

라쉬분석을 이용한 세 가지 요통 장애 설문지의 비교

김경모

연세대학교 대학원 재활학과

박소연

전주대학교 대체의학대학 물리치료학과

이충휘

연세대학교 보건과학대학 물리치료학과, 보건환경대학원 인간공학치료학과

Abstract

A Comparison of Three Low Back Disability Questionnaires With Rasch Analysis

Gyoung-mo Kim, M.Sc., P.T.

Dept. of Rehabilitation Therapy, The Graduate School, Yonsei University

So-yeon Park, Ph.D., P.T.

Dept. of Physical Therapy, College of Alternative Medicine, Jeonju University

Chung-hwi Yi, Ph.D., P.T.

Dept. of Physical Therapy, College of Health Science, Yonsei University

Dept. of Ergonomic Therapy, The Graduate School of Health and Environment, Yonsei University

The purpose of this study was to review existing assessment tools for patients with low back pain and improve them through combination. A total of 314 patients with low back pain participated. Their condition was assessed using the Oswestry Disability Questionnaire (ODQ), the Quebec Back Pain Disability Scale (QBPD), and the Back Pain Functional Scale (BPFS). Rasch analysis was applied to identify inappropriate items, item difficulties, and the separation index. In this study, the 'sex life' item of the ODQ (10 items) and the 'sleeping' item of the BPFS (12 items) showed misfit statistics, whereas all items of the QBPD (20 items) were appropriate. After combining the ODQ, QBPD and BPFS, Rasch analysis was applied. The 'pain intensity', and the 'sex life' item of the ODQ and the 'throw a ball' item of QBPD showed misfit statistics. These 3 items were retained for further analysis. The remaining 42 combined ODQ-QBPD-BPFS items were arranged according to difficulty. For all subjects, the most difficult item was 'pain intensity', whereas the easiest was 'take food out of the refrigerator'. As the separation index of 42 combined ODQ-QBPD-BPFS was higher than that of the three questionnaires separately, difficulty of items varied with some need for rearrangement. The results of this study confirmed the possibility and need for a new back pain disability assessment tool, and produced one. Further study is needed to refine the questionnaire in consideration of psychosocial and occupational factors.

Key Words: Disability assessment; Low back pain; Rasch analysis.

I. 서론

최근 요통 환자들의 기능 측정을 위해 관절가동범위나 근력 등의 평가보다 자기기립식 신체기능 평가도구

등과 같은 주관적인 평가도구를 사용하는 연구가 증가하고 있다(Holm 등, 2003). 요통의 평가 및 검사에 대해 검토한 연구(Dagenais 등, 2010)에 의하면 요통을 평가하는데 있어서 척추의 병리적 손상, 신경학적 손상

여부와 더불어 환자의 병력, 신체 및 신경학적 검사를 통한 증상의 심각성, 기능적 제한정도를 평가하는 것이 중요하다고 강조하였다.

일반적으로 요통의 진단과 평가에는 X-ray와 MRI 등의 방사선 진단 장비를 사용하고 있지만 구조적 이상이나 신경학적 증상을 동반하지 않는 비특이성 요통의 경우에는 원인규명이 어려운 것으로 알려져 있다(Boos 등, 1995). 이는 요통 등 다양한 근골격계 질환의 경우에는 해부학적 문제와 다른 원인으로 통증이나 기능 이상이 발생하는 경우가 많기 때문이며, 방사선 노출의 위험 외에도 실제 요통 발생 원인과 관련 없는 해부학적 진단으로 인해 환자의 증상과 맞지 않는 치료가 제공되어 요통이 악화될 위험이 있다(Chou 등, 2009; Modic 등, 2005). 따라서 물리치료 분야에서는 평가 시 요통을 발생시키는 구조적 이상을 발견하는 것 이외에도 환자의 기능 상태를 정확히 평가하여 적합한 재활운동치료를 제공하고, 통증 및 기능 악화를 예방하고, 일상생활 복귀를 위한 훈련을 실시하여 정상적인 기능을 회복하는 것이 최상의 목표라 볼 수 있다(Lewis 등, 2005).

환자의 기능 정도는 실제 수행하는 양상을 관찰하여 평가하는 것이 정확할 수 있지만 임상현장에서는 쉽지 않은 일이므로, 활동 제한과 관련된 항목으로 구성된 표준화된 검사 도구를 환자에게 작성하게 하는 방법으로 평가하고 있다(Davidson과 Keating, 2002). 표준화된 자기기입식 설문지는 신체적 활동 제한에 대한 정보를 수집하고 통합하는데 편리할 뿐만 아니라(Beattie와 Maher, 1997; Delitto, 1994), 치료효과에 대한 근거를 제시하고 신체활동의 변화를 검증하는데 필요한 정보를 제공하는데 효과적이다(Haywood, 2006; Kopec 등, 1995). 또한, 설문지를 이용한 평가도구는 같은 내용의 평가항목을 환자 및 시간의 제약을 받지 않고 일관되게 평가 할 수 있어 신뢰성도 높다고 하였다(Waddell, 1998).

Grotle 등(2005)은 요통이 환자의 생활에 미치는 영향을 알아보기 위해 타당도와 신뢰도, 치료효과에 대한 반응성(responsiveness) 등 심리측정학적 특성을 만족하며 기능을 평가할 수 있는 측정도구가 필요하다고 하였으며, 최근 평가도구들이 이러한 특성을 만족하는지 여부를 비교한 연구가 다수 출판되고 있다. Davidson과 Keating(2002)이 요통으로 인한 장애를 평가하는 자기기입식 평가도구에 대해 MEDLINE과 CINAHL 데이터

베이스를 통해 검토해 본 결과 2000년 현재 오스웨스트리 장애 설문지(Oswestry Disability Questionnaire; ODQ)(Fairbank, 1980), 퀘벡 요통 장애 척도(Quebec Back Pain Disability Scale; QBPD)(Kopec 등, 1995), 롤랜드-모리스 장애 설문지(Roland-Morris Disability Questionnaire; RMDQ)(Roland와 Morris, 1983), 와들 장애 지수(Waddell Disability Index; WDI)(Waddell과 Main, 1984), 요통 기능 척도(Back Pain Function Scale; BPF)(Stratford 등, 2000) 등 24개의 평가도구를 사용한 논문이 출판되었다고 기술하였다. 이 중 내용 타당도가 입증되었다고 판단한 5개의 평가도구에 대해 신뢰도와 반응성에 대해 연구한 결과, 오스웨스트리 장애 설문지, 퀘벡 요통 장애 척도, SF-36 신체기능 척도(SF-36 Physical Functioning Scale)는 요통환자들에게 임상적으로 사용하기에 측정의 신뢰도가 충분하다고 입증되었지만, 와들 장애 지수와 롤랜드-모리스 장애설문지는 요통환자를 평가하기에 적절한 신뢰도를 보이지 않아 추천하기는 어렵다고 제시하였다(Davidson과 Keating, 2002). Rocchi 등(2005)이 9개의 평가도구에 대한 심리측정학적 특성에 대해 비교한 결과에 의하면, 오스웨스트리 장애 설문지, 퀘벡 요통 장애 척도, 롤랜드-모리스 장애설문지만이 이들 특성을 만족하는 평가도구라고 제시하였다. Bombardier(2000)는 요통평가에 있어서 추천할 만한 평가도구는 오스웨스트리 장애 설문지와 롤랜드-모리스 장애설문지라 제시하였다.

지금까지 요통환자의 신체적 장애를 평가하기 위해 다수의 평가도구들이 개발되어 사용되고 있지만, 국내에서는 오스웨스트리 장애 설문지의 적용에 관한 연구(김경모 등, 2011; 박소연 등, 2008) 이외에는 심리측정학적 특성에 대해 실시한 연구를 찾아보기 어렵다. 또한 요통환자의 신체적 장애를 평가하는 유사한 내용의 설문지들이 증가하기 때문에 임상에서 하나의 설문지를 지속적으로 사용하기에 어려움이 있다(Beattie와 Maher, 1997; Bombardier, 2000).

따라서 기존의 연구에서 신뢰도와 타당도가 높다고 검증된 세 가지 평가척도, 즉 오스웨스트리 장애 설문지와 퀘벡 요통 장애 척도 그리고 비교적 최근에 개발되었으며 항목수가 적어 임상에서 적용하기 쉬운 장점이 있는 요통 기능 척도(Rocchi 등, 2005; Stratford 등, 2000)를 통합한 후 라쉬분석을 통하여 항목의 타당도, 난이도 등을 검토하여 새로운 요통 평가도구를 개발하는 것이 가능한지를 알아보았다.

II. 연구방법

1. 연구대상자

본 연구는 서울과 경기 지역의 13개 병원에서 현재 요통으로 인해 치료를 받고 있는 18세 이상 성인 요통 환자 450명을 대상으로 본 연구의 목적과 평가도구에 대하여 설명한 후 동의를 구한 자에 한해 평가를 진행하였으며, 그 중 설문에 충실하게 응답한 314명(응답율 69.7%)을 대상으로 실시하였다. 연구대상자 중 남자는 140명(44.6%), 여자는 174명(55.4%)이었고, 평균 연령은 40.6세(표준편차 14.3, 범위 18~86)였다(표 1).

2. 평가도구

본 연구에서 사용한 오스웨스트리 장애 설문지의 경우 전창훈 등(2005)의 연구에서 사용되었던 한국어판 오스웨스트리 장애 설문지를 사용하였다. 퀘벡 요통 장애 척도와 요통 기능 척도는 영문판 평가도구를 본 연구의 목적에 대해 인지하지 않은 물리치료사 3인에게 각각 1차 번역을 의뢰하고, 저자가 비교 확인 후 물리치료학과 교수 2인에게 재의뢰하여 수정 확인하는 과정을 거쳐 한국어판으로 번역하여 사용하였다.

모든 평가는 환자에게 사전 동의를 구한 후, 각 평가도구에 대해 환자 본인이 느끼는 불편함이나 장애 정도를 직접 표시하도록 하였다. 평가는 서울 및 경기 지역에 근무하는 물리치료사 13인이 실시하였으며, 평가자에게 설문지 배포시 연구자가 미리 작성한 평가지침을 교육하여 동일한 지침에 따라 평가를 실시하도록 하였다.

가. 오스웨스트리 장애 설문지(Oswestry Disability Questionnaire; ODQ)

오스웨스트리 장애 설문지는 Fairbank 등(1980)이 요통환자의 기능적 장애를 측정하기 위해 자기기입 형태로 개발한 평가도구로 검사-재검사 신뢰도가 급내상관계수 .94로 높은 신뢰도를 보이고(Rocchi 등, 2005) 국제적으로 가장 많이 사용되는 평가도구이다(Davidson, 2008; Grotle 등, 2005) 통증 정도, 개인위생, 물건 들기, 걷기, 앉기, 서있기, 잠자기, 성생활, 사회생활, 여행의 10개 영역에서 0점에서 5점까지의 서

열척도로 구성되어 있으며, 환자 스스로 느끼는 불편이나 장애 정도를 표시한다. 총점은 50점으로 항목별 응답한 점수를 백분율(%)로 표시하고, 점수가 높을수록 장애 정도가 크다는 것을 의미한다(Fairbank 등, 1980; Grotle 등, 2005).

나. 퀘벡 요통 장애 척도(Quebec Back Pain Disability Scale; QBPD)

퀘벡 요통 장애 척도는 Kopec 등(1995)이 개발한 요통 장애 평가도구로 잠자리에서 일어나기, 밤 시간 동안 수면, 침대에서 몸 돌아눕기, 차에 타기, 20~30분간 서있기 등 총 20개 영역으로 구성되어 있고, 0점에서 5점까지의 서열척도로 구성되어 있다. 검사-재검사 신뢰도는 급내상관계수 .92로 높은 신뢰도를 보이며(Rocchi 등, 2005), 오스웨스트리 장애 설문지와 마찬가지로 점수가 높을수록 장애 정도가 큰 것을 의미한다(Kopec 등, 1995).

다. 요통 기능 척도(Back Pain Function Scale; BPFs)

요통 기능 척도는 Stratford 등(2000)이 개발한 요통 장애 평가도구로 평상시 업무나 집안일(학교생활), 취미-스포츠 활동, 집안에서의 힘든 활동, 구부리거나 웅크리기, 신발 또는 양말신기, 박스 들어올리기, 잠자기, 1시간 동안 서있기, 1.6 km 걷기, 2층까지 계단 오르기, 1시간 동안 앉아 있기, 1시간 동안 운전하기의 12개 영역으로 구성되어 있고, 0점에서 5점까지의 서열 척도로 구성되어 있으며, 검사-재검사 신뢰도는 급내상관계수는 .88로 높은 신뢰도를 보이는 평가도구이다(Rocchi 등, 2005). 총점은 60점으로 점수가 낮을수록 장애 정도가 더 크다는 것을 의미한다.

3. 분석방법

연구대상자의 일반적 특성은 윈도우용 SPSS ver. 18.0 프로그램을 이용하여 분석하였고, 세 가지 요통 평가도구는 Winsteps ver. 3.71.0(Winsteps, Chicago, IL, U.S.A.)을 사용하여 각각 라쉬분석을 실시하여 단일구성개념, 평정척도, 항목 난이도, 분리 신뢰도를 알아보았으며, 세 가지 평가도구의 항목을 합친 뒤

(N=314)

표 1. 연구대상자의 일반적 특성

일반적 특성	평균±표준편차
나이(세)	40.6±14.3(범위 18~86)
남/여(명)	140(44.6%)/174(55.4%)

다시 재분석 하였다. 평가도구의 각 항목의 내적합지수(infit)의 평균자승잔차(mean square residual; MNSQ)값이 .6보다 작거나 1.4보다 큰 동시에 Z-값이 -2보다 작거나 2보다 큰 경우 단일구성개념을 만족하지 않는 부적합한 항목으로 판정하였고(이충휘, 2007), 단일구성개념을 만족하는 항목은 항목 난이도를 알아보기 위해 서열척도를 등간척도로 변환하여 로짓값으로 나타내어 난이도 순서를 알아보았다(김경모 등, 2011; 박소연 등, 2008).

평정척도(rating scale)는 낮은 점수에서 높은 점수로 척도가 배열되고, 외적합지수(outfit)의 평균자승잔차가 2이하를 보이는 경우 척도가 적합한 것으로 판정하였다. 또한 각 평가도구의 대상자 분리지수(person separation index)와 항목 분리지수(item separation index)를 알아보았다. 라쉬분석에서는 낮은 점수일수록 장애 정도가 심한 것으로 가정되지만, 오스웨스트리 장애 설문지와 퀘백 요통 장애 척도는 낮은 점수(0점)인 경우보다 높은 점수(5점)일수록 장애 정도가 심한 것으로 표시되기 때문에 라쉬분석을 적용하기 위해 낮은 점수(0점)일수록 장애 정도가 심하고, 높은 점수(5점)일수록 장애 정도가 낮은 것으로 척도를 변환하였다(김경모 등, 2011; 박소연 등, 2008).

III. 결과

1. 단일 구성 개념 검정

가. 오스웨스트리 장애 설문지의 단일 구성 개념 검정
오스웨스트리 장애 설문지의 10개 항목에 대해 라쉬분석을 사용하여 적합도 검정을 한 결과, 내적합지수의 평균자승잔차의 평균값은 1.06이고, 평균 Z-값은 .5이었다. 부적합 판정을 받은 항목은 '생활' 항목으로 평균자승잔차는 1.47, Z-값은 3.9이었다.

나. 퀘백 요통 장애 척도의 단일 구성 개념 검정
퀘백 요통 장애 척도의 20개 항목에 대해 라쉬분석을 사용하여 적합도 검정을 한 결과, 내적합지수의 평균자승잔차의 평균값은 1.02이고, 평균 Z-값은 .1이었으며, 모든 항목에서 단일 구성 개념을 만족하였다.

다. 요통 기능 척도의 단일 구성 개념 검정
요통 기능 척도의 12개 항목에 대해 라쉬분석을 사용하여 적합도 검정을 한 결과, 내적합지수의 평균자승잔차의 평균값은 1.10이고, 평균 Z-값은 .0이었다. 부적합 판정을 받은 항목으로는 '잠자기' 항목이었으며, 평균자승잔차는 1.49, Z-값은 4.3이었다.

표 2. 부적합 판정을 받은 통합된 ODQ-QBPD-BPFS 42항목의 적합도 판정 결과

항목	항목	로짓값	오차	내적합지수	
				평균자승잔차	Z-값
ODQ ^a 1	통증의 정도	1.36	.07	1.60*	6.3
ODQ ^a 8	성 생활	.01	.09	1.41*	3.4
QBPD ^b 11	공 던지기	-.52	.08	1.40*	3.9

^aOswestry Disability Questionnaire: 오스웨스트리 장애 설문지,

^bQuebec Back Pain Disability Scale: 퀘백 요통 장애 척도,

*부적합 항목: 내적합지수의 평균자승잔차는 .6보다 작거나 1.4보다 큰 동시에 Z-값이 -2보다 작거나 2보다 큰 경우.

표 3. 통합된 ODQ-QBPD-BPFS 42항목의 6점 평정 척도 분석

척도	관찰수(%)	관찰된 평균값	내적합지수 평균자승잔차	외적합지수 평균자승잔차	구조적 추정값
0	474(4)	-1.89	1.22	1.30	없음
1	700(5)	-.91	1.13	1.20	-1.83
2	1248(10)	-.17	.96	.98	-1.14
3	2215(17)	.80	.97	1.10	-.26
4	3776(29)	1.85	1.03	.85	.81
5	4510(35)	3.29	.98	.98	2.41

표 4. 통합된 ODQ-QBPD-BPFS 42항목의 난이도와 적합도 지수

번호	문항	로짓값	오차	내적합지수		
				평균자승잔차	Z-값	
QBPD ^a	13	냉장고에서 음식 꺼내기	-1.29	.09	1.05	.5
ODQ ^b	4	걷기	-.98	.09	1.06	.6
QBPD	15	양말(스타킹) 신기	-.91	.09	1.06	.6
BPFS ^c	7	잠자기	-.84	.09	1.24	2.3
QBPD	14	잠자리 정리하기	-.82	.08	.89	-1.1
QBPD	10	높은 선반에 손 뻗기	-.74	.08	1.10	1.0
QBPD	3	침대(바닥)에서 몸 돌아눕기	-.72	.08	1.12	1.3
BPFS	5	신발 또는 양말(스타킹) 신기	-.63	.08	1.06	.7
QBPD	8	300~400 m 걷기	-.61	.08	1.05	.5
QBPD	2	밤 시간 동안 수면	-.57	.08	1.26	2.6
ODQ	7	잠자기	-.56	.08	1.12	1.3
QBPD	4	차에 타기	-.55	.08	.67	-3.9
QBPD	11	공 던지기*	-.52	.08	1.40	3.9
QBPD	7	계단 한 층 오르기	-.50	.08	.96	-.4
QBPD	1	잠자리에서 일어나기	-.45	.08	.86	-1.6
QBPD	17	의자로 이동하기	-.36	.08	.94	-.7
ODQ	2	개인 위생(씻기, 옷입기 등)	-.33	.08	.87	-1.5
QBPD	5	20~30분간 서있기	-.13	.08	.93	-.8
BPFS	10	2층까지 계단(20개) 오르내리기	-.10	.08	.97	-.3
ODQ	9	사회생활(친목활동, 취미, 여가활동 등)	-.06	.08	1.00	.1
QBPD	18	무거운 문을 밀고 당기기	-.02	.08	.99	-.1
ODQ	8	성 생활*	.01	.09	1.41	3.4
BPFS	1	평상시 업무나 집안일, 학교생활	.05	.07	.79	-2.6
ODQ	10	여행, 이동(나들이)	.26	.07	.94	-.7
BPFS	4	구부리거나 웅크리기	.26	.07	.88	-1.5
BPFS	11	1시간 동안 앉아 있기	.38	.07	1.11	1.3
QBPD	16	구부려서 욕조 닦기	.38	.07	.77	-2.9
QBPD	19	두 개의 장바구니 들기	.40	.07	.81	-2.4
ODQ	6	서있기	.44	.07	1.00	.1
ODQ	5	앉기	.47	.07	1.32	3.5
BPFS	12	1시간 동안 운전하기	.51	.07	1.12	1.3
ODQ	3	들기(물건들기)	.58	.07	1.09	1.1
QBPD	9	몇 km 걷기	.59	.07	1.05	.7
BPFS	8	1시간 동안 서 있기	.62	.07	1.01	.2
BPFS	3	집안에서의 힘든 활동	.63	.07	.73	-3.5
QBPD	6	몇 시간 동안 의자에 앉기	.68	.07	1.16	1.8
BPFS	9	1.6 km 걷기	.72	.07	1.00	.0
QBPD	12	100 m 뛰기	.73	.07	1.25	2.8
BPFS	2	평상시 취미, 레크리에이션, 스포츠 활동	.87	.07	1.08	1.0
BPFS	6	바닥에서 박스 들어올리기	.95	.07	.66	-4.6
QBPD	20	무거운 여행가방 들고 옮기기	.99	.07	.82	-2.3
ODQ	1	통증의 정도*	1.36	.07	1.60	6.3
평균			.00	.08	1.03	.2

^aQuebec Back Pain Disability Scale: 퀘벡 요통 장애 척도,

^bOswestry Disability Questionnaire: 오스웨스트리 장애 설문지,

^cBack Pain Function Scale: 요통 기능 척도,

*부적합 항목: 내적합지수의 평균자승잔차는 .6보다 작거나 1.4보다 큰 동시에 Z-값이 -2보다 작거나 2보다 큰 경우.

표 5. 요통 장애 설문지의 분리신뢰도

요통 설문지 종류		평균의 표준오차	분리지수 (separation index)	분리신뢰도 (separation reliability)
ODQ ^a	대상자	.58	3.06	.90
	항목	.09	8.37	.99
QBPD ^b	대상자	.57	3.60	.93
	항목	.09	7.89	.98
BPFS ^c	대상자	.65	3.61	.93
	항목	.09	7.50	.98
통합된 ODQ-QBPD-BPFS	대상자	.31	5.63	.97
	항목	.08	7.84	.98

^aOswestry Disability Questionnaire: 오스웨스트리 장애 설문지,

^bQuebec Back Pain Disability Scale: 퀘벡 요통 장애 척도,

^cBack Pain Function Scale: 요통 기능 척도.

라. 통합된 ODQ-QBPD-BPFS 42항목의 단일 구성 개념 검정

오스웨스트리 장애 설문지 10항목, 퀘벡 요통 장애 척도 20항목, 요통 기능 척도 12항목을 통합한 ODQ-QBPD-BPFS 42항목에 대해 라쉬분석을 실시하여 적합도 검정을 한 결과, 오스웨스트리 장애 설문지의 '통증의 정도'와 '성생활' 항목, 그리고 퀘벡 요통 장애 척도의 '공 던지기' 항목이 부적합한 항목으로 판정되었고, 나머지 39항목은 적합한 것으로 판정되었다(표 2). 42개의 항목 중 부적합한 항목으로 판정받은 3항목은 기준은 벗어났지만, 제외하지 않고 포함시켜 이후의 결과에 제시하였다.

2. 평정 척도 분석

라쉬분석을 이용하여 오스웨스트리 장애 설문지와 퀘벡 요통 장애 척도, 그리고 요통 기능 척도의 6점 평정 척도를 분석 한 결과 각각 관찰된 평균값이 낮은 점수에서 높은 점수 순으로 배열되었고, 외적합지수의 평균자승간차도 모두 2이하로 나타났으며, 구조적 추정값도 순서적으로 배열되었다. 또한, 통합된 ODQ-QBPD-BPFS 42항목의 6점 평정척도를 분석 한 결과 관찰된 평균값이 낮은 점수에서 높은 점수 순으로 배열되었고, 외적합지수의 평균자승간차도 모두 2이하로 나타났으며, 구조적 추정값도 순서적으로 배열되었다(표 3).

3. 항목 난이도

통합된 ODQ-QBPD-BPFS 42항목을 난이도 순서대로 배열하였다. 42개의 항목 중 오스웨스트리 장애 설

문지의 '통증의 정도'와 '성생활' 항목, 그리고 퀘벡 요통 장애 척도의 '공 던지기' 항목 등 3항목은 단일 구성 개념 검정에서 부적합 항목으로 판정받았으나, 제외하지 않고 항목 난이도에 포함시켰다. 가장 어려운 항목은 '통증의 정도' 항목이었고, 가장 쉬운 항목은 '냉장고에서 음식 꺼내기' 항목이었다(표 4).

4. 분리 신뢰도

오스웨스트리 장애 설문지의 대상자 분리지수(separation index)는 3.06이고, 분리신뢰도(separation reliability)는 .90이었고, 항목의 분리지수는 8.37이었고, 분리신뢰도는 .99이었다. 퀘벡 요통 장애 척도의 대상자 분리지수는 3.60, 분리신뢰도는 .93이었고, 항목 분리지수는 7.89, 분리신뢰도는 .98이었으며, 요통 기능 척도는 대상자 분리지수가 3.61, 분리신뢰도는 .93이었고, 항목 분리지수는 7.50, 분리신뢰도는 .98이었다. 통합된 ODQ-QBPD-BPFS 42항목의 대상자 분리지수는 3.61, 분리신뢰도는 .93이었고, 항목 분리지수는 7.50, 분리신뢰도는 .98이었다(표 5).

IV. 고찰

요통으로 인한 기능적 장애를 평가하는 것은 치료의 결과를 확인하는 방법뿐만 아니라 환자의 신체적 제한을 양적으로 평가하는 방법으로서 매우 중요한 과정이다. 요통 평가에 적합한 평가도구의 최종 선택은 임상적인 관련성이 우선적으로 고려되어야 하고 국제적으로

널리 사용되는 측면도 고려하여야 한다(Rocchi 등, 2005). Bombardier(2000)와 Rocchi 등(2005)은 여러 가지 요통 평가도구 중에서 타당성을 고려한 결과, 요통으로 인한 장애를 평가하기 위해 롤런드-모리스 장애 설문지와 오스웨스트리 장애 설문지 중 하나를 선택하는 것이 적합하다고 하였다. 퀘벡 요통 장애 척도 또한 충분히 검증되고 타당성 있는 평가도구지만 과학적 근거에 대해 제한적인 부분을 보인다고 하였고, 요통환자의 장애를 평가하기 위해서 오스웨스트리 장애 설문지와 롤런드-모리스 장애 설문지 두 종류 중 한 가지의 사용을 권장하지만, 천장 효과(ceiling effect)와 바닥 효과(floor effect)를 방지하기 위해 두 평가도구를 통합하여 사용하는 것을 제안하였다.

본 연구는 요통 환자의 평가도구로 널리 쓰이는 오스웨스트리 장애 설문지와 퀘벡 요통 장애 척도, 그리고 요통 기능 척도 등 세 가지의 평가도구에 대해 라쉬분석을 통해 단일 구성 개념 검정과 평정 척도, 항목 난이도, 그리고 분리신뢰도에 대해서 알아보았다. 그리고 다양한 연구 배경을 통해 개발된 세 가지의 평가도구를 통합한 후 재분석 하여 새로운 평가도구 제작의 가능성을 알아보기 위하여 실시하였다. 오스웨스트리 장애 설문지는 평가시간이 5분 이내로 사용하기 쉽고 신뢰도와 타당도가 높아 국내외에서 가장 많이 사용되고 추천되고 있는 평가도구이며(Bombardier, 2000; Rocchi 등, 2005), 퀘벡 요통 장애 척도 역시 신뢰도와 타당도가 높고 시간의 흐름에 따라 장애 정도 변화에 대해 환자 개별적으로 관찰이 가능하도록 개발되어 여러 나라에서 번역되어 사용되고 있다(Kopec 등, 1995; Rocchi 등, 2005). 요통 기능 척도는 평가시간이 짧고 장애의 정도를 평가하는데 사용하기 쉽도록 구성되어 있는 장점이 있다(Rocchi 등, 2005; Stratford 등, 2000).

단일 구성 개념 검정에서 부적합 항목으로는 오스웨스트리 장애 설문지 10개 항목 중에서 '성생활' 항목과 요통 기능 척도 12항목 중에서 '잠자기' 항목이 부적합 항목으로 판정되었고, 퀘벡 요통 장애 척도 20항목은 부적합 항목이 없는 것으로 나타났다. 그러나 통합된 ODQ-QBPD-BPFS 42항목에서는 오스웨스트리 장애 설문지의 '통증의 정도'와 '성생활' 항목, 퀘벡 요통 장애 척도의 '공 던지기' 항목이 부적합 항목으로 판정되었다. 오스웨스트리 장애 설문지 평가도구를 분석 했을 때와 세 가지 평가도구를 통합하여 재분석 한 경우 '성생활' 항목이 부적합 항목으로 같은 판정을 받은 것은 라쉬모

형이 기대되는 피험자들의 반응과 실제로 피험자들의 응답하는 반응을 비교 분석하여 적합도 통계치를 산정하는데(지은림과 채선희, 2000), '성생활' 항목에서는 연구대상자들이 응답하지 않은 비율(33%)이 적합도 판정에 영향을 준 것으로 생각된다(김경모 등, 2011; Dallmeijer 등, 2005). 그리고 '통증의 정도' 항목은 0점(현재 통증이 전혀 없다), 1점(현재 매우 가벼운 통증이 있다), 2점(현재 통증이 조금 있다), 3점(현재 통증이 조금 심하다), 4점(현재 통증이 아주 심하다), 5점(현재 통증이 상상 할 수 없이 심하다)로 구분하는데, 이는 통증의 정도에 대한 개인차가 있고, 대상자가 표기하기에는 기준이 불명확하게 기술되어 대상자들에게 혼동을 유발하는 문항 제작의 오류 가능성이 적합도에 영향을 주었을 것으로 생각된다(지은림과 채선희, 2000).

라쉬분석에서 난이도는 계산된 로짓값 0을 기준으로 양(+)방향일 경우 어려운 항목, 음(-)방향일 경우 쉬운 항목으로 나누게 된다. 일반적으로 난이도 로짓값이 -5미만인 경우 쉬운 문항, -5이상에서 .5미만인 경우 중간정도 난이도 문항, .5 이상은 어려운 문항으로 해석된다(성태제, 2010). 본 연구에서는 부적합 항목으로 판정된 3항목을 포함하여 통합된 ODQ-QBPD-BPFS 42항목 난이도를 살펴본 결과 가장 어려운 항목은 '통증의 정도'였고, 가장 쉬운 항목은 '냉장고에서 음식 꺼내기'이었다. '통증의 정도' 항목은 오스웨스트리 장애 설문지를 이용한 박소연 등(2008)과 김경모 등(2011)의 연구에서와 마찬가지로 가장 어려운 항목으로 분류되고 있음을 알 수 있다. 그리고 42항목의 난이도 분포를 살펴본 결과 오스웨스트리 장애 설문지 2문항, 퀘벡 요통 장애 척도 9문항, 요통 기능 척도 2문항이 쉬운 문항으로 분류되었고, 중간정도 난이도 문항은 오스웨스트리 장애 설문지 6문항, 퀘벡 요통 장애 척도 7문항, 요통 기능 척도 5문항으로 분류되었으며, 어려운 문항은 오스웨스트리 장애 설문지 2문항, 퀘벡 요통 장애 척도 4문항, 요통 기능 척도 5문항으로 분류되었다. 항목 난이도 분류로 보았을 때 오스웨스트리 장애 설문지는 중간정도 난이도, 퀘벡 요통 장애 척도는 중간 정도 이하에서 쉬운 난이도 문항, 요통 기능 척도는 중간 이상에서 어려운 문항으로 구성되어 세 가지 평가도구 중에서 퀘벡 요통 장애 척도가 가장 쉬운 평가도구이고 요통 기능 척도가 가장 어려운 평가도구로 해석할 수 있다. 하지만 이런 결과는 평가도구 간의 상대적인 차이를 보여주는 것으로 직접적인 근거는 되지 않는다.

라쉬분석에서 분리지수는 신뢰도 값이 같아 구별이 어려워 상대적인 값을 비교하는 경우 사용하게 된다(이충휘, 2007). 이는 대상자 분리지수와 항목 분리지수 두 가지로 추정하고, 추정의 분리지수는 표준오차의 개념으로 설명되는데 항목이나 환자의 차이 구분이 잘 되었거나 독립적이라는 의미를 설명한다(박소연, 2008). 대상자 신뢰도(person reliability)는 대상자가 문항을 난이도에 따라 명확하게 구분해 낼 수 있는 척도를 말하며, .8이하인 경우 대상자가 응답을 제대로 하지 않았을 가능성이 높다. 항목 신뢰도(item reliability)는 측정하고자 하는 항목에서 대상자의 능력을 정확하게 구분하였는가를 표시하고 .8이상 이면 적당한 것으로 간주한다(이충휘, 2007). 본 연구에서는 모두 .8이상의 지수와 신뢰도를 나타내어 환자와 항목이 잘 구분되고 독립적이라고 할 수 있다. 또한 통합된 ODQ-QBPD-BPFS 42항목의 대상자 분리지수가 5.63으로 각각의 평가도구보다 높게 나타나고, 항목 분리지수는 모두 비슷한 수치를 보이는데, 일반적으로 문항의 수가 많을수록 대상자의 분리도가 높아지고, 대상자의 수가 많을수록 항목 분리지수가 높아지는 경향을 보이기 때문이다(지은림과 채선희, 2000).

본 연구에서는 요통의 특성이나 분류를 고려하지 않고 모든 요통 환자를 대상으로 자료를 수집하였기 때문에 평가도구의 특성이 정확하게 반영되지 못했을 가능성이 있고, 요통에 영향을 줄 수 있는 정신적, 사회적, 직업적 특성 등과 같은 다양한 배경도 함께 고려되지 못하였다. 또한 각각의 평가도구 항목 중에서는 중복되는 내용의 항목이 포함되어 있어 대상자의 응답에 영향을 주었을 가능성도 있다. 그러나 세 가지의 평가도구를 통합하여 분석한 결과 평가도구 항목의 난이도가 쉬운 것부터 어려운 문항까지 배열되어 통합된 하나의 요통 장애 평가도구의 개발이 가능함을 보여주었다. 이 후의 연구에서는 중복되는 항목의 재구성과 항목 난이도 등을 고려하고, 요통에 영향을 줄 수 있는 다양한 배경을 고려하여 우리나라 요통환자에 맞는 요통 장애 평가도구를 개발하는 연구가 필요하다고 생각된다.

V. 결론

본 연구는 18세 이상 성인 요통환자 314명을 대상으로 오스웨스트리 장애 설문지, 퀘벡 요통 장애 척도, 요통 기능 척도 등 세 가지 요통 장애 설문지를 이용하여

평가한 후 라쉬분석을 통해 단일구성개념, 항목 난이도와 분리지수 등을 알아보았다. 또한 세 평가도구를 통합하여 ODQ-QBPD-BPFS 42항목을 구성하여 라쉬분석을 실시한 결과, 42항목 중에서 가장 어려운 항목은 오스웨스트리 장애 설문지의 '통증의 정도' 이었고 가장 쉬운 항목은 퀘벡 요통 장애 척도의 '냉장고에서 음식 꺼내기'이었다. 통합된 ODQ-QBPD-BPFS 42항목의 항목 분리지수는 5.63으로 각각의 설문지의 분리지수보다 높아 항목의 난이도가 다양하게 구성되었지만, 같은 내용을 물어보는 항목의 난이도가 다르게 분석되어 항목의 재구성이 필요하다고 생각된다. 이후의 연구에서는 항목 적합도와 난이도를 고려하여 항목을 재구성하고, 요통에 영향을 끼칠 수 있는 정신적, 직업적, 사회적 요인을 고려하여 우리나라에 맞는 새로운 요통 평가도구를 개발하는 연구가 필요하다고 생각된다.

인용문헌

- 김경모, 박소연, 이충휘. 요통 환자의 일반적 특성에 따른 한국어판 오스웨스트리 장애 설문지의 라쉬분석. 한국전문물리치료학회지. 2011;18(2):35-42.
- 박소연. 노인에 낙상에 영향을 주는 요인을 평가하기 위한 ABC-BBS의 적용: 사전연구. 한국전문물리치료학회지. 2008;15(2):44-53.
- 박소연, 오재섭, 이충휘. 수정된 오스웨스트리 허리기능 장애 설문지의 라쉬분석: 산업장에서의 업무관련 요통환자를 대상으로. 한국전문물리치료학회지. 2008;15(3):26-34.
- 성태제. 현대교육평가 3판. 서울, 학지사, 2010.
- 이충휘. 물리치료사와 작업치료를 위한 연구방법론 3판. 서울, 계축문화사, 2007.
- 전창훈, 김동재, 김동준 등. 한국어판 Oswestry Disability Index(장애지수)의 문화적 개작. 대한척추외과학회지. 2005;12(2):146-152.
- 지은림, 채선희. 라쉬모형의 이론과 실제. 서울, 교육과학사, 2000.
- Beattie P, Maher C. The role of functional status questionnaires for low back pain. Aust J Physiother. 1997;43(1):29-38.
- Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: Summary and

- general recommendations. *Spine (Phila Pa 1976)*. 2000;25(24):3100-3103.
- Boos N, Rieder R, Schade V, et al. 1995 Volvo award in clinical science. The diagnostic accuracy of magnetic resonance imaging, work perception and psychosocial factors in identifying symptomatic disk herniations. *Spine (Phila Pa 1976)*. 1995;20(24):2613-2625.
- Chou R, Fu R, Carrino JA, et al. Imaging strategies for low-back pain: Systematic review and meta-analysis. *Lancet*. 2009;373(9662):463-472.
- Dagenais S, Tricco AC, Haldeman S. Synthesis of recommendations for the assessment and management of low back pain from recent clinical practice guidelines. *Spine J*. 2010;10(6):514-529.
- Dallmeijer AJ, Dekker J, Roorda LD, et al. Differential item functioning of the functional independence measure in higher performing neurological patients. *J Rehabil Med*. 2005;37(6):346-352.
- Davidson M. Rasch analysis of three versions of the Oswestry Disability Questionnaire. *Man Ther*. 2008;13(3):222-231.
- Davidson M, Keating JL. A comparison of five low back disability questionnaires: Reliability and responsiveness. *Phys Ther*. 2002;82(1):8-24.
- Delitto A. Are measures of function and disability important in low back care? *Phys Ther*. 1994;74(5):452-462.
- Fairbank JC, Couper J, Davies JB, et al. The Oswestry low back pain disability questionnaire. *Physiotherapy*. 1980;66(8):271-273.
- Grotle M, Brox JI, Vøllestad NK. Functional status and disability questionnaire: What do they assess? A systematic review of back-specific outcome questionnaires. *Spine (Phila Pa 1976)*. 2005;30(1):130-140.
- Haywood KL. Patient reported outcome I: Measuring what matters in musculoskeletal care. *Musculoskeletal Care*. 2006;4(4):187-203.
- Holm I, Friis A, Storheim K, et al. Measuring self-reported functional status and pain in patients with chronic low back pain by postal questionnaires. *Spine (Phila Pa 1976)*. 2003;28(8):828-833.
- Kopec JA, Esdaile JM, Abrahamowicz M, et al. The Quebec Back Pain Disability Scale: Measurement properties. *Spine (Phila Pa 1976)*. 1995;20(3):341-352.
- Lewis JS, Hewitt JS, Billington L, et al. A randomized control trial comparing two physiotherapy intervention for chronic low back pain. *Spine (Phila Pa 1976)*. 2005;30(7):711-721.
- Modic MT, Obuchowski NA, Ross JS, et al. Acute low back pain and radiculopathy: MR imaging findings and their prognostic role and effect on outcome. *Radiology*. 2005;237(2):597-604.
- Rocchi MB, Sisti D, Benedetti P, et al. Critical comparison of nine different self-administered questionnaires for the evaluation of disability caused by low back pain. *Eura Medicophys*. 2005;41(4):275-281.
- Roland M, Morris R. A study of the natural history of back pain, Part I: Development of a reliable and sensitive measure of disability in low back pain. *Spine (Phila Pa 1976)*. 1983;8(2):141-144.
- Stratford PW, Binkley JM, Riddle DL. Development and initial validation of the Back Pain Functional Scale. *Spine (Phila Pa 1976)*. 2000;25(16):2095-2102.
- Waddell G. *The Back Pain Revolution*. Edinburgh, Churchill Livingstone, 1998.
- Waddell G, Main CJ. Assessment of severity in low-back disorders. *Spine (Phila Pa 1976)*. 1984;9(2):204-208.

논문접수일 2011년 6월 22일

논문게재승인일 2011년 8월 5일