

원저

견관절의 임상적 평가도구에 관한 고찰

양동훈* · 김건형** · 김행범** · 최양식** · 박영배*

*경희대학교 한의과대학 진단·생기능의학과학교실

**경희대학교 한의과대학 침구학교실

Abstract

A Review of Clinical Shoulder Assessment Scales

Yang Dong-hoon*, Kim Kun-hyung**, Kim Haeng-beom**, Choi Yang-sik**
and Park Young-bae*

*Department of Biofunctional Medicine and Diagnosis, College of Oriental Medicine,
Kyung Hee University

**Department of Acupuncture & Moxibustion, College of Oriental Medicine,
Kyung Hee University

Objectives : A number of instruments have been developed to measure the quality of life in patients with various shoulder disorders. Much progress has been made in this area, and currently an appropriate instrument exists for each shoulder state. The purpose of this study is to review the instruments that are currently in use for assessing the shoulder joint.

Methods : A literature research was performed to choose appropriate scales for assessment of function and the disability of the shoulder. Theoretically based scales were selected for review. Therefore, 11 scales were reviewed. The status of scales involved in shoulder treatment of acupuncture throughout several countries was evaluated.

Results : 11 scales: The American shoulder and elbow surgeons evaluation form(ASES), Constant Shoulder Score, The disabilities of the arm, shoulder and hand(DASH), Shoulder Disability Questionnaire (SDQ), The Shoulder Pain And Disability index(SPADI), The simple shoulder test(SST), Oxford Shoulder Questionnaire(OSQ), The Rotator Cuff quality-of-life Measure(RC-QoL), Western Ontario Shoulder Instability Index(WOSI), Western Ontario Osteoarthritis of the Shoulder Index(WOOS), Western Ontario Rotator Cuff Index(WORC), wereevaluated. Each measurement has its own composition and characteristics. Their validity, reliability, responsiveness and practical characteristics were already evaluated. We found 3 domestic and 10 overseas papers about shoulder treatments using acupuncture assessed with shoulder scales.

· 접수 : 2007년 11월 9일 · 수정 : 2007년 11월 24일 · 채택 : 2007년 11월 25일
· 교신저자 : 박영배, 서울시 동대문구 회기동 경희대학교 부속한방병원 진단·생기능의학과
Tel. 02-958-9195 E-mail : bmppark@khu.ac.kr

Conclusions : In clinical research, the selection of the measurement scale should take account of the condition of disease, the patient's traits and the characteristics of the research. Moreover, appropriate scales, which have validity, reliability, responsiveness and practical characteristics, are needed to enhance the quality of research.

Key words : shoulder, questionnaire, scale, self-report

I. 서 론

다양한 견관절 상태를 가진 환자들에게서 삶의 질을 평가하기 위해 많은 수의 도구들이 개발되고 있다. 오래된 도구들은 적은 양의 정보만이 가능했을 때 도구 개발 방법론을 토대로 개발되어 왔지만, 현재 이 분야에서 많은 발전이 있었고, 현재 각각의 주요 어깨 상태에 대한 적절한 도구들이 개발되어 있다.

견관절 상태를 평가하는 도구의 'gold standard'는 존재하지 않으며, 임상연구에서는 각 질환과 환자의 특성이나 평가도구의 특징을 고려한 선택이 이루어지고 있다. 그러므로 임상연구를 진행할 때에는 정확성과 신뢰성을 바탕으로 만들어진 도구들을 선택해야 하며, 해당연구의 특성에 맞는 도구를 선별할 필요가 있다. 현재 유용한 어깨상태 측정 도구들에는 The American shoulder and elbow surgeons evaluation form (ASES), Constant Shoulder Score, The Disabilities of the arm, shoulder and hand(DASH), Shoulder Disability Questionnaire(SDQ), The Shoulder Pain And Disability index(SPADI), The simple shoulder test(SST), Oxford Shoulder Questionnaire(OSQ), The Rotator Cuff quality-of-life Measure(RC-QoL), Western Ontario Shoulder Instability Index(WOSI), Western Ontario Osteoarthritis of the Shoulder Index(WOOS), Western Ontario Rotator Cuff Index(WORC) 등이 있으며 이 도구들의 내용을 고찰해보고자 한다. 그리고 국내외에서 진행중인 견관절관련 침입상연구(clinical acupuncture study)에서 이 도구들이 사용되고 있는 현황을 살펴보고자 한다.

II. 본 론

1. 연구방법

한국연구정보서비스(RISS4U), NDSL 한국과학기술정보원, National center for biotechnology Information의 Pubmed, Science direct, Scopus에서 Keyword가 'shoulder'를 포함한 'outcome measurement', 'self-report', 'questionnaire', 'functional assessment', 'scoring system', 'index', 'validity', 'reliability', 'responsiveness'와 각 평가도구의 이름으로 검색하였다. 그 후 검색된 평가도구 중에서 설계방법이 적절하고 타당도, 신뢰도, 반응성, 편이성이 갖추어진 도구를 선별하였다. 선별기준은 설계방법에 대하여 논문에 명확한 방법과 기준을 제시하고, 내용타당도와 구성타당도를 갖추고 있으며, 논문으로 밝혀진 신뢰도를 확보한 도구를 선정하였고, 기준에 미치지 못하는 도구는 제외하였다. 도구의 구성과 특징을 요약 정리하였고, 각 평가도구의 타당도, 신뢰도, 반응성, 편이성에 대한 내용을 표로 정리하였다.

전부질환에 대한 침구치료 연구동향을 살펴보기 위해서, 국내논문 검색에는 KERIS 학술연구정보서비스, KISS 한국학술정보, DBPIA의 논문검색서비스와 대한침구학회의 홈페이지에서 1997년부터 2007년까지 발간된 논문 중, Keyword 검색으로 'acupuncture'와 'shoulder'로 검색된 논문을 대상으로 사용된 결과도구 등을 살펴보았다. 국외논문으로는 Pubmed를 이용하여 검색어 'shoulder AND acupuncture'로 검색된 논문 중 'Languages : English', 'default tag : Title'을 세부조건으로 설정한 후 1997년에서 2007년까지 발간된 것을 검색하여 정리하였다.

2. 도구의 구성과 특징¹⁻³⁾

1) The American shoulder and elbow surgeons evaluation form(ASES)^{4,5)}

1993년 미국견주관절외과학회(The Society American shoulder and elbow surgeons)는 어깨 기능 평가를 위한 표준양식을 개발했다. 사용하기 간편하고 환자의 자가평가 부분을 포함한 일상적 활동의 평가방법으로서 고안되었고 이 도구는 의사 평가 부분과 환자 자가평가 부분으로 구성되어 있다.

의사 평가 부분은 이학적 검사와 ROM(range of motion), 힘(strength), 불안정성, 특정 이학적 검사 등의 기록으로 이루어져 있다. 이 부분은 어떤 점수도 계산되지 않도록 되어 있다. 환자의 자가평가 부분은 11개의 항목으로 이루어져 있으며 이를 점수화시킬 수 있다. 이는 2개의 영역(통증 1항목, 기능 10항목)으로 나뉘어져 있다. 1개의 통증질문에 대한 대답은 10cm VAS(visual analogue scale)로 되어 있으며 0-10cm 사이 1cm씩 나뉘어져 작성하게 되어 있다.

기능 영역에서 10개의 항목은 일상적 삶의 활동성(예를 들어 화장실 가기, 코트 입기 등)으로 이루어져 있다. 또한 10파운드의 무게를 어깨 높이보다 높게 들기 또는 손 위로 공 던지기 등도 있다. 결국 2개의 일반적 항목이 있는 셈이다. 일상적 업무와 일상적 운동을 수행하는 항목에 대한 대답의 선택은 4개의 범주로 분류되어 있으며 0(할 수 없음)에서 3(어렵지 않음)까지이다. 이 때문에 개별 항목의 반응성(responsiveness)은 빈약할 수 있으며, 특히 기능이 잘 유지되고 있는 환자들에서 그러하다.

최종 점수화는 통증 점수(최대 10점)에 5를 곱하고(최대 50점), 활력도 측정의 합(최대 30점)에 5/3을 곱하여(최대 50점) 총점 100점으로 계산한다. 그러나 이 도구에서의 이러한 가중치 부여에 대한 설명이나 개념은 밝히지 않았다. 이 도구는 수술 후의 회복이나 경과에 대한 평가나 기능이상을 위주로 평가하며, 류머티즘으로 인한 견관절 이상을 평가하는 데는 근거나 증례가 부족하다.

2) The Constant score^{6,7)}

The Constant score는 유럽에서 어깨 평가 도구로 가장 폭넓게 사용되어 왔다. 이 척도방법은 환자의 주관적 평가와 함께 이학적 검사결과를 혼합한 것이다. 주관적 평가는 35점이 배점되어 있고, 이학적 검사에는 65점이 배점되어 있다.

주관적 평가 면에는 1개의 통증 점수(15점)와 4개의 일상생활 활동성(업무 4점, 운동 4점, 수면 2점, 손의 위치(positioning the hand in space) 10점)로 구성되어 있다.

객관적 평가 면에는, ROM(전방거상 10점, 외방 거상(lateral elevation) 10점, 내회전 10점, 외회전 10점), 힘(외전에 저항하여 환자가 밀 수 있는 파운드의 수에 기반한 점수화. 최대 25점) 따라서 총 점수는 100점 만점이다.

이러한 방법이 어떻게 개발되었고, 항목 선택의 기준 및 항목의 비교 가중치 부여에 대한 설명은 포함되어 있지 않지만 이 도구의 강점은 이 도구의 적용을 위한 방법이 꽤 명료히 기술되었다는 점이다.

이 도구의 개발자들이 왜 항목별로 다양한 가중치(통증 15%, 기능 20%, ROM 40%, 힘 15%)를 두었는지 알 수 없고 근거가 부족하다고 할 수 있다.

이 도구는 4개의 기능 항목과 5개의 이학적 검사 항목을 혼합하였다. 이는 근본적으로 다른 특성을 평가하는 것으로, 합쳐서 총점을 내는 방식과는 다르게, 분리되어 평가되어야 할 것이다. 이 도구는 ROM(40%)와 힘(25%)에 많은 가중치를 두고 있다. 이는 명확한 회전근개질환이나 골관절염 환자들을 감별해내는데 유용할 수는 있겠지만, 불안정성을 가진 환자들에겐 유용하지 않을 수 있다.

3) The Disabilities of the arm, shoulder and hand(DASH)^{8,9)}

최근 미국 골관절수술협회(AAOS)는 Institute for Work and Health와 함께 상지의 어떤 관절 어떤 질환에 대해서도 사용할 수 있는 도구를 개발했다. 이 도구는 팔·어깨·손의 장애 평가척도(DASH)라 불리며, AAOS에 의해 사용가능하게 되었다. 항목의 생성과 기본항목삭제에 대해 간단히 기술된 방법론이 출간되었다. 1999년에는 AAOS와 Institute는 DASH 척도의 사용자 매뉴얼을 개발하여 출간하였다.

DASH는 상지와 관련된 증상 및 장애의 정도에 따른 기능상태를 평가하기 위해 만들어진 30개 문항의 설문지이다. 장애는 “건강이나 신체적 문제 때문에 삶의 어떤 영역에서도 생활하는데 생기는 어려움”으로 정의하고 있다. DASH에서 다루고 있는 개념은 증상(통증, 위약, 강직, 저림/마비감), 신체적 기능(일상생활, 집안일, 쇼핑, 심부름, 레저활동, 자가관리, 옷 입기, 식사, 성생활, 수면, 운동/예술활동), 사회적 기능(가사업무, 친구/가족과의 사회활동), 정신적 기능(self-

image)들이다.

이 도구는 상지의 모든 관절에 질환을 가진 환자에게 적용하기 위한 것이며 타당도와 신뢰도 모두 확보하고 있다. 이는 환자가 현재 별 특정질환이 없는 경우 임상적으로 사용하는 것에 적합하다. 하지만 이 도구의 보다 넓은 영역은 임상시험에서 사용하기엔 부족한 면이 있다. 상지질환을 평가하기 위해 설정된 많은 항목들은 특정 상태에 놓인 환자들에게 모두 적합한 질문일 수는 없다. 또한 이 도구는 기타 어깨 전용 또는 어깨 상태의 특정 도구에 비해 연구도구로 덜 효율적일 수 있다는 단점이 있다.

4) Shoulder Disability Questionnaire (SDQ)¹⁰⁾

Van der Windt DA 등은 견관절에 관련된 기능장애를 평가하기 위해서 어깨장애 설문(SDQ)을 개발하였다. 16개의 통증관련 장애 질문으로 구성되어 있고, '예' 혹은 '아니오'로 대답하는 형식이다. 13개 항목은 업무와 관련된 상지기능을 평가하며, 나머지 3항목은 수면장애, 통증완화를 위해 어깨를 문지르는 정도, 불편한 기분을 측정한다. '예'이면 1점, '아니오'이면 0점으로 점수가 높을수록 상태가 불량하다.

이 도구는 간단하고 작성하기 쉽다. 그러나 일상생활에서의 연부조직 질환에서만 특정지어 사용될 수 있고, 거의 모든 항목이 통증과 관련되어 있기 때문에 통증이 적거나 심각한 질환을 가지고 있지만 아직 통증을 발현하지 않은 경우, 위약이나 움직임 제한으로 통증을 느낄 수 없는 경우에 이 도구는 유용하지 않다.

5) The Shoulder Pain And Disability index(SPADI)¹¹⁻¹⁵⁾

Roach는 1991년 어깨 통증 및 장애 척도(SPADI)의 개발과 평가에 대한 논문을 발표하였다. 저자는 "SPADI는 어깨 통증으로 인한 임상 증후와 관련된 장애와 통증을 반영하는 자가평가척도로서 개발되었다"라고 언급하였다. 이것은 감별도구(discriminative) 및 평가도구로서 설계되었다.

이 도구는 2개의 부속척도로 나뉘어진 13개의 평가항목으로 구성되어 있다. 이 도구에서 선택된 응답양식은 10cm VAS이다. 그러나 선 위에 점수를 매기는 일반적인 VAS와는 다르게, 저자들은 수평선을 동일한 등분으로 12등분하는 방법을 사용한다. 1-11까지의 번호가 각 등분에 부착되어, 각 항목의 점수를 평가하게 되어 있다.

개별항목의 점수화는 각 영역 내에서는 동일한 가중치를 가지며, 영역의 점수는 100점으로 환산된다. 0은 완벽한 상태이며 100은 가장 나쁜 상태를 가리킨다. 이 도구의 총점은 통증과 장애 2가지 영역의 점수를 평균 낸 것으로 결정된다. 이 도구는 높은 타당도(내용, 구성)와 민감도를 보인다.

6) The simple shoulder test(SST)¹⁶⁾

Lippitt, Harryman, Matsen은 1992년 SST의 개발과 검사에 대해 보고하였다. 이 도구의 목적은 진단에 대한 특정 외과외에 의해 시행된 수술로 인한 기능적 호전을 평가하고 문서화하기 위한 것과 상태의 심각한 정도를 정량화하기 위한 것이었다.

SST는 '예' 혹은 '아니오'의 12개 질문으로 구성되어 있다. 이 도구는 주관적 항목과 환자가 실제로 이학적 기능을 수행할 것을 요구하는 항목들로 혼합되어 있다. 예를 들어 "어깨로 인해 잠을 설치지 않고 편안히 잘 수 있습니까?" 등 주관적 질문과, "팔꿈치를 굽히지 않고 8파운드의 무게를 어깨 높이까지 들어 올릴 수 있습니까?" 등 실제 그 기술을 수행하길 요구하는 질문들이 있다.

SST는 임상적으로 중요한 변화에 대해 민감하지 않은 것 같다. 왜냐하면 2가지의 대답양식 '예' 혹은 '아니오' 밖에 없기 때문이다. 또한 이 도구는 같은 질환에서 다양한 중증도(심한 정도)를 나타내는 환자들 간을 구별하는 감별기능이 낮은 것으로 보인다.

7) Oxford Shoulder Questionnaire (OSQ)¹⁷⁾

Dawson, Fitzpatrick, Carr 등은 어깨 수술에 대한 환자들의 인식(perceptions)을 다루는 2개의 설문지를 발표하였다. 첫 번째 설문지는 Oxford shoulder questionnaire(OSQ)은 1996년 발간되었으며, 이는 어깨 수술을 받은 환자들을 위한 것이었다. 두 번째 설문지는 1999년 발간되었으며, 견관절 불안정성을 나타내는, 원래 설문지 대상에서 제외된 환자그룹을 위한 것이었다. 두 가지 설문지 모두 12개 항목으로 구성되었으며, 각 항목은 1에서 5까지, 최소한에서 최대한의 어려움이나 심각함까지이며, 혼합되어 12점(최고점)에서 60점(최저점)까지의 단일한 점수들을 산출한다.

2가지 설문지 모두 구성 타당도(construction validity)는 판별도구(discriminative instrument)로서의 다른 도구들과 비교한 전향적 조사를 시행하였다. 다른 어깨 도구와 적절한 상관성(correlation)을 보이고

Table 1. 평가도구의 구성과 특징

평가도구	개발자	구성 (문항개수-최고점수-총점대비 비율)	척도	문항수	범위 (점수)	응답 방식
ASES	Research Committee of the American Shoulder and elbow Surgeon	통증 : 1-10-50% 기능 : 10-30-50%	2	11	0-100	통증 : 10cm VAS(visual analog scale.) 기능 : 4항목 리커트척도
Constant Shoulder Score	Orthopedic surgeon	자가보고식(35%) 통증 : 1-15-15% 기능 : 4-20-20%			0-100	통증 : 4항목 리커트척도 기능 : 업무 0-4 ; 여가활동 0-4수면 0-2, 작업시 팔위치 2-10 ROM : 0-10점 힘 : 1점/1lb, 25lb까지.
		임상적 측정(65%) ROM : 4-40-40% 힘 : 1-25-25%				
DASH	American Academy of Orthopedic Surgeons, and Institute for Work and Health	증상 : 5-30-16.7% 기능이상 : 25-125-83.3%	1	30	0-100	증상과 이상 : 5점 리커트 척도
SDQ	Researchers Physical therapists	통증-관련 이상 : 16-16-100%	1	22	0-100	‘예/아니오’ 형식
SPADI	Physicians	통증 : 5-50-50% 기능/이상 : 8-80-50%	2	13	0-100	통증과 기능 : 10cm VAS
SST	Shoulder Service : University of Washington	기능 : 12-12-100%			0-12	‘예/아니오’ 형식
OSQ	University of Oxford and the Nuffield Orthopaedic Centre	통증-관련이상 : 12-5-100%	1	12	12-60	객관식
RC-QoL	University of Calgary sport medicine centre	신체적불편 : 16-100-46%, 운동/여가 : 4-100-12%, 업무관련 : 4-100-12%, 생활방식 : 5-100-15%, 사회적 감정적 활동 : 5-100-15%	1	34	0-100	100점 VAS
		신체증상 : 10-10-48%, 운동/여가/업무 : 4-10-19%, 생활기능 : 4-10-19%, 감정 : 3-10-14%	5	21	0-2100	10cm VAS
		신체기능 : 6-10-33%, 운동/여가 : 5-10-26%, 생활기능 : 5-10-26%, 감정 : 3-10-15%	4	19	0-1900	10cm VAS
The Western Ontario shoulder tools	Fowler·Kennedy sport medicine clinic, University of Western Ontario, London Ontario Canada	신체증상 : 6-10-29%, 운동/여가 : 4-10-19%, 업무기능 : 4-10-19%, 생활기능 : 3-10-14%	5	21	0-2100	10cm VAS

견관절평가의 전반적인 도구로서 적당해 보인다. 결국, 반응성이나 민감도(sensitivity)는 수술적 고정술을 받은 환자들에 대한 SF-36, HAQ, Constant score, Rowe score 등과 비교함으로써 측정되었다. 그 결과 일반적인 도구보다 이 도구가 보다 민감하였다. 이 설문지는 신뢰할 수 있고, 타당도 있으며, 반응성있는 정보를 제공할 수 있을 것으로 보인다.

8) The Rotator Cuff quality-of-life Measure(RC-QoL)^{18,19)}

Hollinshead는 2000년 10월 심하고 광범위한 회전근개 파열의 6년 추적을 보고하는 논문을 발표하였다. 이 글에서 회전근개 질환이 있는 환자들의 새로운 증상, 특이적 삶의 질 평가도구를 소개하였다. 이 도구는 Guyatt이 기술한 방법과 비슷한 방법을 사용하여 개발, 시험되었다. 이 도구는 '전 범위의 회전근개 질환' 환자들에 대한 결과 도구로서 사용되고 있다.

이 도구는 5개 영역으로 이루어진 34개 항목으로 구성되어 있다. 증상 및 신체적 불편(16개 항목), 체육/여가(4항목), 업무관련(4항목), 생활방식(5항목), 사회적 감정적 활동(5항목)이다. 답변양식은 가장 높은 점수가 100mm, 가장 낮은 점수가 0mm이다. 저자는 원 점수(raw score; 0에서 3,400까지. 0=최저 점수. 3,400=최고 점수)를 백분율로 표시하는 것을 권장하고 있다.

9) The Western ontario shoulder tools

Kirkley는 1998년 WOSI²⁰⁾(어깨를 위한 질환특징적 삶의 질 평가도구 시리즈 제1차)를 발간하였다. 이 도구는 Kirschner와 Guyatt에 의해 기술된 방법론을 사용하여 개발, 평가되었다. 이 도구는 어깨 불안정성을 가진 환자들에 대한 임상연구에서 치료를 평가하는 제1차 결과 평가도구로서 고안되었다.

2001년, WOOS²¹⁾(어깨를 위한 질환특징적 삶의 질 평가도구 시리즈 중 제2차)가 발간되었다. 저자들은 이 도구가 WOSI를 개발할 당시의 방법론과 유사한 방법을 이용하여 개발되고 평가되었다고 기술하였다. WOOS는 어깨의 증상이 있는 원발성 관절염 환자들에 대한 임상연구 평가의 제1차 결과평가 도구로 쓰일 목적이라고 하였다.

가장 최근인 2003년 시리즈 중 제3차로 WORC²²⁾가 개발되었다. 이는 회전근개의 퇴화 환자들에 대한 치료를 평가하는 임상연구의 제1차 결과평가도구로 쓰일 것이었다. WORC 역시 Kirschner와 Guyatt에 의해 기술된 방법론을 이용하여 개발, 평가되었다.

고정된 탈구를 제외하고 임상적으로 연관된 모든 범주에서의 유증상성 어깨 불안정성에 대한 해당/배제기준에 맞는 환자들의 데이터베이스가 마련되었다. 또 회전근개건염, 파열이 없는 회전근개건염, 회전근개 파열, 회전근개 관절병증(arthropathy)을 포함한 유증상성 회전근개질환에 대한 해당/배제기준에 맞는 환자들의 데이터베이스가 마련되었다. 비슷하게 원발성 견관절염 진단을 받은 모든 연령대 환자들의 데이터베이스도 마련되었다.

각 도구들은 환자들에 대한 설명, 각 항목에 대한 설명, 임상가를 위한 상세한 점수와 설명 등을 포함하고 있다. 저자들은 임상연구에서 1차 결과를 위해 전체 점수를 사용할 것을 권장하지만, 각 개별 영역의 점수 역시 보고할 것을 권장하고 있다. 점수는 원상태 그대로 볼 수도 있고, 백분율로 환산해서 볼 수도 있다. 가장 좋은 점수는 100%(원점수=0)이며, 이는 어깨와 관련된 삶의 질에서 아무런 저하가 없음을 나타낸다. 가장 나쁜 점수는 0%(원점수=WOSI와 WORC에서 2100점, WOOS에서 1900점)이며 어깨와 관련된 삶의 질에서 극심한 저하가 있음을 나타낸다

10) 요약

각 도구의 구성과 특징은 Table 1과 같다. 개발자, 구성(문항개수-최고점수-총점대비 비율), 척도, 문항수, 범위(점수)를 조사하였다.

3. 도구의 측정 평가

1) 타당도

(1) 내용

연구의 목적을 달성하기 위해서는 연구에서 밝히고자 하는 속성을 측정해야 하며 직접 측정하는 것이 불가능할 경우에는 검사 도구를 사용한다. 타당도(validity)²³⁻²⁵⁾는 측정하고자 하는 것을 얼마나 충실히 측정하였느냐, 즉 검사점수가 검사의 사용목적에 얼마나 부합되는느냐의 문제다. 다시 말해 연구에서 측정하고자 하는 변수를 검사가 제대로 측정하였느냐가 타당도이며, 이는 검사도구의 목적에 대한 적합성에 해당한다.

(2) 종류

① 내용타당도(content validity)

실험 또는 연구설계에 있어, 측정하고자 하는 내용이 조사대상의 주요 국면을 대표할 수 있느냐 하는 판단과 관련된 타당성을 말한다. 측정도구 자체가 측

정하고자 하는 속성이나 개념을 측정할 수 있는지와 연관된다.

② 기준관련타당도(criterion-related validity)

기준관련타당도는 이미 타당성이 입증되거나 지금까지 널리 사용되고 있는 기준을 타당성의 평가 잣대로 사용한다. 통계적인 유의성을 평가하는 것으로 어떤 측정도구와 측정결과인 점수간의 관계를 비교하여 타당도를 파악하는 방법이다. 이미 타당성이 검증된 기준과 관련시켜서 타당성을 검증하기 때문에 경험적 타당성이라고 한다.

③ 구성타당도(construct validity)

조사설계에 있어 처리, 결과, 모집단 및 상황들에 대한 이론적 구성요소들이 성공적으로 조작화된 정도를 의미한다. 즉 조사항목 등이 관련된 이론의 구성(theoretical construct) 또는 가설과 얼마나 부합하는가 하는 문제와 관련된 타당성을 말한다.

2) 신뢰도(reliability)

(1) 신뢰성의 개념²³⁾

측정된 결과치의 안정성, 일관성, 예측가능성, 정확성 등이 내포된 개념 측정도구가 측정하고자 하는 현상을 일관성 있게 측정하는 능력 또는 동일한 개념에 대해 측정을 반복했을 때 동일한 측정값을 얻을 가능성을 의미한다.

(2) 신뢰도의 추정방법

① 재검사법(test-retest method)

동일한 측정도구를 이용하여 동일한 상황에서 동일한 대상에게 일정기간을 두고 반복 측정하여 최초의 측정치와 재측정치가 동일한지의 여부를 평가하는 방법으로, 측정기간동안 대상의 속성이 변하지 않는 것을 전제로 한다.

그룹내 상관계수(ICC, intraclass correlation coefficient)나 카파 계수(kappa coefficient)같은 통계적인 방법이 이용된다. 재검사법의 계수가 0.75보다 클 경우 통계적으로 유의하다. 장점은 측정도구 자체를 직접 비교할 수 있어 적용이 간편하나, 검사요인효과(처음 측정이 재검사점수에 영향을 미치는 효과), 성숙요인효과(측정간격이 길 때에 조사대상집단의 특성변화에 따른 효과), 역사요인효과(측정기간중에 발생한 사건의 영향)로 인해 오류가 발생할 수 있다.

② 내적 일관성 분석(internal consistency analysis)²⁴⁾

개별문항을 하나하나의 검사점수로 하여 상관관계를 구하고 이들의 평균상호상관관계를 신뢰도의 추정치로 삼는 방법으로 평균상호상관은 모든 가능한 반복법을 사용하여 구한 신뢰도 계수의 평균을 의미한다.

신뢰계수 추정법 중 가장 알려진 방법은 크론바(cronbach)의 알파계수(α coefficient)^{26,27)}인데 0~1범위이며, 1에 가까울수록 항목들 사이에 긴밀한 관계나 상관성이 있음을 의미한다. 일반적으로 집단수준인 경우에는 크론바 α 계수가 0.6 이상이 되어야 하고, 개별수준인 경우에는 0.9 이상이 되어야 한다.

3) 반응성(responsiveness)²⁸⁻³⁰⁾

반응성은 상태의 변화를 반영하는 정도한다. 반응성이 높다는 것은 치료나 중재, 시간의 흐름에 따른 변화에 따른 민감한 차이를 측정할 수 있음을 의미한다.

반응성을 측정하는 방법으로는 표준반응평균(SRM, standardized response mean)³¹⁾, 효과 크기(ES, effect size)³⁰⁾, 상대적 효율성(relative efficiency)³²⁾, ROC(receiveroperating characteristic)곡선, 반응성 Guyatt 척도(Guyatt index of responsiveness)³³⁾가 있다.

SRM과 ES는 가장 많이 사용하는 방법이며 식은 아래와 같다. ES는 표준편차(standard deviation)를 사용하는 점을 제외하고 SRM과 유사하다.

SRM과 ES의 해석은 표준화된 방법이 없다. 일반적으로 점수가 높으면 반응성이 높다고 평가된다. Guyatt 등은 SRM이 0.8을 기준으로 반응성을 인정할 수 있다고 제시하였다³⁴⁾. 코헨의 경험규칙(Cohen's rule of thumb)에 따르면, ES나 SRM이 0.2는 약한 반응성, 0.5 중간정도의 반응성, 0.8 이상은 강한 반응성을 의미한다³⁵⁾. SRM와 ES의 공식은 아래와 같다.

$$SRM = \frac{\text{점수변화의 평균}}{\text{변화점수의 표준편차}}$$

$$ES = \frac{\text{점수변화의 평균}}{\text{초기점수의 표준편차}}$$

4) 편이성

환자의 기능과 이상을 제대로 평가하기 위한 검사도구의 중요한 특성은 편이성이다.

검사도구는 환자가 쉽게 작성할 수 있도록 짧고, 이해하기 쉬워야 하며, 검사자가 알아보기 편해야 한다. 그러나 환자의 운동제한이나 이상을 측정할 수 있는 적절한 항목들은 반드시 포함되어 있어야 한다.

검사자는 환자가 작성한 설문지를 쉽고 빠르게 이해하고, 정량화할 수 있어야 한다. 예를 들어 시각상사척도(VAS, visual analogue scale)의 경우 이상이 없는 경우부터 가장 심한 상태까지 선 위에 표시하게 되는데 검사자가 이를 수치화시키기 위해서는 반드시 자가 필요하게 되는데, 이보다 수치등급평가(NRS, numeric rating scale)는 수치를 직접 확인할 수 있으므로 VAS보다 높은 편이성을 갖는다고 할 수 있다.

또한 검사도구는 임상이나 연구에 적용하기 용이하

도록 고안·설계되어야 하며, 환자의 상태를 파악할 때 유용하여 치료의 방향을 결정하는데 임상적으로 쓸모 있어야 한다.

5) 요약

각 도구의 타당도, 신뢰도(재검사법), 신뢰도(내적 일관성), 반응성을 조사하였고 편이성과 관련된 작성시간(분), 점수계산시간(분)을 정리하였는데 Table 2와 같다.

Table 2. 도구의 측정 평가

Scale	타당도	신뢰도		반응성	편이성	
		재검사법	내적 일관성		작성시간(분)	점수계산시간(분)
ASES	내용타당도 구성타당도 판별타당도	ICC : 0.84	α : 0.86	SRM : 1.6 ES : 1.4	3-5	2
Constant Shoulder Score	내용타당도 구성타당도 판별타당도	ICC : 0.80	*	SRM : 0.59	자가보고식 3	1
					임상적측정 1	1
DASH	내용타당도 구성타당도 판별타당도	ICC : 0.92	α : 0.96	SRM : 1.13 ; 0.70	6	3
SDQ	내용타당도 구성타당도	Inferred, no statistics ⁴⁹⁾	*	Guyatt index of responsiveness : 1.89-2.22, 1.89-2.14	3	1
SPADI	내용타당도 구성타당도 판별타당도	ICC범위 : 0.64-0.66	α range : 0.86-0.95	SRM : 1.38, 1.23	7	8
SST	내용타당도 구성타당도	ICC : 0.99	*	SRM : 0.8	3	1
OSQ	내용타당도 구성타당도	*	α : 0.94	*	7	1
RC-QoL	내용타당도 구성타당도	overall average error difference : 5.05% ⁵⁰⁾	*	SRM : 1.43	15	5
WOSI	내용타당도 구성타당도	ICC범위 : 0.949-0.911	*	SRM : 0.931	10	10
WOOS	내용타당도 구성타당도 판별타당도	ICC : 0.964	*	SRM : 1.910	10	10
WORC	내용타당도 구성타당도	ICC : 0.96-0.98	α : 0.92	SRM : 1.44	10	10

SRM : standardized response mean. ES : effect size, ICC : intraclass correlation coefficient. α : Cronbach alpha coefficient. Values calculated on the basis of reported values in original paper. * : not Scored.

Table 3. 국내 견부질환 침구치료 임상논문 중 평가도구 사용논문

Study	Outcome measure					
	shoulder-specific scale	QoL	10-point VAS	ROM	PPT	기타
남 2007 ³⁶⁾	SPADI CSA		O	O		DITI
남 2006 ³⁷⁾	SPADI CSA		O			
이 2006 ³⁸⁾	SPADI		O	O		Patient's global assessment

QoL : Quality of Life Questionnaire, ROM : Range of Motion. PPT : Pressure Pain Threshold. DITI: Digital Infrared Thermal Imaging.

Table 4. 국외 견부질환 침구치료 논문 중 평가도구 사용 논문

Study	Outcome measure						
	shoulder scale	QoL	10-point VAS	ROM	PPT	etc.	Impact Factor
Dyson-Hudson 2007 ³⁹⁾	WUSPI		O			intake questionnaire weekly self questionnaire	1.826 a
Shin 2007 ⁴⁰⁾				O		motor power level	0.831 a
Ma T 2006 ⁴¹⁾		SF-36	O	O			0.710
He D 2005 ⁴²⁾			O			questionnaire -Quality of sleep (9-point likert scale) -pain-related activity at work -pain-related activity at home	b
Guerra 2004 ⁴³⁾	SPADI	COOP/WONCA chart	O	O		Lattinen index(pain intensity) Number of pills	4.601
He D 2004 ⁴⁴⁾			O		O	questionnaire - intensity of pain - frequency of pain - degree of pain	4.061
Jin D 2003 ⁴⁵⁾						3-point likert scale (pain & movement)	b
Nabeta 2002 ⁴⁶⁾			O		O		1.507
Sun KO 2001 ⁴⁷⁾	CSA						b
Dyson-Hudson 2001 ⁴⁸⁾	WUSPI		O	O		Intake questionnaire weekly log. 6-point verbal rating scale	1.371

QoL : Quality of Life Questionnaire, ROM: Range of Motion, PPT: Pressure Pain Threshold, DITI : Digital Infrared Thermal Imaging, WUSPI : Wheelchair User's Shoulder Pain Index, COOP/WONCA chart : Dartmouth Primary Care Cooperative Information Project chart (WONCA means World Organization of National Colleges, Academies, and Academic Associations of General Practices/Family Physicians). COOP/WONCA chart is international version of COOP chart. a : 동일 Journal의 2006년도 Impact Factor. b : not SCI journal.

4. 견부질환에 대한 침치료 연구동향

1) 국내연구

국내의 경우 견부질환에 대한 평가도구를 사용한 연구는 남^{36,37)} 및 이³⁸⁾ 등 3개가 있다. 3개 논문 모두 유착성 관절낭염 환자를 대상으로 하여 침치료 전후의 효과를 SPADI 혹은 CSA 등을 이용하여 비교, 평가하였으며, 10점 VAS 및 ROM 등은 2차 평가도구로서 보조적으로 사용하였다(Table 3).

2) 국외연구

총 18개의 논문이 검색되었으며 이 중 침 임상연구 논문은 10개였다. 국외의 경우 10개 침 임상연구 논문 중 Impact Factor가 산정된 SCI 저널에 등재된 논문은 7개였다. 견부질환에 대한 평가도구가 사용된 것은 WUSPI가 사용된 Trevor³⁹⁾ 및 SPADI가 사용된 Juan⁴⁵⁾의 논문 3가지였다. 비 SCI 논문의 경우 CSA를 사용한 Ko⁴⁷⁾의 연구 1례가 있었다. 기타 견부질환의 증상 및 원인의 다양성에 따라 VAS, ROM, pressure algometry를 이용한 pressure pain threshold, weekly log 및 해당연구에 맞게 변형 및 자체제작된 설문지와 평가척도를 조합하여 사용하였다(Table 4).

III. 결 론

견관절 평가도구는 전문가 집단이나 대학, 위원회 등에 의해 개발되며, 평가도구의 목적에 따라 다양한 내용과 구성을 가진다. 의사 평가 부분과 환자 자가평가 부분으로 구성되어 있으며 자가보고식만으로 구성되기도 한다. 도구에 따라 평가항목은 기본적으로 통증, 기능, ROM, 힘(strength)같은 이학적 평가는 물론, 신체적 불편, 업무, 사회적 감정적 활동, 운동/여가같은 삶의 질을 평가하는 내용을 포함하고 있다.

어깨 상태를 평가하기 위한 오래된 도구들은 거의 정보가 없었거나 이러한 노력들의 적당한 방법론에 대한 관심이 별로 없을 당시에 만들어진 경우가 많다. 그러나, 이제는 어깨의 특정 상태를 평가하는데 뛰어난 많은 수의 도구들이 존재한다. 특별한 환자집단에서 이 도구들을 평가하고, 각 도구들간 최소한 임상적으로 중요한 차이를 확인하고, 각 국가에 맞는 국제적으로 사용될 수 있는 타당한 번역을 이행해야 할 것이다.

임상연구를 계획하는 조사자들은 환자와 해당질환에 맞는 도구를 선택하고, 타당도(validity)와 신뢰도(reliability), 반응성, 편이성이 갖추어진 적절한 도구를 선택해야 할 것이다. 견관절의 임상적 평가도구는 국내외에서 견부질환의 임상과 연구분야에서 치료의 평가를 위해 활발하게 사용되고 있고, 좀더 정확한 평가와 편리한 사용이 가능한 도구의 개발과 기존 도구의 개선을 위해 앞으로 지속적인 연구와 발전을 요할 것으로 사료된다.

IV. 참고문헌

1. Lori A Michener, Brian G Leggin. A review of self-report scales for the assessment of functional limitation and Disability of the Shoulder. *Journal of Hand Therapy*. 2001 ; 14(2) : 68-76.
2. Philip McClure, Lori Michener. Measures of adult shoulder function: The American Shoulder and Elbow Surgeons Standardized Shoulder Form Patient Self-Report Section(ASES), Disabilities of the Arm, Shoulder, and Hand(DASH) Shoulder Disability Questionnaire, Shoulder Pain and Disability Index(SPADI) and Simple Shoulder Test. *Arthritis & Rheumatism(Arthritis Care & Research)*. 2003 ; 49(5) : 50-8.
3. Bot SDM, Terwee CB, van der Windt DAWM, Bouter LM, Dekker J, de Vet HCW. Clinimetric evaluation of shoulder disability questionnaires : a systematic review of the literature. *Ann Rheum Dis*. 2004 ; 63, 335-41.
4. Richards RR, An KN, Bigliani LU, Friedman RJ, Gartsman GM, Gristina AG et al. A standardized method for the assessment of shoulder function. *J Shoulder Elbow Surg*. 1994 : 347-2.
5. Michener LA, McClure PW, Sennett BJ. American Shoulder and Elbow Surgeons Standardized Shoulder Assessment Form, patient self-report section : reliability, validity, and responsiveness. *J Shoulder Elbow Surg*. 2002 ; 11 : 587-4.
6. Constant CR, Murley AHG. A clinical method of functional assessment of the shoulder. *Clin Orthop*. 1987 ; 214 : 160-4.

7. Conboy VB, Morris RW, Kiss J, Carr AJ. An evaluation of the Constant-Murley Shoulder Assessment. *J Bone Joint Surg Br.* 1996 ; 78 : 229-32.
8. Hudak P, Amadio PC, Bombardier C, and the Upper Extremity Collaborative Group. Development of an upper extremity outcome measure : The DASH(Disabilities of the Arm, Shoulder, and Hand). *Am J Ind Med.* 1996 ; 29 : 602-8.
9. Hunsaker FG, Cioffi DA, Amadio PC, Wright JG, Caughlin B. The American Academy of Orthopedic Surgeons Outcomes Instruments : normative values from the general population. *J Bone Joint Surg Am.* 2002 ; 84 : 208-5.
10. Van der Windt DA, van der Heijden GJ, de Winter AF, Koes BW, Deville W, Bouter LM. The responsiveness of the Shoulder Disability Questionnaire. *Ann Rheum Dis.* 1998 ; 57 : 82-7.
11. Roach KE, Budiman-Mak E, Sangsiridej N, Lertrantanakul Y. Development of a shoulder pain and disability index. *Arthritis Care Res.* 1991 ; 4 : 143-9.
12. Williams JW, Holleman DR, Simel DL. Measuring shoulder function with the Shoulder Pain and Disability Index. *J Rheumatol.* 1995 ; 22 : 727-2.
13. Beaton DE, Richards RR. Assessing the reliability and responsiveness of 5 shoulder questionnaires. *J Shoulder Elbow Surg.* 1998 ; 7 : 565-2.
14. Beaton DE, Richards RR. Measuring function of the shoulder : a cross-sectional comparison of five questionnaires. *J Bone Joint Surg Am.* 1996 ; 78 : 882-90.
15. Roddey TS, Olson SL, Cook KF, Gartsman GM, Hanten W. Comparison of the University of California-os Angeles Shoulder Scale and the Simple Shoulder Test with the Shoulder Pain and Disability Index : single administration reliability and validity. *Phys Ther.* 2000 ; 80 : 759-68.
16. Lippitt SB, Harryman DT, Matsen FA. A practical tool for evaluation of function : the simple shoulder test. In : Matsen FA, Fu FH, Hawkins RJ(eds). *The Shoulder. a balance of mobility and Stability.* Rosemont(IL). American Academy of Orthopedic Surgery ; 1993.
17. Dawson J, Fitzpatrick R, Carr A. Questionnaire on the perceptions of patients about shoulder surgery. *J Bone Joint Surg Br.* 1996 ; 78 : 593-600.
18. Hollinshead RM, Mohtadi NG, Vande Guchte RA, Wadey VM. Two 6-year follow-up studies of large and massive rotator cuff tears : Comparison of outcome measures. *J Shoulder Elbow Surg.* 2000 ; 9 : 373-81.
19. Guyatt GH, Townsend M, Berman LB, Keller JL. A comparison of Likert and visual analogue scales for measuring change in function. *J Chronic Dis.* 1987 ; 40 : 1129-33.
20. Kirkley A, Griffin S, McLintock H, Ng L. The development and evaluation of a disease-specific quality of life measurement tool for shoulder instability : The Western Ontario Shoulder Instability Index(WOSI). *Am J Sports Med.* 1998 ; 26 : 764-72.
21. Lo IKY, Griffin S, Kirkley A. The development and evaluation of a disease-specific quality of life measurement tool for osteoarthritis of the shoulder : The Western Ontario Osteoarthritis of the Shoulder Index(WOOS). *Arthritis Cartilage.* 2001 ; 9 : 771-8.
22. Kirkley A, Griffin S, Alvarez C. The development and evaluation of a disease-specific quality of life measurement tool for rotator cuff disease : The Western Ontario Rotator Cuff Index(WORC). *Clin J Sport Med.* 2003 ; 13 : 84-92.
23. Portney LG, Watkins MP. *Foundations of Clinical Research : Application to Practice.* 2nd Ed. Upper Saddle River, NJ : Prentice Hall. 2000.
24. Nunnally JC, Bernstein IH. *Psychometric Theory.* New York : McGraw-Hill. 1994.
25. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med.* 1993 ; 118, 622-9.
26. Cronbach LJ. coefficient alpha and the internal structure of tests. *Psychometrika.* 1951 ; 16 : 297-334.
27. 한상숙, 이상철. *간호·보건 통계분석.* 서울 : 현문사. 2004.

28. Guyatt GH, Walter S, Norman G. Measuring change over time : assessing the usefulness of evaluative instruments. *J Chronic Dis.* 1987 ; 40 : 171-8.
29. Kantz ME, Harris WJ, Levitsky K, Ware JE, Davies AR. Methods for assessing condition-specific and generic functional status outcomes after total knee replacement. *Med Care.* 1992 ; 30 : 240-52.
30. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change : An analogy to diagnostic test performance. *Journal of Chronic Diseases.* 1986 ; 39(11) : 897-906.
31. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Medical care.* 1990 ; 28(7) : 632-42.
32. Liang MH, Larson MG, Cullen KE, Schwartz JA. Comparative measurement efficiency and sensitivity of five health status instruments for arthritis research. *Arthritis and Rheumatism.* 1985 ; 28(5) : 542-7.
33. Guyatt GH, Kirshner B, Jaeschke R, Williams, JI, Naylor CD. Measuring health status : What are the necessary measurement properties? *Journal of Clinical Epidemiology.* 1992 ; 45(12) : 1341-55.
34. Guyatt G, Walter S, Norman G. Measuring change over time : Assessing the usefulness of evaluative instruments. *Journal of Chronic Diseases.* 1987 ; 40(2) : 171-8.
35. Cohen J. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* Hillsdale, NJ : Lawrence Erlbaum. 1983.
36. 남동우, 임사비나, 김종인, 김건식, 이두익, 이재동, 이윤호, 최도영. 유착성 관절낭염에 대한 침 및 신경차단술 처치의 임상적 관찰. *대한침구학회지.* 2007 ; 24(4) : 143-55.
37. 남동우, 정인태, 김주희, 박유선, 임사비나, 이두익, 이재동, 이윤호, 최도영. 유착성 관절낭염 치료에 있어서 한방 치료와 양방 치료의 임상적 고찰-동서협진 모델 개발을 위한 기초 연구를 중심으로-. *대한침구학회지.* 2006 ; 23(5) : 177-85.
38. 이현, 홍권의, 김영일, 임윤경, 안택원, 강위창, 최선미. Frozen Shoulder의 침치료에 대한 임상연구. *대한침구학회지.* 2006 ; 23(1) : 165-77.
39. Dyson-Hudson TA, Kadar P, LaFountaine M, Emmons R, Kirshblum SC, Tulsy D, Komaroff E. Acupuncture for chronic shoulder pain in persons with spinal cord injury : a small-scale clinical trial. *Arch Phys Med Rehabil.* 2007 ; 88 : 1276-83.
40. Shin BC, Lim HJ, Lee MS. Effectiveness of combined acupuncture therapy and conventional treatment on shoulder range of motion and motor power in stroke patients with hemiplegic shoulder subluxation : a pilot study. *Int J Neurosci.* 2007 ; 117 : 519-23.
41. Ma T, Kao MJ, Lin IH, Chiu YL, Chien C, Ho TJ, Chu BC, Chang YH. A study on the clinical effects of physical therapy and acupuncture to treat spontaneous frozen shoulder. *Am J Chin Med.* 2006 ; 34(5) : 759-75.
42. He D, Høstmark AT, Veiersted KB, Medbø JI. Effect of intensive acupuncture on pain-related social and psychological variables for women with chronic neck and shoulder pain-an RCT with six month and three year follow up. *Acupunct Med.* 2005 ; 23(2) : 52-61.
43. Guerra de Hoyos JA, Andrés Martín Mdel C, Bassy Baena de Leon E, Vigára Lopez M, Molina López T, Verdugo Morilla FA, González Moreno MJ. Randomised trial of long term effect of acupuncture for shoulder pain. *Pain.* 2004 ; 112(3) : 289-98.
44. He D, Veiersted KB, Høstmark AT, Medbø JI. Effect of acupuncture treatment on chronic neck and shoulder pain in sedentary female workers : a 6-month and 3-year follow-up study. *Pain.* 2004 ; 109(3) : 299-307.
45. Jin D, Li Z. Acupuncture and the elevation manipulation of massage for treatment of frozen shoulder. *J Tradit Chin Med.* 2003 ; 23(3) : 212-3.
46. Nabeta T, Kawakita K. Relief of chronic neck and shoulder pain by manual acupuncture to tender points-a sham-controlled randomized trial. *Complement Ther Med.* 2002 ; 10(4) : 217-22.
47. Sun KO, Chan KC, Lo SL, Fong DY. Acupunc-

- ture for frozen shoulder. Hong Kong Med J. 2001 ; 7(4) : 381-91.
48. Dyson-Hudson TA, Shiflett SC, Kirshblum SC, Bowen JE, Druin EL. Acupuncture and Trager psychophysical integration in the treatment of wheelchair user's shoulder pain in individuals with spinal cord injury. Arch Phys Med Rehabil. 2001 ; 82(8) : 1038-46.
49. van der Heijden GJ, Leffers P, Bouter LM. Shoulder disability questionnaire design and responsiveness of a functional status measure. J Clin Epidemiol. 2000. Jan ; 53(1) : 29-38.
50. Hollinshead RM, Mohtadi NG, Van de Guchte RA, Wadey VM. Two 6-year follow-up studies of large and massive rotator cuff tears : comparison of outcome measures. J Shoulder Elbow Surg. 2000 ; 9 : 373-81.