neck fracture. *Arch Orthop Trauma Surg.* 2018;138(12):1671-7.
- Üstün TB, Chatterji S, Bickenbach J, et al. The ICF: A new tool for understanding disability and health. *Disabil Rehabil.* 2003;25(11-12):565-71.
- Wienecke T, Jacobsen AB, Husted SS. Inter-rater reliability of 6MWT and 10MWT in acute stroke. *Phys Ther Theory Pract.* 2023;39(5):1024-32.
- Youdas JW, Madson TJ, et al. Usefulness of the Trendelenburg test for identification of hip OA. *Phys Ther Theory Pract.* 2010;26(3):184-94.
-