

요통의 기능 평가방법에 대한 고찰

윤유석 · 문상현 · 류한진 · 이종수

경희대학교 한의과대학 한방재활의학과학교실

The Study Of Functional Disability Scales for Low Back Pain

You-Suk Youn, O.M.D. Sang-Hyun Moon, O.M.D., Han-Jin Ryu, O.M.D., Jong-Soo Lee, O.M.D.

Dept. of Oriental Rehabilitation Medicine, College of Oriental Medicine, Kyung Hee University

Objectives :

The object is to know the difference among the funtional disability scale for low back pain and to use standard and meaningful forms of assessment

Methods :

I researched the contents of funtional disability scale for low back pain

Results & Conclusions:

1. The funtional disability scale for low back pain is specific-condition and general-health
2. A proposal of set core of funtional disability scale for low back pain is back specific function, generic health state, pain, work disability, satisfaction
3. Oswestry disability index or Rolland disability questionnaire wase used for standard specific-condition scale and SF-36 was used for standard general-health scale
4. SF-36 is able to use the funtional disability scale for low back pain independence
5. In the future, scale for low back pain shulod be standardized in multiple dimension by computer

Key Words : Functional disability, Low back pain

I. 서론

요통은 인류의 역사와 더불어 생겨난 것이며 일어서서 두발로 걸어다녀 생리학적인 부담을 허리에 받게 되어 발생한다. 요통은 국제적인 관심사가 되어 있으며 전 인류의 약 80%는 일생동안 한번 이상의 요통을 경험하는 것으로 되어있다¹⁾.

요통의 평가하기 위해서 능동, 수동 관절가동범위, 이학적 검사, 영상 검사, 통증의 정도 등을 평가한다. 그러나 근육의 길이나 척추의 움직임도 환자가 생각하는 것과 상관이 없는 경우도 흔하며, 통증의 하나만으로도 요통의 기능과 연관시키는 것은 무리이다. 따라서 Waddell, Main²⁾이 통증, 신체적 장애, 장애에 의해서 요통을 평가하기 시작한 이후로 여러 평가법이 등장하게 되었다.

■ 교신저자 : 윤유석, 서울 동대문구 회기동 1 경희의료원 한방재활의학과(130-702)
Tel : (02) 958-9226, Fax : (02) 963 -4983, E-mail : fourtype@empal.com

최근 20여년동안 요통의 평가의 연구에 있어서 근력이나 운동의 범위 같은 객관적인 지표보다도 환자 자신의 기능적인 평가(self-report of physical function)에 중점을 두고 있다³⁾. 이는 신체적 장애의 정도가 평가자의 검사만으로 환자의 기능 제한에 충분히 설명하지 못하여서 나타나는 현상이기도 하다⁴⁾.

이러한 기능적인 평가법은 특정 질병에 특징적인(specific) 척도와 전체적(general)인 척도로 나눌 수 있다. 이에 저자는 요통 환자의 기능적 평가방법도 2가지로 나누어 비교 고찰하고자 한다.

II. 본 론

1. 요통의 기능적 평가법의 구성

요통의 기능적 평가법의 구성에는 각 scale마다 통증의 강도, 특정한 행위의 정도, 정신상태, 직업유지유무 등 여러 항목이 각 평가법마다 다르게 구성되어 있으나 1998년 Spine의 추천한 통증증상(Pain symptom), 기능(function), 좋은상태(well-being), 장애(disability), 치료의 만족도(satisfaction with care)⁵⁾를 들 수 있다(Table I).

2. 요통의 기능적 평가법의 종류

요통의 평가의 척도는 두 가지로 대별할 수 있다. 요통에 보다 특징적인(specific) 척도와 전체적인 건강상태(generic) 척도로 나눌 수 있다. 특징적인 척도는 요통의 기능적 평가를 위해서 만들어진 평가법이며, 전체적인 척도는 신체적 기능(physical function), 정신건강(mental health), 활동제한(role limitation) 등의 여러 변수를 모두 평가하는 것이다. 본 고찰에서는 검증을 획득하여 타당도, 신뢰도, 민감도를 인정받은 평가법을 중심으로 고찰하여 보고 이들을 비교하여 보고자 한다.

1) ODI (The Oswestry Disability Index)

Oswestry Disability Index(이하 ODI)는 1976년 O'Brien에 의해 시작되어 1980년 Fairbank 등에 의해 발전된 것으로 일상 생활의 각각의 동작과 관련된 10개의 항목으로 구성되어 있다. 평가되는 항목은 통증 정도, 수면 방해, 자가 치료, 걷기, 앉기, 서 있기, 옮기기, 성생활, 사회적인 활동, 여행 등이 포함되어 있다. 검사 시간은 5분 정도이며 각 항목에서 장애를 0-5점으로 6가지 단계로 기술한다. 이 장애

Table I. A proposed 'Core Set' of Six Questions, Practical for Routine Clinical Use, Quality Improvement, and as a Component of More Formal Research

Domain	Specific instrment
Pain symptom	Bothersomeness or severity and frequency of low back pain and leg pain(sciatica)
Back related function	Roland-Morris Disability Questionnaire(or adaptation) or Oswestry Disability Index(or adaptation)
Generic well being	SF-12 or EuroQoL: also "If you had to spend the rest of your life with the symptoms you have right now, how would you feel about it?"
Disability(sacial role)	Days of work absenteeism, cut down activities, bed rest
Satisfaction with care	Single question on overall satisfaction(optional)

지표는 총 점수를 답한 항목의 수로 나누고 100을 곱하여 계산하므로 점수가 높을수록 장애가 많은 것을 의미한다. 지표에 대한 타당도, 민감도, 신뢰도 연구도 여러 논문에서 시행되어 흔히 쓰이고 있는 평가방법으로 1980년 이후 개발된 이후로 200여편의 발표된 논문에서 사용되었다⁶⁾. 그에 따라 척도의 정확성을 위해서 1988년, 1994년, 1995년, 1996년, 1999년에 각각 7편에 새로운 버전이 나왔으며 국제적으로 번역되어 여러 나라에서 활용되어 지고 있다.

2) RDQ(Roland Disability Questionnaire, 이하RDQ)

요통으로 인한 신체적 장애를 평가하기 위해 1983년 Roland와 Morris 의해 고안되어 환자들의 임상적 효용성을 인정받아 지금에 이르고 있다. 136가지의 항목을 가지고 있는 Sickness Impact Profile(SIP)의 항목중에서 신체적, 정신적 기능에 해당된다고 여겨지는 24가지 항목을 선택하여 만든 검사법으로 각 검사항목에 '요통 때문에(because of my back pain)'를 추가하였다(Table II)⁷⁾. yes/no의 대답형식으로 구성되어 해당되는 항목에 대답하는 형식으로 되어있어 각각의 항목을 더하여 점수는 0(no disability)에서 24(severe disability)까지 다양하다. 처음 버전은 12개국에 번역되어 활용되고 있다. Patrick 등⁸⁾은 23개의 항목을, Stratford 등⁹⁾에 18개의 항목을 제거하였으며, Andrew M. 등은 2003년 Rasch analysis에 의해 24개의 항목 중에서 4가지 항목을 빼는 것을¹⁰⁾ 제거하였으나 추후의 논문은 없었다.

3) Waddle Disability Index

Waddle Disability Index는 무거운 것 옮기기, 앉기, 걷기, 서있기, 사회 활동, 여행, 성생활, 양말 신기 등의 아홉 개의 주요 요소들이 제시되었는데 예를 들면 '양말을 신는데 대개 도움이 필요함' 등이다. 이러한 각각의 요소에 대하여서는 '이런 행동이 감소

하였는지 아니면 이런 행동을 하는데 도움을 필요로 하는지'를 물어보게 된다. 두 명의 검사자가 30명의 환자를 각각 인터뷰한 결과 높은 관찰자 상호간의 일치도를 나타내었으며 일상적인 용도를 위해서는 아홉가지 중 긍정적인 항목만 더한 것만으로도 만족스러운 결과를 보였다¹¹⁾. 임상가가 요통 환자를 자주 보지 못하고 처음 평가로 초진인 경우에 자주 쓰이며, 최근 신뢰성 확보하고자 하는 논문의 발표는 없는 상황이다. 2002년 Davidson⁸⁾ 등에 의한 연구에 있어서도 Waddell Disability Index는 5가지의 평가법중 신뢰도의 평가가 임상에 활용될 수 있다고 평가 받았으나 다른 3가지의 척도인 Modified ODQ, SF-36 physical functional scale, Quebec보다 뒤떨어져 활용가치가 떨어진다고 여겨진다.

4) Million Index

Million Index는 VAS로 기록되고, 요통의 심한 정도를 반영하는 15개의 주관적인 변수(주관적인 부분)와 하지 직거상 검사를 단계별로 측정하는 것에서부터 센티미터로 기록되는 요추의 굴곡을 포함하는 10개의 객관적인 측정(객관적인 부분)의 두 가지 부분으로 구성되어 있다. 하지 직거상 검사라는 이학적 검사가 들어가 있는 점이 다른 검사와 다르다. 주관적인 변수로는 통증 강도, 일상 생활의 방해 정도, 일의 방해 정도, 그리고 전반적인 장애를 평가한다¹²⁾.

5) Resumption of Activities of Daily Living Scale

Resumption of Activities of Daily Living Scale은 환자가 손상을 받은 이후의 일상 생활의 회복 정도를 평가한다. 평가하는 활동은 수면, 성생활, 자가 치료, 집안 허드렛일, 쇼핑, 사회화, 여행, 레크레이션활동, 직업활동 등을 포함한다(Table II). 각각의 항목은 0%에서 100%의 그래프 척도가 있다.

Table II. Summary of Functional Status Questionnaire with Low Back Pain

Name	Number of items; response formats	subscales	Self-administered formats
Daily Activity Diary	Position and location in 1/2 hour time block for entire time	Activity; ordinal pain, tension and mood scales	Yes
Million Index	15 questions, visual analog; 10 physical measurement	Subjective: no subscales; questions concern pain with several activities Objective: SLR, flexion, extension in various planes	No
Sickness Impact Profile	136 items yes/no	Ambulation Body care/movement Mobility Emotional behavior Social behavior Alertness behavior Communication Work Sleep and rest Eating Home management Recreational activities	Yes
Waddell Disability Index	9 items yes/no	Items cover lifting sitting, standing walking, sleep social activity sexual activity dressing	?

6) Low Back Outcome Scale

Low Back Outcome Score는 요통을 빠르게 실용적으로 설문지를 사용하여 통증과 장애를 평가하는 방법이다. 기능적인 장애와 통증이 오지 않는 범위 내에서 수동적 움직임을 평가하는 것이 특징으로 수동적인 평가로 인해 병의 심각도와 치료의 범위를 정하는데 장점이 있다. 요통 환자에게 포괄적인 평가 체계로서 의도된 13개의 항목으로 구성되어 있다. 이것은 현재의 통증, 고용상태, 가사 및 스포츠활동, 약이나 의료서비스의 사용, 휴식, 성생활, 일상생활과 관계된 질문에 무게를 두고 구성되어 있다. 처음에는 직장인의 평가로 시작되어 주부, 학생, 은퇴한 개인 등의 평가를 위해 항목이 추가되었다.

7) Dallas Pain Questionnaire

통증의 강도, 감정 손상, 장애의 정도. 일상 생활

의 평가 등 만성 통증의 평가를 위해서 개발되어진 척도이다. 만성 요통의 기능적, 정서적인 면을 모두 고려한 척도법¹³⁾으로 16개의 항목으로 구성되어진 시각적 상사척도이며(Table III). 만성 통증을 평가하므로 작업장에서 근무를 여전히 하는 사람을 대상으로 평가하는 장점이 있다. 처음의 10가지 항목은 기능적인 면을 평가하고, 나머지 6개가 정서적인 면을 평가한다. 일상 생활(daily activities), 직업 활동(work activities), 두려움과 우울(anxiety-depression), 사회적 관심(social interest) 네 가지 영역에서 발전되어 왔다. Lawlis¹⁴⁾ 등은 143개의 세부 항목으로 나누어 기능적인 면과 정서적인 면으로 나누어 발전시켰으며, Thomsen¹⁵⁾ 등은 척추 수술의 장기적인 평가로 사용하였으며, 1998년에는 Marty 등에 의해 프랑스어로 번역되어 평가되어 활용되고 있다.

Table. III Summary of Dallas Pain Questionnaire(1989), Low Back outcome score(1991)

Name	Format(scale)	Subscale(Item)
Dallas Pain Questionnaire	4 scales 16 Items VAS	Daily activities(pain and intensity, personal care, lifting, walking, sitting, standing, sleeping) Work and leisure activities(social life, traveling, vocational) Anxiety/depression(Anxiety and mood, emotional control, depression) Social Interest(interpersonal relationship, social support, punishing responses)
Low Back Outcome Score	13 Scales/Items, 3-4 options	Current pain, employment, domestic chores and 'odd jobs', sports/active social(dancing), resting, treatment or consultation, anagelsia, sex life, sleeping, walking, sitting, travelling, dressing

8) Aberdeen Low Back Pain Scale(ABPS)

APBS는 임상 의 관찰속에서 등장한 평가법으로 요통 환자의 건강 상태를 통증, 이학적인 장애, 그리고 기능적인 장애의 차원에 걸쳐 평가하기 위한 19 가지 항목의 설문지이다. duration of pain, areas of pain, duration of anagelsia, effect of lying down, areas of weakness, loss of feeling, bending, sleep, work, confinement to bed, sex life, leisure, self-care 등으로 구성되어 있으며, 이는 요통을 치료하는 정형외과의사, 류마토이드의사, 물리치료사 등에 의해서 쓰여지고 있다. 질문에 대한 답은 단수의 선택과 복수의 선택의 문항이 나누어져 있으며 19가지의 답을 합하여 요통의 심각성을 0에서 100까지의 백분율로 변환하는데 0이 장애가 가장 적은 것이다. Kopec¹⁶⁾, Rurra¹⁷⁾ 등에 의해서 타당성, 신뢰성, 민감성을 인정 받아왔으며, 심리적, 임상적 시험을 거친 척도이다.

9) Quebec Back Pain Disability Scale

Quebec Back Pain Disability Scale은 급성 혹은 만성 의 요통 환자를 평가하기 위하여 고안된 척도로서 기능적인 장애를 평가하기 위해 개발되었다. 이 척도는 요통과 관련된 20개의 항목으로 구성되

어 get out of bed, sleep through the night, turn over in bed, ride in a car, stand for 20 to 30 minutes, sit a chair for severl hours, climb one flight of stairs, walk s few blocks, walk several miles, reach up to high shelves, throw a ball, run one block, take food out of the refrigator, make your bed, put on socks or panty house, bend over to clean the bothtub, move a chair, pull or push heavy doors, carry two bags of gloceries, lift and carry a heavy suitcase 등으로 구성되며 감정 상태, 통증 정동, 사회적인 활동, 성생활 등의 항목은 제외되었다. 각각의 항목은 전혀 어렵지 않는 0에서 부터 동작을 할 수 없는 5까지의 점수로 표현되며 질문들은 금일의 상태에 대한 특정 질문들로 구성되어 있다. 총 점수는 각각의 항목의 점수를 합쳐서 100으로 나누어 구하며¹⁸⁾ 높은 점수일수록 장애가 심한 것이다.

10) Lumbar Spine Questionnaire

Lumbar Spine Questionnaire은 North American Spine Society와 American Academy of Orthopedic Surgeons에서 장려되는 scale로 ODQ의 수정된 버전을 포함하고 있다. 게다가 이 척도는 신경학적 척도, 환자의 현재 상태에 대한 만족도 질문, 기대 적

도를 포함하고 있다.

11) SF-36(Short Form - 36)

전체적인 척도에 쓰이는 것으로 상대적으로 간단한 검사법이어서 세계 여러 나라에서 활용되고 있다. 19). 8가지의 척도로 구성되어 있다. 8개의 척도는 Physical Functioning(PF), Mental Health(MH), General Health(GH), Bodily Pain(BP), Role Physical (RF), Role Emotional(RE), Social Functioning (SF), Vitality(VT)로 구성되어 있으며 각 척도는 2에서 10까지의 항목으로 구성되어 있으며 점수는 0-100%로 되어 있어 0은 심각한 장애를 의미하며 100은 장애가 없음을 의미한다(Table IV)¹⁹⁾. 또한, 두가지의 변환으로 나누어 평가하는 방법이 있는데 Physical Component summary(PCS)와 Mental Component Summary(MCS)로 나누어 평가하는 방법이다. 높은 점수가 건강한 상태를 나타내며, 50을 평균적인 건강한

사람이 나타내는 수치가 되도록 설정을 하여 점수가 50보다 크게 나르면 건강상태를 나타내는 것이고 50보다 적게 나오면 평균이하의 건강상태를 나타낸다. 결과치를 두개로 하는 것이 8개로 하는 것보다 간단하므로 최근에는 PCS와 MCS를 이용하는 추세에 있으나 아직은 논란이 되는 추세이다. 검사시간은 전체적인 검사이지만 5-10분 정도면 끝난다. SF-36는 요통에 특징적인 검사로 평가할 수 있다는 주장도 제기되고 있다.

검사 시간의 단축을 위해서 SF-12²⁰⁾가 등장하였다. Ware 등에 의해서 1996년 제기되어 검사시간을 2분 정도로 단축시키는 결과를 가져왔다. 이는 Simon²¹⁾ 등에 의해서 세세한 것에는 일치할 수 없다는 주장을 하였다. 2003년 Riddle²²⁾ 등에 의해서 PCS의 평가에 있어서는 SF-36과 SF-12가 똑같은 민감도를 가지나 개별의 환자의 평가에는 유용하지 못하여 PCS-12는 쓰지 않을 것을 권하고 있다.

Table IV. Summary of SF-36(1994)

Summary Measures	Scales	Items(Question No.)
Physical Health	Physical Function(PF)	3a-3j
	Role-Physical(RF)	4a-4d
	Bodily Pain(BP)	7-8
	General Health(GH)	1, 11a-11d
Mental Health	Vitality(VT)	9a, 9e, 9g,9i
	Social Functioning(SF)	6, 10
	Role-Emotional(RE)	5a-5c
	Mental Health(MH)	9b-9d, 9f, 9h

12) SIP(Sickness Impact Profile)

1976년에 고안되고 1981년 Bergner에 의해 발표되어진 평가법으로 SIP는 sicknee(불건강) 때문에 나타나는 행동의 변화를 평가하여 서로 다른 의학 영역 간에 공통적으로 적용 할 수 있는 평가법으로 일반적인 건강 상태의 측정을 위해서 만들어졌다. 이는 136가지의 항목으로 구성되어 있으며 12개의 범주로 나눌 수 있다. Ambulation, Body care/movement, mobility, Emotional behavior, Social behavior, Alertness behavior, Communication, Work, Sleep and rest, Eating, Home management, Receational activities 나눌 수 있으며 신체적(이학적), 정신적, 독립적인 나머지의 3개의 큰 범주로 나눈다(Table II, V). Yes or No의 형태로 답을 합하며 높은 점수일수록 기능장애가 심하거나 건강하지 못함을 의미한다. 설문은 작성하는데 20-30분이 걸리며 계산기로 계산하는데 5분정도 걸린다.

SF-36과 더불어 타당도, 신뢰도, 민감도의 검증을 받아 요통에 있어서 전체적인 척도로 가장 일반적으로 쓰이는 검사법²³⁾이었으나 현재는 추천받은

검사척도가 아니다.

13) NHP(Nottingham Health Profile)

NPH는 진료시 초진을 위하여 만들어진 간단한 설문지이다. 38개의 항목으로 구성되어 있으며 이는 6개의 항목으로 나눌 수 있다. Physical ability, pain, sleep, social isolation, emotional reaction, energy 이다. 일부 임상가들을 제외하고는 추천을 받지 못하는 평가법이다²⁴⁾. yes/no 대답의 형태를 취하며 SIP에 비하면 행동적인 면보다는 직접적으로 감정적인 면에 중점을 두고 있다.

14) DUKE(Duke Health Profile)

Duke-University of North Carolina health profile의 개정되고 짧아진 평가법으로 초진을 위하여 개발되었다. 17개의 질문으로 구성되어 있으며 건강과 기능장애의 2분류로 나누어진다. 건강은 6개의 항목으로 기능장애는 4개의 항목으로 나누어진다(Table VI.).

Table V. Summary of SIP

Name	Category	Item
Sickness Impact Profiles	Physical dimension	Ambulation, mobility, Body care
	Emotional dimension	social interaction, alertness, emotiona, communication
	Independent	Sleep and rest, Eating, Home management, Receational activities

Table VI. Summary of DUKE

Name	Category	Item
Duke Health Profiles	Health score	Physical health, mental health, social health, perceived health, self-esteem
	Dysfunction score	anxiety, depression, pain, disability

16) EuroQol

건강상태의 기술과 가치 판단을 포함하는 척도이다. 1990년 유럽인에 적용을 목적으로 하여 개발된 것으로써 5개의 항목으로 구성되어 가동성(mobility), 자기근심(self-care), 일상생활(usual activities), 고통과 불편(pain/discomfort), 두려움과 우울(anxiety/depression)으로 구성되어 이들의 세부항목은 243가지의 건강상태를 표시하고 있다. 2001년 Andrew 등에 의해서 민감도의 평가로 공인을 받았으나 요통의 적용에 있어서는 APBS와 RDQ에 비해서 민감도가 떨어지는 것으로 확인되었다²⁵⁾.

III. 고 찰

요통의 평가법은 증상과 장애를 포함한 기능적 평가법, 이학적 검사, 일반 방사선 촬영, 근력검사 등 여러 가지가 존재한다. 요통과 이로 인한 부가적인 증상은 기능장애를 흔히 일으키지만 객관적인 생리학적 검사에서 요통과 부가적인 증상에 일치하는 경우는 많지 않다. 이러한 이유로 기능의 제한에 중요시 하여 요통을 평가하는 척도가 등장하기 시작하였다. 그러나 활동의 평가는 임상적으로는 비효율적인 경우가 많아서 환자의 자기평가(self-report)를 포함하는 평가가 많이 등장하였으며, 이런 경우 하나의 기본이 되는 척도법이 나오면 많은 활동의 제한에 있어서 정보가 분석되고 모여질 수 있게 되어 1980년 이후 요통 평가의 기준이 되기 여러 척도들이 개발되고 연구되기 시작하였다. 2003년 현재 MEDLINE, CINAHL 등에 등장한 요통의 특정한 검사를 평가를 위한 척도만 24가지이며, 전체적인 척도를 합치면 더욱 많아진다.

요통의 기능에 의한 평가법은 요통에만 특징적인(condition-specific) scale과 신체 전체의(generic-health) scale로 나누어 설명 할 수 있다. 요통의

scale은 먼저 요통에 특징적인 scale이 먼저 등장하였으며 이후 건강상태를 나타내는 전체적인 검사법이 등장하였으며, 전체적인 scale이 요통의 검사법에 이용되기 시작하면서 현재 이들은 서로 같이 쓰이기도 하고 전체적인 검사법이 요통에 있어서도 특정한 평가법을 대응되기 위한 연구가 지속되고 있다. 이상적인 평가법이란 타당도(validity), 신뢰도(reliability), 간편도(simplicity), 민감도(responsiveness)를 갖춘 평가법을 말한다. 타당도라는 것은 측정하고자 하는 것을 얼마나 정확하게 측정하고 있는가를 말하는 것이고, 신뢰도라는 것은 반복 측정시 변화가 적어야 즉, 재현성이 있어야 하는 것을 말하는데 동일인(intra-observer)의 반복 측정시, 검사자간(inter-observer)의 반복 측정시 같은 결과가 나와야 함을 말한다. 간편도라는 것은 평가법이 시행하기 쉬워야 하고 짧은 시간 내에 이루어질 수 있어야 하는 것을, 민감도라는 것은 시간의 변화에 따른 변화를 잘 반영할 수 있어야 함을 말한다²⁶⁾.

지금까지 많은 설문지가 요통과 관계된 기능 장애를 평가하는데 사용되어 왔다. 요통을 기록할 수 있는 이상적이고 명확한 척도가 존재하지 않는 상황에서 많은 저자들이 여러 측정 체계를 개발하였으며 이 척도들은 주관적인 그리고 객관적인 척도를 서로 다른 정도로 포함하고 있다. 그러나 대부분의 scale은 임시적으로 연구마다 결과를 평가하기 위해 발달되었으며, 서로의 scale 사이에는 비교가 거의 없는 실정이다. 검사를 위한 척도는 많은 시간이 걸려서는 임상적으로 의미가 없다. 현재 쓰이는 대부분의 검사는 타당도, 신뢰도, 민감도를 공인을 받았지만 30분 이상의 시간이 걸리는 검사는 없음을 알아야 한다. 기존의 공인받은 scale로는 ODI, RDQ, Waddle Disability Index, Functional Nating Scale, Resumption of Activities of Daily Living Scale, Back Outcome Score, Dallas Pain Questionnaire, Aberdeen Low Back Pain Scale, Quebec

Back Pain Disability Scale, Lumbar Spine Questionnaire, SF-36, SIP, NPH, DUKE, EuroQol 등이 있다. 이러한 척도들의 대부분에 있어서 물론 타당도, 신뢰도, 민감도 등의 연구가 시도되었으며 자체적으로는 공인받았다.

특정적인 scale에 있어서 수많은 척도의 평가로 이용되고 있는 것이 ODI와 RDQ이다. 1998년 요통을 위한 국제 심포지엄에서 ODI와 RDQ는 임상활용을 추천 받은 검사법이기도 하다²⁷⁾. 1980년 일상생활의 각각의 동작과 관련된 평가를 위해 Fairbank에 의해 ODI가 1983년 Roland, Morris에 의해 SIP로부터 RDQ가 요통으로 인한 신체적 장애를 평가하기 위해 만들어졌다. 따라서 상대적으로 RDQ가 전체적인 기능 평가와 민감도에서 우수하다. 이후 두 scale은 여러 차례 점점 더 개선을 위해 여러 학자들에 의해 개정이 되었다. 설문과 측정기간의 차이에 의해 개정이 이루어졌지만 기본적인 것은 변하지 않았다. 1997년 Leclaire 등²⁸⁾은 ODI와 RDQ를 비교하였는데 두 평가방법이 기능적인 능력의 평가라고 여겨 공통적인 구성을 가진다고 하였으며 심각한 환자에게는 ODI가 더 민감하다고 주장하였는데, 이는 2000년 RDQ와 ODQ는 Roland와 Fairbank에 의해 비교된, 두 scale간에는 큰 차이가 없지만 다만 심한 장애의 환자에서는 RDQ가 최대값을 가질 때, ODQ는 변화를 보여줄 수 있으며, 심하지 않는 장애에서 ODQ가 최저값을 가질 때, RMDQ는 변화를 보여줄 수 있어 Roland와 상대적으로 심한

환자에서는 ODQ를 적용하며, 상대적으로 경한 환자에서는 RMDQ를 적용할 것을 추천한 것과 일치한다(Table VII). Beurskens 등²⁹⁾은 ODI와 RMDQ를 비교하여 ODI가 유의하다고 평가하였다.

이학적 검사를 요통의 기능적인 평가에 포함하고자 한다면 Million Index, Low Back Outcome Scale, ABPS를 사용할 수 있다. Low Back Outcome Scales은 기능적인 장애와 통증이 오지 않는 범위내에서 수동적 움직임을 평가하는 것이 특징으로서 수동적인 평가로 인해 병의 심각도와 치료의 범위를 정하는데 장점이 있으며, 처음에는 직장인의 평가를 위해 개발된 척도이며 Aberdeen Low Back Pain Scale는 임상과 기존의 검사척도를 기본으로 하여 만들어진 검사법이어서 이학적 검사가 포함되어 있으며, 상대적으로 Million Index는 최근 잘 쓰이지 않는 검사법이어서 최근에는 논문에 요통의 기능평가로 이용되지 않고 있다.

환자의 기대척도를 알면서 치료하고자 한다면 Lumbar Spine Questionnaire이 좋으며, 감정적인 면을 고려하여 요통의 환자를 평가하고자 한다면 Dallas Pain Questionnaire가 좋다. 만약 전체적인 검사척도를 사용한다면 정서적인 요소가 포함되어 있는 전체적인 검사척도를 모두 이용할 수 있다. 이는 환자가 심리적으로 우울하다든지, 감정의 기복이 심하다면 추천될 만 하다.

요통에 특징적인 검사법은 민감도, 신뢰도, 타당도의 비교 연구가 기본적으로 ODI, RMQ와의 비교

Table VII. Comparison of the Roland-Morris Disability Questionnaire and the Oswestry Disability Index

Instrument	No. of Item (response options)	Score (best of worse)	Time to complete, min	comparison	Dimensions
Roland-Morris	24(yes/no)	0-24	5	mild	physical activities, housework, mobility, dressing, getting, help, appetite, irritability, pain
Oswestry	10(6 levels)	0-100	5	severe	pain intensity, personal care, lifting, walking, sitting, standing, sleeping, sex life, social life, traveling

라는 사실을 알 수 있다. 각각의 민감도, 타당도, 신뢰도의 평가된 논문을 살펴보면 1992년 Greenough CG³⁰⁾ 등은 Low Back Outcome과 ODI를 비교하여 Low Back Outcome이 ODI 등과 동등한 타당도를 가지고 있다고 평가되었으며, 그 이후 다양하게 자주 쓰이고 있다.

기존의 ODI와 RDQ가 우수하다는 논문의 결과와 달리 1995년, 1996년 Kopec 등³¹⁾은 Quebec 척도가 ODQ, RDQ 보다 타당도가 우수하다고 주장하였으나 그 이후의 지지를 하는 후속적인 3가지를 비교하는 논문은 나타나지 않고 있다. 이러한 Kopec의 주장은 Quebec과 ODI를 비교시 Quebec은 ODI에 비하면 특정한 행동(수마일 걷기, 공을 던지기, 의자를 옮기기)에 대해서 질문을 함으로 ODI의 단순한 행동평가(걷기, 던지기, 옮기기)보다 점수가 낮게 나오는 경우가 많으며 Quebec은 당일 시행하는 것을

평가하며 ODI는 기억에 의존함으로 시간에 따른 변이가 있어 검사의 내용에 따라 충분히 달라질 수 있다.

2001년 Garratt³²⁾ 등은 RDQ, Aberdeen Low Back Pain Scale과 EurQol를 비교하여 민감도는 RDQ, Aberdeen Low Back Pain Scale가 더 좋다고 결론을 내렸으며 RDQ가 가장 우수하다고 하였다.

전체적인 평가로는 SF-36, SIP, NPH, DUKE, EuroQol 등이 있다. 전체적인 건강 측정은 건강(health), 장애(disability), 삶의 질(quality)을 평가하기 위해 고안되었다. 이들의 각각의 항목은 다르지만 전체적인 평가는 신체 전반의 건강 상태의 평가를 위해 만들어졌으나 요통의 평가방법으로도 쓰이고 있으며, 요통에 특정한 검사법과 비교 논문이 나오고 있다. 각 항목은 검사평가법마다 다양하다 (Table VIII).

Table VIII. Comparison of Dimension and Subscales in Generic Health Outcome

Dimensions/Categories	SF-36	SIP	NHP	Duke
Alterness behavior		X		
Ambulation		X		
Anxiety/depression		X		X
Body/self care		X		
Communication		X		
Change in health		X		X
Disability				
Eating				
Emotional behavior		X	X	
Energy/Vitality	X	X	X	
General/Perceived health	X			X
Home management				
Mental health	X	X		
Mobility				
Pain	X			
Physical ability/health	X		X	X
Recreation		X	X	X
Role function/Usual activity	X			
Self-esteem				
Sleep/rest		X	X	
Social interaction	X		X	
Work		X		
Combined Subscales				
Physical		X		
Mental	X	X		
General health	X			X

SIP는 SF-36의 등장 이후 잘 활용되지 않고 있다. 이들은 요통 뿐만 아니라 여러 질병의 평가에 사용되어지고 있다. SF-36는 크게는 Physical Component Summary(PCS)와 Mental Component Summary(MCS)로 나누어 평가하기도 하며, 8개의 척도로 Physical Functioning(PF), Mental Health(MH), General Health(GH), Bodily Pain(BP), Role Physical(RF), Role Emotional(RE), Social Functioning(SF), Vitality(VT)로 평가되기도 한다. 36가지의 설문지여서 5분정도면 끝낼수 있으나 시간의 단축을 위하여 SF-12가 등장하였으며 SF-36과 SF-12의 효용성에 대한 논문도 제기되고 있으며 효용성의 논란은 여전하다.

또 다른 논문의 방향은 요통에 특정한 scale과 검사법과 전체적인 검사법의 비교 논문이다. 요통의 검사로 전체적인 scale 쓰여 질 수 있는가의 문제와 더 나아가서 요통만의 평가로 단독으로 쓰여 질 수 있는가의 비교 논문이다. 지금까지 비교가 많이 되고 이들의 요통에 대한 신뢰도, 타당도, 민감도의 평가가 많이 이루어져 왔다. 1999년 Taylor 등³⁹⁾은 Low Back Outcome Score의 반응성과 ODI와 SF-36의 반응성을 비교하였는데, The Low Back Outcome Score는 SF-36보다는 반응성이 뛰어났지만, RDQ보다는 반응성이 떨어진다고 보고하였다.

ODI와 SF-36는 일반적으로 추천받은 검사법이다. 구조적으로 정확하게 평가받기 위한 분석의 연구가 일반적인 대조군과 요통 혹은 방산통을 가지고 있는 환자들에게 계속되어 왔는데 이는 ODI, SF-36과의 비교를 통해 일관성, 신뢰성, 타당도를 획득하기 위해서이다³⁴⁾.

1995년 Kopec 등³⁵⁾에 의해 Quebec scale이 RMDQ, ODQ, SF-36 Physical function scale과 신뢰성, 차별성, 반응성의 항목에서 비교되었는데, 네 가지 scale 모두 표준 평가법으로 의미를 갖출 수 있었으나 The Quebec scale, RMDQ, ODQ가

SF-36 Physical Function scale보다 반응성에서 뛰어났다는 평가를 받았으며, 이는 2001년 Garratt 등³⁶⁾은 RDQ, Aberdeen Low Back Pain Scale과 EurQol를 비교하여 민감도는 RDQ, Aberdeen Low Back Pain Scale가 더 좋다고 결론을 내렸는데 요통에 있어서는 전체적인 평가법보다는 특정적인 평가법이 우수하다는 주장과 일치한다. 이와는 반대적인 주장도 제기되어 있는데, 기존의 주장과는 달리 SF-36이 ODI보다 민감도에서 우수하다고 하였으며, 2003년 Walsh 등³⁷⁾에 의해서는 특정한 질병에 맞는 척도인 ODI와 전체적인 척도인 SF-36이 민감도에 있어서 특이한 차이를 발견할 수 없다는 결론에 도달하여 요통의 평가에 있어서 전체적인 척도의 단독 사용의 가능성을 제기하였다. 이는 전체적인 평가법 중에서 SF-36은 요통의 특정한 평가척도 보다 동등한 혹은 우수할 수 있다는 것을 의미한다. 즉, 건강상태의 평가의 항목은 두가지로 나누어지는데 일반적인 상황과 특정한 상황의 평가로 나누어진다. 일반적인 상황의 평가는 건강의 다양한 면을 살핀다. 특정한 상황이 평가는 특정한 증상 등의 평가에 이용된다. 요통에 특정한 검사는 일반적으로 요통의 평가에 특이성이 있는 것으로 알려져 있다. 또한, 전체적인 검사는 특히 동일한 상황의 환자의 비교나 다른 상황의 환자의 비교에 유용하다. 전체적인 scale은 다양한 상태를 평가하므로 건강의 현상에 좀 더 한정된 검사보다 유의한 결과를 갖게 되며 한정된 검사가 놓치기 쉬운 부정확한 반응을 평가할 수 있는 장점이 있다. 따라서 이는 치료효과가 불확실할 때 부작용(side effect)을 평가할 수 있게 된다. 특징적인 평가법이 민감성이 전체적인 보다 뛰어나며 특히 통증의 다양성의 평가에서는 더욱 그러하다고 옹호한다. 따라서 두가지 방법을 같이 평가하는 것이 가장 적당하다는 주장이 제기되고 있다. 그러나 두가지 모두를 이용하기 위해서는 더욱 많은 후발적인 연구가 필요하다. 만약, 전체적인

평가척도가 특정한 평가법과 동등한 혹은 그 이상의 민감성을 획득한다면 일반적인 평가의 장점을 되살릴 수 있게 될 것이다. 한의학은 인체의 전체성을 대상으로 하는 학문이다.

아직까지도 동일 그룹을 대상으로 하는 scale간의 직접적인 비교가 시작되는 단계이므로 평가 방법상의 차이를 평가할 수는 없다. 이러한 이유로 특정의 요통의 평가법이 다른 것보다 더 나은 평가법이라고 단정할 수 있는 상황이 아니다. 이에 Outcome Assessment의 General Recommendation에서는 Low back pain의 functional outcome scale로 가장 널리 사용되고 있는 RDQ와 ODI를 추천하고 있으며 The Quebec Back Pain Disability Scale이 그 다음의 평가법으로 추천되었다. 다른 scale의 경우도 타당성, 신뢰성, 반응성 등에서 표준 scale로는 의미가 있으나 현재 비교할 수 있는 자료가 제한되어있는 실정으로 연구자의 필요성과 선호도에 의해 사용을 고려할 수도 있다. 전체적인 평가법으로는 SF-36이 추천되고 있다. 2000년에 Bombardier C.³⁸⁾은 Spine에 Outcome Assessment의 General Recommendation을 발표하여 Deyo RA 등이 제안한 다섯가지의 범주에 각각의 scale을 추천하고 있는데, Function에서는 RDQ와 ODQ를, Generic health status에서는 SF-36 version 2.0을, Pain에서는 Bodily pain scale of SF-36과 GCPS를, Work disability에서는 work status와 일에 대한 손실 날짜수, 일에 대한 복구 시간을, Satisfaction에서는 치료결과와 치료과정에 대한 만족도의 설문을 제안하고 있다. 향후 전체적인 연구방법상의 scale 적용의 표준화와 규격화의 첫걸음이라고 사료되며 2003년 현재 아직 다시 제시된 것은 없는 상황이다.

환자에게 치료 결과를 평가하는 가장 보편적인 이유는 치료에 따른 환자의 반응을 평가하기 위해서다. 환자가 더 좋아졌는가? 무엇이 임상적으로 변한 증상과 가장 관계가 있는가? 어떤 항목이 증상

과의 관계에서 가장 유의성이 있는가? 여기에 대해서는 현재 정해진 답은 없다. Patrick DL 등³⁹⁾은 The Maine Lumbar Spine Study의 자료에 의해 임상적으로 중요한 최소한의 변화는 RDQ의 24개의 문항에서 2-3point의 변화이며 SF-36 Physical Function subscale에서는 대략 7point라고 보고하였으며, Bombardier C. 등⁴⁰⁾은 RDQ의 24개의 문항을 적용하였을 때, 2-8 point의 변화는 서지학적인 면에서 의미가 있으며 2-3point의 변화는 임상적으로 의미 있는 최소한의 변화라고 하였다. 점수(score)변화의 의미는 시간의 변화에 따른 통증의 강도나 신체적 장애의 변화를 알 수 있게 해 주는 것이다. 이는 척도의 신뢰도와 민감도를 검사시 변화의 기준을 삼기 위해서 등장하였다. 점수는 의미있는 변화가 나타나야 하는데 이는 두가지의 의미가 있는데 통계학적 또는 임상적인 것을 의미한다. 통계학적인 의미는 최소한의 감지되는 변화(minimum detectable change), 가장 적은 감지되는 차이(smallest detectable difference), 최소한의 믿을만한 변화(minimum reliable change) 등으로 여러 논문에서 표현되고 있다. 임상적인 변화는 최소한의 임상적인 중요한 변화(minimum clinically important change, MICD)라고 불린다⁴⁰⁾.

과거 10년동안 Item Response Theory의 사용에 큰 관심을 가져 왔다. 이것은 연구자에게 환자가 왔을 때 측정된 다른 장애의 수준이 다른 항목에 어떻게 반응할 것인가를 결정할 수 있게 한다. 적절한 통계적 방법만 사용한다면 연구자는 환자 상태에 따라 주어진 항목에서 어떤 평가를 받을 것인가의 가능성을 평가할 수 있다. 그리고 이러한 scale이 컴퓨터 작업화 된다면 Computerized adaptive testings가 적용될 수 있다. 현재처럼 종이와 연필로 하는 scale이라면 Computerized adaptive testings은 쉽게 적용될 수 없지만, 이런 Item Response Theory의 발전은 항목의 선택, scale의 발전, 점수의 평가에 유

용할 것이다. 현재 이런 Item Response Theory의 연구는 요통에서는 Quebec scale에 먼저 적용된 이래로 Item Response Theory 연구의 적용 예가 빠르게 늘어나고 있는 상황이다. 또한, 컴퓨터를 이용한 조사는 모은 자료를 간단화 할 수 있으며, 자료의 질을 향상시킬 수 있다. 환자로 하여금 논리에 어긋나거나 둘 이상의 답을 못 하게 할 수 있으며 같은 질문에 대하여 서로 다른 답을 못하게 할 수도 있다. 2002년 Hanscom 등⁴¹⁾에 의해서 SF-36, ODI의 척도는 컴퓨터로 작업하는 것이 기존의 검사법보다 우수함을 보여주었다. 또한 컴퓨터는 전체적인 평가척도의 장기적인 평가에도 도움을 줄 것으로 사료된다.

IV. 결 론

문헌을 통해 요통의 기능적 평가법을 연구한 결과 다음과 같은 결론을 얻었다.

1. 요통의 기능적 평가법은 요통에 특정한 평가법과 전체적인 건강 상태의 평가법이 있다.
2. 요통의 기능적 평가법의 기본 요소로 back specific function, generic health state, pain, work disability, satisfaction이 사용될 수 있다.
3. 요통만을 위한 기능적 평가법으로는 특정한 평가법인 RDQ와 ODI가 전체건강 상태의 평가법으로는 SF-36이 사용되고 있다.
4. 요통의 기능적 평가에 SF-36 단독으로 쓰일 수 있다.
5. 미래의 low back pain scale의 연구는 다면적인 입장에서 전체적인 평가를 포함한 표준화와

규격화이며, 이는 IRT를 사용한 컴퓨터를 이용한 방법으로 지향될 것이다.

참고문헌

1. 김남현, 이환모. 요통의 평가지수. 대한정형외과학회지. 1990;25(3):927-32
2. Waddell G, Main CJ. Assessment of severity in low back disorder. Spine. 1984;9:204-8
3. Bouter LM, van Tulder MW, Koes BW. Methodologic issues in low back pain research in primary care. Spine. 1998;23:2014-20
4. Mayer TG, Mooney V, Gatchel RJ, et al. Quantifying postoperative deficit of physical function following spinal surgery. Clin Orthop. 1989;244:147-57
5. Deyo RA, Battie M, Beurskens M, Bombardier C, Croft P, Koes B, Malmivaara A, Roland M, Von Koff M, Waddell G. Outcome measures for low back pain research. spine. 1998;18:2003-13
6. Fairbank J, Pynsent P. The Oswestry Disability Index. Spine. 2000;25:2940-52
7. Garratt AM, Moffett JK, Farrin AJ. Responsiveness of generic and specific measures of health outcome in low back pain. Spine. 2001;26:71-77
8. Patrick DL, Deyo RA, Atlas SJ, et al. Assessing health-related quality of life in patients with sciatica. Spine. 1995;17:1899-909
9. Stratford PW, Binley JM. Measurement properties of the RM-18: A modified

- version of the Roland-Morris Disability Scale. Spine. 1997;22:2416-71
10. Andrew M, Garratt A. Rasch analysis of the Roland disability questionnaire. Spine. 2003;28:79-84
 11. Waddle G, Main CJ. Assessment of severity in low back disorder. Spine. 1984;9:204-12
 12. Million R, Hall W, Haavick-Nilsen K, Baker RD, Jayson MIV. Assessment of the progress of the back pain patient. Spine. 1982;7:204-12.
 13. Ozguler A, Gueguen A, Leclerc A, et al. Using the Dallas pain questionnaire to classify individuals with low back pain in working population. Spine. 2002;27:1783-89
 14. Lawlis GF, Cuencas R, Selby D, et al. The development of the Dallas pain questionnaire: An assessment of impact of spinal pain on behavior. Spine. 1989;14:511-6
 15. Thomasen K, Christensen FB, Eiskjaer SP, et al. The effect of pedicle screw instrumentation of functional outcome and fusion rates in postolateral lumbar spinal fusion. Spine. 1997;22:2813-22
 16. Kopec JA. Measuring functional outcome in person with back pain: A Review of back-specific questionnaire. Spine. 2000;25:3100-3
 17. Ruta D, Garratt A, Wardlaw D, et al. Developing a valid and reliable measure of health outcome for patient with low back pain. Spine. 1994;19:1887-96
 18. Kopec JA, Esdaile JM, Abrahamowicz M, Abenhaim L, Wood-Dauphinee S, Lamping D, et al. The Quebec back pain disability scale: Measurement properties. Spine. 1995; 20:341-50
 19. National Health Survey: SF-36 population Normas. Canberra Australian Capital Territory, Australian Bureau of Statistics. 1997
 20. Ware JE, Kosinski M, Keller SD. A 12-item short form health survey. Med care. 1996; 34:220-33
 21. Simon GE, Revicki D, Grothaus L, et al. SF-36 summary scores Are physical and mental health truly distinct?. Med care 1998;36:567-72
 22. Taylor SJ, Tayler AE, Foy MA, et al. Responsiveness of common outcome measure for patients with low back pain. Spine. 1999;24:1805-12
 23. Deyo R. Comparative validity of the sickness impact profile and shorter scales for functional assessment in low back pain. Spine. 1986;11:951-4
 24. Beaton DE, Hogg-Johnson SA. Measuring health in injured workers: a cross-sectional comparison of five generic health status instrument in workers with musculoskeletal injuries. Am J Ind Med. 1996;29:618-31
 25. EuroQol Group. Euroqol: A new facility for measurement of health related quality of life. Health police. 1990;16:199-208
 26. Adrain white. Measuring pain. Acupuncture in medicine. 1998;16(2)
 27. Deyo RA, Battie M, Beurskens AJ, et al. Outcome measures for low back pain research: a proposal for standardized use. Spine. 1998;23:2003-13
 28. Leclaire R, Blier F, Fortin M, et al. A cross-sectional study comparing Oswestry

- and Roland-Moris functional disability scales in two populations of patients with low back pain of differing levels of severity. *Spine*. 1997;22:68-71
29. Beurskens AJ, de Vet HC, Koke AJ, Lindeman E, van der Heijden GJ, Regtop W, Knipschild PG. A patient-specific approach for measuring functional status in low back pain. *J Manipulative Physiol Ther*. 1999;22(3):144-8.
30. Greenough CG, Fraser SD. Assessment of outcome in patients with low back pain. *Spine*. 1992;17:36-41
31. Kopec JA, Esdaile JM, Abrahamowicz M, et al. The Quebec back pain disability scale: conceptualization and development. *J Clin Epidemiol*. 1996;49:151-61
32. Garratt A, Moffett J, Farrin A. Responsiveness of generic and specific measures of health outcome in low back. *Spine*. 2001;26:71-7
33. Taylor SJ, Taylor AE, Foy MA, et al. Responsiveness of common outcome measure for patients with low back pain. *Spine*. 1999;24:1805-12
34. Mchorney CA, Ware JE, Raczek AE. The MOS 36-Item Short-Form Health Survey (SF-36): II. Psychometric and clinical tests of validity in measuring physical and mental health constructs. *Med Care*. 1993;31:247-63
35. Kopec JA, Esdaile JM, Abrahamowicz M, Abenhaim L, Wood-Dauphinee S, Lamping D, et al. The Quebec back pain disability scale: Measurement properties. *Spine*. 1995;20:341-50
36. Garratt A, Moffett J, Farrin A. Responsiveness of generic and specific measures of health outcome in low back. *Spine*. 2001;26:71-7
37. Walsh TL, Hanscom B, Lurie JD, Weinstein JN. Is a condition-specific instrument for patients with low back pain/leg symptom really necessary? The responsiveness of the Oswestry disability index, MODEMS, and SF-36. *Spine*. 2003;15(6):607-15
38. Bombardier C, Hayden J, Beaton DE. Minimal clinically important difference. Low back pain: Outcome measure. *J Rheumatol*. 2001;28(2):431-8
39. Patrick DL, Deyo RA, Atlas SJ, et al. Assessing health-related quality of life in patients with sciatica. *Spine*. 1995;17:1899-909
40. Bombardier C. Outcome assessment in the evaluation of treatment of spinal disorder. *Spine*. 2000;25:3100-3
41. Hanscom B, Lurie JD, Homa K, Weinstein JN. Computerized questionnaires and quality of survey data. *Spine*. 2002;27:1797-1801