



족부족관절 분야의 환자 자가 보고 결과 평가법

김지연, 김범수*

고신대학교 의과대학 정형외과학교실, *인하대학교 의과대학 정형외과학교실

Patient-Reported Outcome Measures of the Foot and Ankle

Jiyoun Kim, Bom Soo Kim*

Department of Orthopaedic Surgery, Kosin University College of Medicine, Busan,

*Department of Orthopaedic Surgery, Inha University College of Medicine, Incheon, Korea

The patient-reported outcome measures (PROMs) are important in assessing the patient's overall health, lesion-specific outcomes, and condition-specific outcomes. PROMs are also known as "scoring systems" and are usually in questionnaires. There are almost a hundred different PROMs available in foot and ankle surgery. Each PROM has its merits, demerits, and validity. Selecting an appropriate PROM is important for adequately evaluating a patient's health status. This article summarizes the most frequently used PROMs in the literature on foot and ankle surgery and presents the authors' recommendations.

Key Words: Patient reported outcome, Patient reported outcome measures, Scoring system, Foot, Ankle

서론

의료 행위에 대한 가치(value)를 정확하게 평가하는 것은 의학의 발전을 위해 매우 중요한 일이다. 의료에서 가치란 비용(cost) 대비 결과(outcomes)로 산정할 수 있으며, 비용은 낮을수록, 결과는 좋을수록 가치가 높다.¹⁾ 그렇다면 치료의 결과는 어떻게 평가할 것인가. 정형외과 영역에서 치료의 결과를 나타내는 방법으로 방사선학적 지표와 함께 여러 가지 설문조사 결과를 이용하여 산출한 점수들이 이용된다. 흔히 '스코어(score)'로 불리는 이러한 설문 평가 체계는 환자가 본인 상태를 스스로 보고하기 때문에 '환자 자가 보고 결

과(patient-reported outcome, PRO)'라고 하며, PRO를 측정하기 위해 사용되는 설문지 또는 평가 도구를 '환자 자가 보고 결과 평가법(patient-reported outcome measure, PROM)'이라고 한다. 이러한 PROM은 환자가 스스로 인지하는 본인의 건강 상태, 삶의 질(quality of life), 기능 등을 반영하는 유일한 도구일 뿐만 아니라 의료 행위의 결과를 수치화하여 데이터로 활용할 수 있게 한다.

근거중심 의학과 데이터 시대의 바람을 타고 최근 이러한 PROM에 대한 관심이 커지고 있다. 올바른 데이터의 가치를 인식하기 시작하면서 정확한 평가의 중요성에 대한 공감대가 생긴 것이다. 하지만 PROM은 그 종류도 너무 많고 질환에 따라 검증(validation) 여부도 다르기 때문에 질환마다 어떤 PROM을 사용해야 하는지 혼란스럽기 마련이다. 미국정형외과학회(American Academy of Orthopaedic Surgeons, AAOS)에서는 각 부위별 권장 PROM을 발표하기도 하였다.^{2,3)} 하지만 어느 질환이나 적용할 수 있는 완벽한 PROM이 존재하는 것은 아니기 때문에 특정 질환군을 평가하기 위해 어떤 PROM을 사용할지는 결국 연구자 스스로 판단하여 결정해야 한다. 따라서 본 논문에서는 연구자들이 PRO를 평가할 때 고려해야 할 사항 및 족부족관절 영역에서 사용되는 여러 PROM을 정리하고 질환별 권장 PROM을 제시하고자 한다.

Received February 9, 2022 Revised February 19, 2022

Accepted February 19, 2022

Corresponding Author: Bom Soo Kim

Department of Orthopaedic Surgery, Inha University College of Medicine, 100 Inharo, Michuhol-gu, Incheon 22212, Korea

Tel: 82-32-890-2916, Fax: 82-32-890-3099, E-mail: bskim.md@gmail.com

ORCID: <https://orcid.org/0000-0002-9601-396X>

The point of this thesis was presented at the Spring Conference of the Korean Foot and Ankle Society in 2021.

Financial support: None.

Conflict of interest: Bom Soo Kim is the founder of MDdata Solution, Inc., Seoul, Korea. Except for that, no potential conflict of interest relevant to this article was reported.

Copyright © 2022 Korean Foot and Ankle Society.

© This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

PROM의 역사

근거중심 의학이 발전하면서 의료 행위의 평가 방법에 대한 관심이 높아졌고 이에 PROM의 필요성이 대두되었다. 통증에 대한 PROM으로 visual analogue scale (VAS)이 1976년에 제안된 이후⁴⁾ 정형외과 영역과 관련된 PROM에 대한 연구는 1990년대에 이르러 활발해졌다. 족부족관절 영역에 특화된 PROM인 American Orthopaedic Foot and Ankle Society (AOFAS) 점수와 Foot Function Index (FFI)를 비롯하여 전반적인 건강 상태를 평가하는 SF-36 (36-Item Short Form Health Survey), SF-12 (Veterans RAND 12-Item Health Survey) 등 대다수의 PROM이 1990년대에 개발되었다.⁵⁻⁷⁾

PROM의 분류

PROM은 측정하고자 하는 내용에 따라 크게 전반적인 건강 상태를 평가하는 것과 해부학적인 특정 부위나 관절을 평가하는 것, 그리고 어떤 특정 질환이나 수술 기법 등에 특화된 것으로 분류된다.

1. 전반적인 건강 상태 평가도구(overall health measures)

전반적인 건강과 삶의 질에 대해 평가하는 것으로 SF-36, SF-12, EuroQol-5 Dimension (EQ-5D) 등이 있다.^{5,8)} 정신적, 심리적, 사회적인 부분까지 포함하여 환자의 전체적인 건강 상태를 평가한다는 장점이 있지만 근골격계에 특화되어 있지 않기 때문에 특정 부위나 질환별로 필요한 세밀한 평가는 불가능하다.

2. 부위별 평가도구(region specific measures)

신체 각 부위 또는 관절별로 특화된 PROM으로 해당 부위의 기능이나 운동을 보다 정확하게 평가하기 위해 개발된 것이다. 족부족관절 영역에서 자주 사용되는 PROM에는 AOFAS 점수, Foot and Ankle Outcome Score (FAOS), FFI 등이 있으며 2012년부터 2016년 사이에 족부족관절과 관련된 연구에서 사용된 부위별 PROM은 89종이라고 한다.⁹⁾

3. 조건별 평가도구(condition specific measures)

특정 질환이나 수술 방법의 결과 평가에 특화된 평가 도구이다. 대표적인 것으로는 아킬레스건 파열을 평가하는 Achilles total rupture score,¹⁰⁾ 발목 관절염을 평가하는 ankle osteoarthritis scale¹¹⁾ 등이 있다. 여러 종류의 PROM이 있는데 아킬레스건과 관련된 PROM은 21종,¹²⁾ 족관절 전치환술과 관련된 PROM은 11종이 보고되어 있다.¹⁾

PROM의 평가

PROM에 대한 체계적인 검증은 통합 평가 방법을 이용하여 이루어진다.¹³⁾ 이 방법은 PROM의 방법론적 우수성(methodological quality), 심리평가적 타당성(psychometric properties), 그리고 근거 수준(level of evidence)에 대해 각각 평가한 후 이를 종합하여 해석한다. 방법론적 우수성은 COSMIN (Consensus-based Standards for the selection of health Measurement Instruments) 체크리스트¹⁴⁾를 이용하여 평가할 수 있으며 심리평가적 타당성은 Terwee 등¹⁵⁾의 질 평가 기준을, 근거 수준은 Cochrane Back Review Group이 제안한 근거 수준¹⁶⁾ 등을 이용하여 평가할 수 있다.

COSMIN 체크리스트는 PROM에 대한 체계적인 평가를 가능하게 하며 신뢰도(reliability), 타당도(validity), 반응도(responsiveness), 해석 가능성(interpretability)의 네 가지 영역에 대해 세분화하여 평가한다. 신뢰도는 내적 일관성(internal consistency), 신뢰도, 측정 오류(measurement error)로 나누어 평가하며, 타당도는 내용 타당도(content validity), 구성 타당도(construct validity), 기준 타당도(criterion validity)로 나누어진다. 각 영역의 문항을 우수, 좋음, 보통, 나쁨의 네 가지 등급으로 평가한다(Fig. 1).¹⁴⁾ COSMIN 체크리스트 및 관련 정보는 온라인(<https://www.cosmin.nl/>)으로 얻을 수 있다. PROM의 심리평가적 타당성을 평가하기 위해 Terwee 등¹⁵⁾에 의해 개발된 질 평가 기준을 사용할 수 있는데 평가 척도에 내적 일관성, 신뢰도, 측정 오류, 내용 타당도, 기준 타당도, 구조 타당도(structural validity), 가설 검정(hypothesis testing), 반응도가 있다. 그리고 PROM의 근거 수준에 대한 평가는 Cochrane Back Review Group이 제안한 기준으로 평가가 가능하며,¹⁶⁾ 상기 세 가지 영역에 대한 평가 점수를 합산하여 통합 평가할 수 있다.

즉, 좋은 PROM의 기준은 타당도, 신뢰도, 민감도가 높고 사용하기 쉬운 것이어야 한다. 타당도란 측정하고자 하는 것을 얼마나 정확히 측정할 수 있는지를 나타내는 것이고 정확도(accurate)라고도 한

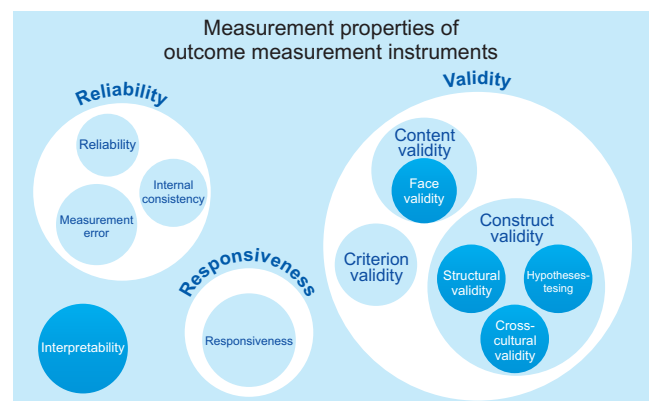


Figure 1. The COSMIN Taxonomy. Measurement properties of outcome measurement instruments.¹⁴⁾

다. PROM을 평가하는 기준 중에 가장 중요한 항목이다. 타당도를 평가할 때 세 가지를 평가하는데, 내용 타당도와 구성 타당도, 그리고 기준 타당도를 평가한다. 내용 타당도란 PROM의 내용이 측정하고자 하는 것을 얼마나 정확히 반영하고 있는가를 나타내며 연구자의 의도에 초점을 둔다. 구성 타당도는 PROM의 구성이 실제 측정하고자 하는 개념을 제대로 측정하는가에 대한 것이며, 추상적인 개념을 계량화하여 검증할 수 있는지를 평가한다. 구성 타당도를 평가할 때 다문화 간 조정(cross-cultural adaptation)이 중요한데, 이는 PROM이 기존의 환경이 아닌 다른 환경에 적용될 때 측정되어야 하는 주요한 속성이며, 기존 PROM을 번역하거나 문화적으로 조정했을 때 얼마나 기존 PROM을 적절히 반영하는가를 평가하는 것이다.¹³⁾ 기준 타당도는 준거 타당도라고도 하며 새로 개발한 PROM의 결과가 이미 타당성이 입증된 기존의 PROM의 결과와 통계적 상관성이 있는지를 평가하며 검사 문항들이 평가 목적에 적절히 부합하는가를 판단한다.

신뢰도란 환자의 상태에 변화가 없을 때 여러 조건하에서 반복 측정하여 동일한 결과가 나오는 정도이며 측정 오차가 얼마나 작은가를 뜻한다. 신뢰도는 내적 일관성, 신뢰도, 측정 오류로 나누어 평가하며 신뢰도와 측정 오류 평가를 위해 시간 경과에 따른 검사-재검사 일관성(test-retest consistency), 동일한 경우에 대한 다른 평가자 간 일관성(inter-rater consistency), 그리고 다른 경우에 대해 단일 평가자 간 일관성(intra-rater consistency)을 평가한다. 또한 신뢰도와 측정 오류를 평가할 때 PROM의 측정 간격(time interval)에 대한 평가를 간과하면 안 되는데, COSMIN 체크리스트에서는 신뢰도의 상승과 측정 오차의 감소를 위해서 PROM의 측정 간격을 2주로 하는 것을 권장하였다.¹⁷⁾

민감도란 측정하고자 하는 대상의 차이 또는 변화를 PROM이 얼마나 잘 감지하는지를 뜻한다. 시간이 경과함에 따라 환자 상태에 주요한 변화가 있을 때 PROM이 그 변화를 반영할 수 있어야 한다. 편의성(easy)도 매우 중요한 요소인데, 문항 수가 지나치게 많다면 환자 편의성이 낮다고 할 수 있다. 또한 평가자도 데이터를 정리하고 분석하기에 용이해야 한다.

족부족관절 분야의 PROM별 사용 빈도

Hunt 등에 따르면 2012년부터 2016년까지 족부족관절 분야 논문에서 가장 빈번하게 사용된 PROM은 AOFAS 점수, VAS, SF-36이며, 이 세 가지 PROM이 전체의 48%를 차지하였다(Fig. 2).²⁾ 다음으로 SF-12와 FAOS가 빈번히 사용되었고 그 외 환자의 만족도와 기능에 관한 맞춤형 설문 평가 체계 등 다양한 PROM이 사용되었다.

족부족관절 분야에서 흔히 사용되는 PROM

1. AOFAS 점수

AOFAS 점수는 족관절-후족부, 중족부, 무지, 소족지의 총 네 가지 유형이 있다. AOFAS 점수는 통증과 기능에 대한 환자의 주관적 평가와 정렬 상태, 안전성, 운동 범위에 대한 의료진의 평가로 구성된다.²⁾

AOFAS 점수는 1994년에 도입된 이후 널리 사용되기 시작했고,⁹⁾ 2002년부터 2011년까지 족부족관절 관련 논문에서 독보적으로 가장 많이 사용되었다.¹⁸⁾ 그러나 2011년 AOFAS는 AOFAS 점수의 타당도와 신뢰도가 낮은 것을 이유로 AOFAS 점수를 권고하지 않는다고 발표하였다.¹⁹⁾ AOFAS 점수는 측정하고자 하는 내용이 분명하지 않은 내용 타당도의 문제와 널리 인정된 SF-36, validated general health score (QUALY), Musculoskeletal Function Assessment (MFA)와의 상관관계가 낮은 기준 타당도의 문제가 있다.^{5,20)} 또한 환자가 아닌 관찰자가 입력하는 신체 검진 영역이 있어 평가자 간 (intra-observer, inter-observer) 오류의 발생 가능성도 있다.¹⁾ 이러한 권고 사항이 발표된 이후 AOFAS 점수의 사용 빈도는 급격히 줄어들고 있는 추세이다. 이미 수집된 과거 데이터를 이용한 논문들이 아직 발표되고는 있지만 향후 전향적 데이터 수집에 AOFAS 점수를 사용하는 것은 지양된다.

2. VAS

VAS는 통증에 대한 PROM이며 1976년에 도입되었다.⁴⁾ VAS는 환자가 스스로 느끼는 통증의 정도를 10 cm 척도에 그림으로 표현

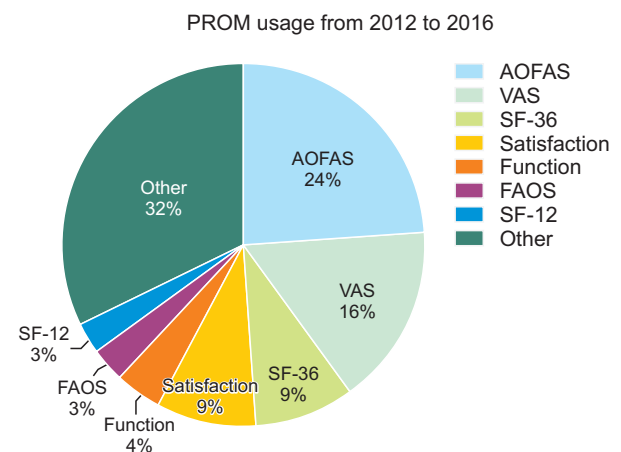


Figure 2. PROM usage from 2012 to 2016. The chart depicts the current distribution of PROM in foot and ankle clinical research.²⁾ PROM: patient-reported outcome measure, AOFAS: American Orthopaedic Foot and Ankle Society Score, VAS: visual analogue scale, SF-36: 36-Item Short Form Health Survey, FAOS: Foot and Ankle Outcome Score, SF-12: Veterans RAND 12-Item Health Survey.

하여 직관적이며 간단하다. 측정 시간이 짧고 측정하기 쉬워 치료 결과를 비교하기에도 적합하다. 이러한 VAS는 근골격계 질환에 광범위하게 사용하기 적합한 것으로 검증되었다.²¹⁾ 하지만 통증의 변화 정도가 선형으로 충분히 반영되지 않는다는 특이도의 문제도 있다.²²⁾

3. SF-36, SF-12

SF-36은 1992년에 개발되었으며 전반적인 건강 상태에 대한 8개 영역, 36개 항목으로 구성된다.⁵⁾ 근골격계 영역에서의 사용 적합성에 대해 검증되었으며 타 PROM을 검증할 때 기준으로 사용되기도 한다.^{23,24)} 하지만 전신적인 건강 상태를 반영하는 만큼 근골격계 영역의 차이를 세밀하고 정확하게 반영하기 어려울 수 있다.

SF-36의 축약형인 SF-12는 1996년에 도입되었으며 불필요한 중복을 줄이고 환자의 부담을 덜기 위해 개발되었다.²⁵⁾ 문항 수를 12개로 줄여 사용의 편리성을 높였으며 바닥 효과 및 천장 효과를 최소화하여 정형외과 환자에게 적합한 것으로 검증되었다.²⁶⁾

SF-36과 SF-12는 모두 훌륭한 평가 도구이며 족부족관절 영역의 PROM에서 큰 비중을 차지하고 있다. 하지만 사용할 때마다 금액을 지불해야 하며 점수를 집계하는 방식이 비공개라는 단점이 있다.²⁷⁾

4. EQ-5D

EQ-5D는 전반적인 건강 상태에 대한 표준화된 척도이다. 최초에 유럽의 5개국이 모여 1987년에 설립한 EuroQol이라는 단체에서 논의하면서 개발이 시작되었고 현재 세계적으로 130개 이상의 언어로 번역되어 대부분의 국가에서 공통으로 사용되고 있다.⁸⁾ 건강과 관련된 삶의 질에 대한 단일 지표값을 제공하며, 운동, 자기 관리, 일상 활동, 통증/불편, 근심/우울에 대한 5개 영역으로 구성된다. 각 영역에 대해 세 가지 수준별 점수(3 levels: no problems, some problems, and extreme problems)를 부여하여 가중치를 곱해 더한 값으로 계산하며 가중치의 추정값은 국가마다 적합한 기준을 제시하여 사용한다.^{8,28)} 다른 PROM에 비하여 상대적으로 평가하기 쉬우며 여러 국가에서 신뢰도와 타당도가 검증되어 있다. 최초의 EQ-5D-3L을 변형한 EQ-5D-5L이 2009년에 개발되어 함께 사용 중이며 EQ-3D-5L은 각 영역에 대해 5가지 수준별 점수(5 levels: no problems, slight problems, moderate problems, severe problems, and extreme problems)로 변경하여 민감도를 높이고 천장 효과를 줄였다.

5. FAOS

FAOS는 5개 영역, 42개 문항으로 구성되어 있으며 족부족관절 분야에서 신뢰도와 타당도가 높다.^{29,30)} 질환별로는 성인의 후천성 편평족, 무지 외반증, 무지 강직증, 인대 재건 등에 대해 검증된 바 있다. 한글 번역본에 대한 검증도 완료되어 사용이 가능하다.³¹⁾

FAOS의 가장 큰 단점은 문항 수가 많기 때문에 환자에게 부담이 크다는 점이며 족관절 관절염에는 검증력이 부족하다는 점도 있다.³⁰⁾

6. FFI

FFI는 1991년에 개발되었으며 통증, 장애 및 활동 제한의 세 가지 영역에서 족부의 기능을 평가하는 23개 문항으로 구성되었다.⁷⁾ 간단한 설문 조사이지만 SF-36과 비교하였을 때 신뢰도 및 반응도가 높은 것으로 검증되었다.²⁾ FFI의 모든 문항은 직선 상에 답을 표시하는 VAS로 평가되며 점수가 높을수록 통증이나 장애가 더 많은 것을 나타낸다. FFI는 류머티즘 관절염과 족관절 관절염 환자의 치료 결과를 평가하기 위해 개발되었으며 족관절 관절염에서 타당도가 검증되었다.³²⁾ FFI도 한글 번역본에 대한 검증이 완료되어 사용이 가능한 상태이다.^{33,34)}

7. FFI-R (revised FFI), FFI-R S (FFI-R short form)

FFI는 통증, 장애, 활동 제한의 세 영역으로 구성되어 있으며 높은 신뢰도와 타당도를 인정받아 세계적으로도 여러 언어로 번역되어 사용되어왔다. 그러나 그 개발 과정에서 사용된 평가 이론의 내재적 제한점과 VAS를 사용하는 평가 방식의 문제점에 대한 보완의 필요성이 대두되었다. 이에 2006년에 FFI-R³⁵⁾로 개정 발표되었는데, 통증, 감각, 장애, 활동 제한, 사회적 문제의 총 5개 영역, 68문항으로 구성되었다. FFI-R은 발의 문제와 연관된 심리적, 사회적 기능과 불편함의 정도까지 평가할 수 있는 포괄적인 평가도구로 발전하였다. FFI-R의 모든 항목은 기존의 VAS식 평가방식이 아닌 사지선다 단답형 평가 방식을 채택하여 응답자에 의한 측정 오류를 줄이고자 하였다. 또한 각 영역의 문항 수를 조정하여 총 34개 문항으로 이루어진 짧은 서식(FFI-R S)도 개발되었으며, 두 서식 모두 정신 건강 평가에 적합한 것으로 보고되었다.³⁶⁾ 다만 국내에서 사용할 수 있는 한글 번역본에 대한 연구가 추가적으로 필요한 상태이다.

8. Foot and Ankle Disability Index (FADI)

FADI는 1999년에 제안되었고 Foot and Ankle Ability Measure (FAAM)의 초기 버전이며 만성 족관절 불안정성에서 민감도와 신뢰도가 높다.³⁷⁾ 2016년에 AAOS에서 족부족관절 영역의 권장 PROM에 포함되어 주목을 받았지만 바닥 효과와 천장 효과가 있는 것으로 보고되었으며³²⁾ 최근에는 자주 사용되지 않는다.²⁾

9. FAAM

FAAM은 2003년에 족부 및 족관절 기능에 장애가 있는 개인의 신체 기능을 평가하기 위해 개발되었으며 일상 기능과 스포츠 기능 모두에 대해 31개 문항으로 평가한다.³⁸⁾ FAAM은 FADI에서 5개 문항을 삭제한 것이며 그 외 내용은 동일하다. 족부족관절 영역에서 광범위하게 검증되었으며 만성 족관절 불안정성과 당뇨병의 PROM으

로 적합하다. FFI, Short Musculoskeletal Function Assessment (SMFA) 및 FAAM은 모두 족부 및 족관절 외상 환자에서 높은 상관관계가 있으며 임상 연구에서는 이 중 한 가지만 사용해도 된다고 한다.²⁾

10. Manchester-Oxford foot questionnaire (MOXFQ)

MOXFQ는 원래 족부 수술을 받은 환자와의 면담 내용을 Manchester Foot Pain and Disability Questionnaire (MFPDQ)에 그려 넣었던 것을 수술적 치료 평가에 적합하게 변형시킨 것이다.³⁹⁾ MOXFQ는 최초로 무지 외반증 수술을 받은 환자들을 평가하기 위해 만들어졌으며 신뢰도, 타당도 및 민감도가 높으며 구조가 훌륭한 것으로 평가되었다. 3개 영역, 16개 문항으로 평가하며 0점에서 4점까지 응답이 가능하다.⁴⁰⁾ 최근에 다양한 족부 및 족관절에 기능 장애가 있는 환자들에서 수술 전후의 삶의 질을 비교할 수 있도록 수정되었으며 타당도가 검증되었다. MOXFQ는 족부족관절 환자에서 일상적으로 사용하기에 적합하며 응답률이 높고 환자가 설문지를 완료하기 쉬운 것으로 나타났다.⁴¹⁾ 아직 한글 번역본과 이에 대한 검증 연구는 이루어지지 않은 상태이다.

11. Ankle Osteoarthritis Scale (AOS)

AOS는 족관절 관절염에 특화하여 간단하면서도 신뢰도와 타당도가 높은 PROM으로 1998년에 개발되었다.¹¹⁾ 총 18개 문항으로 구성되어 있고, 족관절 관절염의 증상과 기능에 대한 평가 도구로 검증되었다.⁹⁾ 그러나 AOS의 심리평가적 타당성을 조사한 몇몇 연구에서 내적 일관성, 신뢰도, 측정 오류, 내용 타당도, 가설 타당도 또는 민감도에 대한 증거가 불충분한 것으로 보고되었다.^{42,43)}

12. Ankle Arthritis Score (AAS)

AAS는 캐나다 족부족관절학회(Canadian Orthopaedic Foot and Ankle Society, COFAS)에서 2017년에 개발하였으며 AOS를 단순화하여 검증한 뒤 발표하였다.⁴⁴⁾ AAS는 말기 족관절 관절염 환자에 특화되어 인공관절 치환술 또는 관절 유합술을 시행 받은 후 결과를 평가하는 PROM으로 개발되었다. AAS는 기본 활동(basic activity)과 고급 활동(advanced activity)의 두 영역, 8가지 질문으로 구성되며 검증에 대한 후속 연구가 발표되고 있다.⁴⁵⁾ 아직 한글 번역본과 이에 대한 검증 연구는 이루어지지 않은 상태이다.

13. Achilles Tendon Total Rupture Score (ATRS)

ATRS는 아킬레스건 완전 파열에 특화되어 환자의 치료 결과의 평가를 위해 2007년에 개발되었으며¹⁰⁾ PROM 검증의 주요한 측면들에 대해 후속 데이터들이 보고되어 있다. 그리고 전문가 의견만 반영한 타 아킬레스건 관련 PROM과는 대조적으로 적절한 방법론을 사용하여 개발되고 평가된 유일한 PROM이다.¹²⁾ 최근에 ATRS의 한글

번역본과 이를 검증한 연구가 발표되었다.⁴⁶⁾

14. OMAS (Olerud Molander Ankle Score)

OMAS는 1984년에 족관절 골절 후 기능적 치료 결과를 평가하기 위해 개발된 PROM이다.⁴⁷⁾ 9개의 질문으로 구성되어 있으며 족관절 골절에 특화되어 있으나 아직 한글 번역본에 대한 연구는 진행되지 않았다.

권장 PROM

기술한 바와 같이 PROM은 그 종류가 많을 뿐만 아니라 각각의 장단점이 다르며 부위별, 질환별 검증 여부에 대한 자료가 방대하여 특정 PROM을 선정하기에 어려움이 있다. 또한 동시대 연구자들에 의한 어느 정도의 합의(consensus)가 이루어져야 하는 부분도 있다. 어떤 PROM을 사용하는지는 연구자의 결정이지만 권장하는 PROM을 참조하는 것은 도움이 될 것이다.

1. AAOS 권장 PROM

AAOS에서는 PROM의 조사 및 평가를 진행하여 권장 PROM 목록을 제작하였으며 2016년에 이사회에서 이를 승인하였다. 이후에도 계속 개정을 거듭하고 있으며 2021년 12월 현재 전반적인 삶의 질과 족부족관절 영역에 관해 권장하는 PROM은 아래와 같다 (Table 1).⁴⁸⁾ 이에 대한 자세한 정보는 온라인(<https://www.aaos.org/quality/research-resources/patient-reported-outcome-measures>)으로 얻을 수 있다.

AAOS에서 사용한 기준은 누구나 자유롭게 무료로 사용이 가능해야 하며, 환자가 직접 입력해야 하고(임상의의 개입 배제), 질문 개수가 20개 이하여야 하며, 임상적으로 의미가 있어야 하고(민감도), 세 가지 이상의 관절이나 질환을 다루면 안 된다는 것이었다. 또 입력 방법이 다양해야 하고(전자식, 지면, 인터넷), CAT (computerized adaptive test) 버전 사용이 가능해야 한다.

Table 1. American Academy of Orthopaedic Surgeons (AAOS) Recommended Patient Reported Outcome Measures for Foot and Ankle Surgeons

Treatment outcome/quality of life	Foot and ankle
Single Assessment Numeric Evaluation (SANE)	Foot and Ankle Ability Measure (FAAM)
Veteran: RAND 12	Foot and Ankle Disability Index (FADI)
PROMIS: (PROMIS 10 or CAT)	
EuroQol 5D (EQ-5D)	

Data from American Academy of Orthopaedic Surgeons (AAOS). Patient reported outcome measures. Quality Measures [Internet]. Available from: <https://www.aaos.org/quality/research-resources/patient-reported-outcomemeasures>.⁴⁸⁾

PROMIS: Patient-Reported Outcomes Measurement Information System, CAT: computerized adaptive test.

Table 2. Recommendation of Patient Reported Outcome Measures for Foot and Ankle Surgeons

Overall health	Region specific	Condition specific	
EQ-5D	MOXFQ	Hallux valgus	MOXFQ
SF-12	FAAM	Ankle instability	FAAM/FFI-R
	FFI-R	Ankle arthritis	AAS
	FAOS*	Achilles rupture	ATRS*
		Fracture	OMAS
		Diabetes mellitus	FAAM

EQ-5D: EuroQol-5 Dimension, SF-12: Veterans RAND 12-Item Health Survey, MOXFQ: Manchester-Oxford foot questionnaire, FAAM: Foot and Ankle Ability Measure, FFI-R: Foot Function Index-Revised, FAOS: Foot and Ankle Outcome Score, AAS: Ankle Arthritis Score, ATRS: Achilles Tendon Total Rupture Score, OMAS: Olerud Molander Ankle Score.

*The PROMs with validated Korean versions.^{31,33,34,46)}

2. 저자 권장 PROM

저자들이 본 논문을 준비하면서 리뷰한 내용을 바탕으로 검증되고 추천할 만한 PROM을 정리하면 Table 2와 같다. 기본적으로 VAS와 함께 전반적인 건강 상태 평가 도구 중에서 하나와, 족부족관절 부위에 비교적 두루 사용할 수 있는 것 중에서 1~2개를 선정하여 사용하는 것이 권장된다. 여기에 조건에 따라 무지외반증, 족관절 불안정성, 족관절 관절염, 아킬레스건 파열, 골절의 경우는 각 조건에 특화되거나 특별히 검증된 평가자들을 조합하여 사용하면 더욱 좋을 것으로 생각된다(Table 2). 단, 아직 한글 버전의 PROM이 검증된 것은 FAOS, FFI, ATRS뿐이다.

전자식 PROM

종이 설문지를 이용하는 방식과 달리 최근에는 PC나 태블릿, 또는 스마트폰을 이용한 전자식 평가 방식(electronic PROM, ePROM)을 사용하기도 한다. AAOS에서는 좋은 PROM의 기준으로 전자식으로 사용이 가능할 것을 꼽고 있다. 아무리 훌륭한 PROM이라도 환자가 사용하기에 편리해야 하고 평가자가 분석하기에 용이해야 하는데, 이러한 편의성은 종이에 비해 전자식이 월등히 높기 때문이다. 종이를 사용하는 경우 설문 진행, 설문지 수거 및 데이터 입력, 결과값 계산, 설문지 보관 및 관리에 많은 시간과 노동력이 소요되는데, 현실적으로 인력과 시간이 부족한 대부분의 진료 환경에서 이 모든 과정을 제대로 수행하는 것은 매우 어렵다. 이에 비해 ePROM은 태블릿이나 스마트폰으로 전송된 설문지에 환자가 스스로 답지를 클릭하면 모든 데이터 입력 과정이 완료될 뿐만 아니라 계산 등의 과정이 자동으로 이루어지기 때문에 시간이 절약되고 데이터 관리가 수월해진다.^{49,50)} 또한 환자의 회복 추이나 통계 분석이 실시간으로 이루어지기 때문에 활용도가 훨씬 높다.

국내에서 사용 가능한 한글 평가지가 제공되는 ePROM 플랫폼으로는 PROscore (www.proscore.kr, MDdatasolution, Inc.,

Seoul, Korea)가 있다. PROscore는 원하는 평가지와 설문 일정을 미리 예약해두면 해당일에 환자의 모바일 기기에 자동 전송된 평가지를 어디서든 작성 가능하기 때문에 원격 데이터 수집도 가능하다. 폴더에 진단명이나 수술명으로 필터링 조건을 설정해 놓으면, 조건에 해당되는 케이스들이 자동으로 폴더에 포함되어 사용하기 편리하고, 폴더 내의 모든 데이터는 엑셀로 다운로드할 수 있어 자료 수집 및 분석에 유리하다.

결론

의학은 과학이며 진리 탐구를 위한 연구는 정확한 데이터에 기반한다. 따라서 의학을 발전시키고 인류를 이롭게 하고자 하는 모든 연구자는 데이터의 정확도에 심혈을 기울여야 할 것이다. 환자가 자신의 상태에 대해 직접 평가하고 알려주는 PROM 데이터는 매우 소중한 자산이다. 따라서 환자나 질환에 따라 평가하고자 하는 것을 정확히 평가할 수 있는 적절한 PROM을 선정하여 사용해야 한다.

안타깝게도 족부족관절 PROM의 대부분은 아직 한글 번역본이 검증되지 않은 실정이다. 이로 인해 원어 설문지를 기관마다 자체적으로 한글로 번역하여 사용하는 경우가 많은데, 이러한 임의 번역본의 사용은 지양해야 하며 권장 PROM에 대한 한글 번역본에 대한 검증 연구가 속히 이루어져야 할 것이다.

아울러 정확한 데이터를 편리하게 수집하고 효율적으로 관리하기 위해 ePROM의 사용을 적극적으로 고려할 것을 권장한다.

ORCID

Jiyoun Kim, <https://orcid.org/0000-0002-4239-3848>

REFERENCES

1. Naal FD, Impellizzeri FM, Rippstein PF. Which are the most frequently used outcome instruments in studies on total ankle arthroplasty? *Clin Orthop Relat Res.* 2010;468:815-26. doi: 10.1007/s11999-009-1036-y.
2. Hunt KJ, Lakey E. Patient-reported outcomes in foot and ankle surgery. *Orthop Clin North Am.* 2018;49:277-89. doi: 10.1016/j.jocl.2017.11.014.
3. American Academy of Orthopaedic Surgeons. Instruments for collection of orthopaedic quality data. Rosemont: American Academy of Orthopaedic Surgeons; 2016.
4. Scott J, Huskisson EC. Graphic representation of pain. *Pain.* 1976;2:175-84. doi: 10.1016/0304-3959(76)90113-5.
5. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care.* 1992;30:473-83. doi: 10.1097/00005650-199206000-00002.
6. Kitaoka HB, Alexander IJ, Adelaar RS, Nunley JA, Myerson MS, Sanders M. Clinical rating systems for the ankle-hindfoot, mid-

- foot, hallux, and lesser toes. *Foot Ankle Int.* 1994;15:349-53. doi: 10.1177/107110079401500701.
7. Budiman-Mak E, Conrad KJ, Roach KE. The Foot Function Index: a measure of foot pain and disability. *J Clin Epidemiol.* 1991;44:561-70. doi: 10.1016/0895-4356(91)90220-4.
 8. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. *Health Policy.* 1990;16:199-208. doi: 10.1016/0168-8510(90)90421-9.
 9. Lakey E, Hunt KJ. Patient-reported outcomes in foot and ankle orthopedics. *Foot Ankle Orthop.* 2019;4:2473011419852930. doi: 10.1177/2473011419852930.
 10. Nilsson-Helander K, Thomeé R, Silbernagel KG, Thomeé P, Faxén E, Eriksson BI, et al. The Achilles tendon Total Rupture Score (ATRS): development and validation. *Am J Sports Med.* 2007;35:421-6. doi: 10.1177/0363546506294856. Erratum in: *Am J Sports Med.* 2011;39:NP8.
 11. Domsic RT, Saltzman CL. Ankle osteoarthritis scale. *Foot Ankle Int.* 1998;19:466-71. doi: 10.1177/107110079801900708.
 12. Kearney RS, Achten J, Lamb SE, Plant C, Costa ML. A systematic review of patient-reported outcome measures used to assess Achilles tendon rupture management: what's being used and should we be using it? *Br J Sports Med.* 2012;46:1102-9. doi: 10.1136/bjsports-2011-090497.
 13. Jia Y, Huang H, Gagnier JJ. A systematic review of measurement properties of patient-reported outcome measures for use in patients with foot or ankle diseases. *Qual Life Res.* 2017;26:1969-2010. doi: 10.1007/s11136-017-1542-4.
 14. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol.* 2010;63:737-45. doi: 10.1016/j.jclinepi.2010.02.006.
 15. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34-42. doi: 10.1016/j.jclinepi.2006.03.012.
 16. van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration back review group. *Spine (Phila Pa 1976).* 2003;28:1290-9. doi: 10.1097/01.BRS.0000065484.95996.AF.
 17. 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. *Qual Life Res.* 2018;27:1147-57. doi: 10.1007/s11136-018-1798-3.
 18. Hunt KJ, Hurwit D. Use of patient-reported outcome measures in foot and ankle research. *J Bone Joint Surg Am.* 2013;95:e118(1-9).
 19. Pinsker E, Daniels TR. AOFAS position statement regarding the future of the AOFAS Clinical Rating Systems. *Foot Ankle Int.* 2011;32:841-2. doi: 10.3113/FAI.2011.0841.
 20. Malviya A, Makwana N, Laing P. Correlation of the AOFAS scores with a generic health QUALY score in foot and ankle surgery. *Foot Ankle Int.* 2007;28:494-8. doi: 10.3113/FAI.2007.0494.
 21. Ferreira-Valente MA, Pais-Ribeiro JL, Jensen MP. Validity of four pain intensity rating scales. *Pain.* 2011;152:2399-404. doi: 10.1016/j.pain.2011.07.005.
 22. Kersten P, White PJ, Tennant A. Is the pain visual analogue scale linear and responsive to change? An exploration using Rasch analysis. *PLoS One.* 2014;9:e99485. doi: 10.1371/journal.pone.0099485.
 23. Busija L, Osborne RH, Nilsdotter A, Buchbinder R, Roos EM. Magnitude and meaningfulness of change in SF-36 scores in four types of orthopedic surgery. *Health Qual Life Outcomes.* 2008;6:55. doi: 10.1186/1477-7525-6-55.
 24. Laucis NC, Hays RD, Bhattacharyya T. Scoring the SF-36 in orthopaedics: a brief guide. *J Bone Joint Surg Am.* 2015;97:1628-34. doi: 10.2106/JBJS.O.00030.
 25. Jenkinson C, Layte R, Jenkinson D, Lawrence K, Petersen S, Paice C, et al. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *J Public Health Med.* 1997;19:179-86. doi: 10.1093/oxfordjournals.pubmed.a024606.
 26. Gosling CM, Gabbe BJ, Williamson OD, Sutherland AM, Cameron PA. Validity of outcome measures used to assess one and six month outcomes in orthopaedic trauma patients. *Injury.* 2011;42:1443-8. doi: 10.1016/j.injury.2011.05.022.
 27. Hiji FY, Schneider AD, Pyper M, Laughlin RT. The popularity of outcome measures used in the foot and ankle literature. *Foot Ankle Spec.* 2020;13:58-68. doi: 10.1177/1938640019826680.
 28. Brooks R. EuroQol: the current state of play. *Health Policy.* 1996;37:53-72. doi: 10.1016/0168-8510(96)00822-6.
 29. Mani SB, Brown HC, Nair P, Chen L, Do HT, Lyman S, et al. Validation of the Foot and Ankle Outcome Score in adult acquired flatfoot deformity. *Foot Ankle Int.* 2013;34:1140-6. doi: 10.1177/1071100713483117.
 30. Golightly YM, Devellis RF, Nelson AE, Hannan MT, Lohmander LS, Renner JB, et al. Psychometric properties of the foot and ankle outcome score in a community-based study of adults with and without osteoarthritis. *Arthritis Care Res (Hoboken).* 2014;66:395-403. doi: 10.1002/acr.22162.
 31. Lee KM, Chung CY, Kwon SS, Sung KH, Lee SY, Won SH, et al. Trans-cultural adaptation and testing psychometric properties of the Korean version of the Foot and Ankle Outcome Score (FAOS). *Clin Rheumatol.* 2013;32:1443-50. doi: 10.1007/s10067-013-2288-1.
 32. Shazadeh Safavi P, Janney C, Jupiter D, Kunzler D, Bui R, Panchbhavi VK. A systematic review of the outcome evaluation tools for the foot and ankle. *Foot Ankle Spec.* 2019;12:461-70. doi: 10.1177/1938640018803747.
 33. In TS, Jung JH, Kim K, Jung KS, Cho HY. The reliability and validity of the Korean version of the foot function index for patients with foot complaints. *J Phys Ther Sci.* 2017;29:53-56. doi: 10.1589/jpts.29.53.
 34. Huh JW, Eun IS, Ko YC, Park MJ, Hwang KM, Park SH, et al. Reliability and validity of the Korean version of the Foot Function Index. *J Foot Ankle Surg.* 2016;55:759-61. doi: 10.1053/j.jfas.2016.03.011.
 35. Budiman-Mak E, Conrad K, Stuck R, Matters M. Theoretical model and Rasch analysis to develop a revised Foot Function Index. *Foot Ankle Int.* 2006;27:519-27. doi: 10.1177/107110070602700707.
 36. Budiman-Mak E, Conrad KJ, Mazza J, Stuck RM. A review of the foot function index and the foot function index - revised. *J Foot Ankle Res.* 2013;6:5. doi: 10.1186/1757-1146-6-5.
 37. Hale SA, Hertel J. Reliability and sensitivity of the Foot and Ankle Disability Index in subjects with chronic ankle instability. *J Athl Train.* 2005;40:35-40.

38. Martin RL, Irrgang JJ, Burdett RG, Conti SF, Van Swearingen JM. Evidence of validity for the Foot and Ankle Ability Measure (FAAM). *Foot Ankle Int.* 2005;26:968-83. doi: 10.1177/107110070502601113.
39. Dawson J, Coffey J, Doll H, Lavis G, Cooke P, Herron M, et al. A patient-based questionnaire to assess outcomes of foot surgery: validation in the context of surgery for hallux valgus. *Qual Life Res.* 2006;15:1211-22. doi: 10.1007/s11136-006-0061-5.
40. Dawson J, Doll H, Coffey J, Jenkinson C. Responsiveness and minimally important change for the Manchester-Oxford foot questionnaire (MOXFQ) compared with AOFAS and SF-36 assessments following surgery for hallux valgus. *Osteoarthritis Cartilage.* 2007;15:918-31. doi: 10.1016/j.joca.2007.02.003. Erratum in: *Osteoarthritis Cartilage.* 2011;19:920.
41. Venkatesan M, Aziz S, Mahadevan D, Korim T, Dias J, Bhatia M. Evaluation of validity, reliability, and feasibility of 3 patient-reported foot and ankle outcome measures. *J Foot Ankle Surg.* 2020;59:507-12. doi: 10.1053/j.jfas.2019.09.033.
42. Pinsker E, Inrig T, Daniels TR, Warmington K, Beaton DE. Reliability and validity of 6 measures of pain, function, and disability for ankle arthroplasty and arthrodesis. *Foot Ankle Int.* 2015;36:617-25. doi: 10.1177/1071100714566624.
43. Madeley NJ, Wing KJ, Topliss C, Penner MJ, Glazebrook MA, Younger AS. Responsiveness and validity of the SF-36, Ankle Osteoarthritis Scale, AOFAS Ankle Hindfoot Score, and Foot Function Index in end stage ankle arthritis. *Foot Ankle Int.* 2012;33:57-63. doi: 10.3113/FAI.2012.0057.
44. Wing KJ, Chapinal N, Coe MP, Daniels TR, Glazebrook M, Dryden P, et al. Measuring the operative treatment effect in end-stage ankle arthritis: are we asking the right questions? A COFAS multicenter study. *Foot Ankle Int.* 2017;38:1064-9. doi: 10.1177/1071100717714953.
45. Croft S, Wing KJ, Daniels TR, Glazebrook M, Dryden P, Younger A, et al. Association of Ankle Arthritis Score with need for revision surgery. *Foot Ankle Int.* 2017;38:939-43. doi: 10.1177/1071100717712891.
46. Park YH, Cho HW, Choi JW, Kim HJ. Validation and cross-cultural adaptation of the Korean translation of the Achilles tendon Total Rupture Score. *BMC Musculoskelet Disord.* 2021;22:876. doi: 10.1186/s12891-021-04765-w.
47. McKeown R, Rabi AR, Ellard DR, Kearney RS. Primary outcome measures used in interventional trials for ankle fractures: a systematic review. *BMC Musculoskelet Disord.* 2019;20:388. doi: 10.1186/s12891-019-2770-2.
48. American Academy of Orthopaedic Surgeons (AAOS). Patient reported outcome measures. Quality Measures [Internet]. Rosemont (IL): AAOS [cited 2021 Dec 30]. Available from: <https://www.aaos.org/quality/research-resources/patient-reported-outcome-measures>.
49. Lee DY, Kim YM, Lee JH, Kim J, Kim JB, Kim BS, et al. Validation of electronic foot function index in patients with foot and ankle disease: a randomized, prospective multicenter study. *J Korean Foot Ankle Soc.* 2019;23:24-30. doi: 10.14193/jkfas.2019.23.1.24.
50. Park JY, Kim BS, Lee HJ, Kim YM, Kim HN, Kang HJ, et al. Comparison between an electronic version of the foot and ankle outcome score and the standard paper version: a randomized multicenter study. *Medicine (Baltimore).* 2019;98:e17440. doi: 10.1097/MD.00000000000017440.