

# La disabilità e la qualità della vita dei pazienti con artrite reumatoide: valutazione e prospettive

## *Disability and quality of life of patients with rheumatoid arthritis: assessment and perspectives*

F. Salaffi, A. Stancati

Clinica Reumatologica e Cattedra di Reumatologia, Università Politecnica delle Marche

### SUMMARY

Functional disability and quality of life in rheumatoid arthritis (RA) are key outcomes that determine patient's demand for care, and influence their compliance and satisfaction with treatment. In the recent years the demand to collect outcome data is ever increasing as a means for the validation of quality care, and the development of effective outcome measures has become a major thrust of health research and has contributed to better understanding the relationship between outcomes and specific elements of health care. There are several disease-generic and specific instruments available that have proven valuable in outcome testing in RA. The first instruments provide a broad picture of health status across a range of conditions, whereas the latter are more sensitive to the disorder under consideration and are therefore more likely to reflect clinically important changes. When necessary, this kind of scales can be supplemented with specialised domain-specific scales (for the assessment of psychological well-being, social role functioning, or other). As in other fields, these measurement instruments mainly focus on:

- a) clinical signs and symptoms (physiologic and biologic);
- b) physical and/or cognitive functioning;
- c) well-being and emotional functioning;
- d) social functioning;
- e) satisfaction with care and other personal constructs (life satisfaction, spirituality, etc.);
- f) health-related quality of life (HRQOL).

Over the past 20 years, there has been a better recognition of the patient's point of view as an important component in the assessment of health care outcomes, and an increasing interest in HRQOL as an important area of research, due to the rising burden of chronic diseases, longer expectation of life, the growing number of health intervention alternatives, and greater emphasis on humanising health care. In addition, decision-making on issues of cost-effectiveness across health inputs and resource allocation across health programs is likely to be more sound if informed by HRQOL evidence. This paper reviews the literature and discusses the major issues regarding mainly measures of physical function (e.g., mobility or daily activities) and health status, including some so called HRQOL instruments.

Reumatismo, 2004; 56 - N. 1 (Suppl. 1):87-106

## 1. INTRODUZIONE

L'artrite reumatoide (AR) per il carattere progressivo, per l'impatto disabilitante e per gli elevati costi sociali diretti e indiretti costituisce uno dei problemi medico-sociali più rilevanti e, al tempo stes-

so, poco considerati (1-7). Troppo spesso viene trascurato l'aspetto cronico e invalidante di tale malattia, che rende indispensabile un corretto approccio diagnostico e un precoce intervento terapeutico. Ciò comporta l'esigenza di una valutazione globale dello stato di salute del paziente, che tenga conto non solo delle caratteristiche clinico-biologiche della malattia, ma anche dei potenziali esiti invalidanti. Gli esiti possono essere misurati a vari livelli: le alterazioni anatomico-patologiche di base, la compromissione funzionale (*impairment*), la disabilità e l'handicap (8). Così, ad esempio, nel caso di un'artrite, il processo anatomico-patologico

Indirizzo per la corrispondenza:

Dott. Fausto Salaffi

Clinica Reumatologica - Università Politecnica delle Marche

Ospedale A. Murri, Via dei Colli, 52

60035 - Jesi (AN)

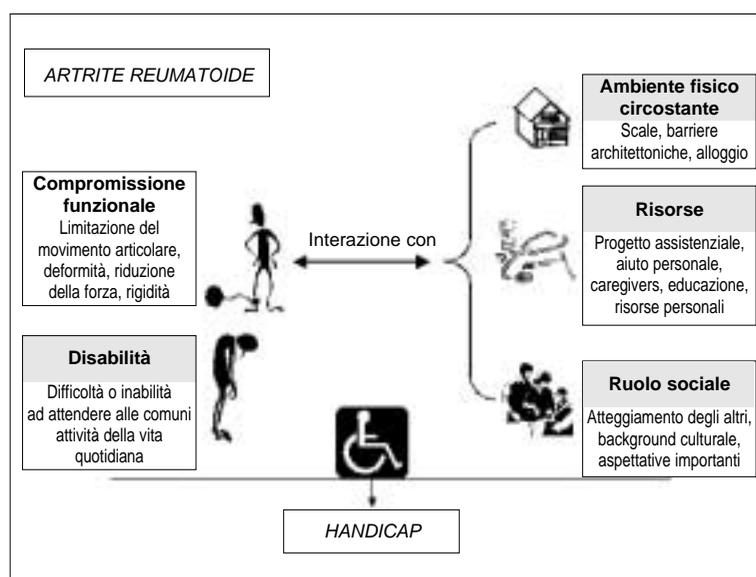
Tel. 0731/534128-32 - Fax 0731/534124

Email fsalaff@tin.it

che coinvolge un'articolazione si traduce in un'alterazione della struttura articolare che, a sua volta, può indurre un deterioramento della funzione con limitazione del movimento, riduzione della forza e determinare, pertanto, la necessità da parte del paziente di evitare, ridurre o modificare determinati tipi di attività, compiti e comportamenti. I cambiamenti nello stile di vita e le modificazioni in termini di comportamento e abitudini che avvengono nei pazienti affetti da malattie croniche, come l'AR, costituiscono ciò che alcuni autori hanno definito "*illness intrusiveness of a (chronic) disease*" (9). Fattori psico-sociali, demografici, assistenziali e altre variabili ambientali possono, in diversa misura, condizionare il grado di impatto di uno stato patologico sulla qualità della vita (QdV) del paziente (10). La disabilità, pertanto, è strettamente correlata alla malattia di base e al deterioramento funzionale conseguente, rappresentando uno stato intermedio del *continuum* che, partendo dall'alterazione anatomo-patologica, si conclude con l'esito socialmente più sfavorevole: l'handicap (8). L'handicap è la risultante dell'interazione tra il deterioramento funzionale, la disabilità e una serie di fattori esterni alla malattia che derivano al paziente dall'ambiente circostante (Fig. 1) (8). La perdita della capacità lavorativa (*work disability*), e quindi l'impossibilità a mantenere il proprio ruolo sociale, rappresenta senza dubbio la connotazione più negativa della disabilità, che si traduce in grave handicap. L'impatto di una malattia cronica sui ruo-

li sociali dipende sia dal tipo di ruolo ricoperto che dalla flessibilità, intesa come capacità di cambiamento, che caratterizza tale ruolo. Ad esempio, in alcuni tipi di attività lavorativa, dove maggiori sono le possibilità di cambiare mansione od orario, sarà più semplice per il paziente affetto da AR mantenere un'accettabile attività lavorativa. Il ruolo lavorativo, in quanto principale determinante del reddito e della posizione sociale, contribuisce in misura significativa all'identità dell'individuo. La disabilità al lavoro, intesa come l'incapacità di proseguire la propria attività lavorativa, mantenere le stesse mansioni, o lavorare lo stesso numero di ore, rappresenta uno degli "outcomes complessi" di maggiore rilevanza nei pazienti con AR. In generale, nell'AR, la prevalenza stimata di disabilità al lavoro è sette volte superiore rispetto alla popolazione generale (11). I numerosi studi pubblicati hanno dimostrato una perdita di capacità lavorativa del 50-60% a 10 anni dall'esordio di malattia (Tab. I) (11-33). Non è del tutto chiaro, tuttavia, se in tutti i casi il motivo dell'abbandono dell'attività lavorativa sia riconducibile a una diretta conseguenza della malattia. Da un'analisi dei dati disponibili in letteratura (11-32) si evince che i fattori più frequentemente associati con la perdita dell'attività lavorativa, e che quindi potrebbero avere un valore predittivo, sono essenzialmente riconducibili a tre categorie principali (Tab. I):

- le caratteristiche dell'attività lavorativa: tipo di lavoro, grado di impegno fisico richiesto, livello di autonomia, inteso come possibilità di gestirsi il proprio ruolo e i periodi di riposo;
- le caratteristiche del paziente: età all'esordio dell'AR, stato civile, appartenenza sociale, grado di scolarità;
- le caratteristiche della malattia: anzianità di malattia, livello di disabilità, attività di malattia (intensità del dolore, numero di articolazioni tumefatte, VES o PCR).



**Figura 1** - World Health Organization (WHO): struttura classificativa della compromissione funzionale, disabilità e handicap.

## 2. LA VALUTAZIONE DELLA DISABILITÀ E DELLA QUALITÀ DELLA VITA

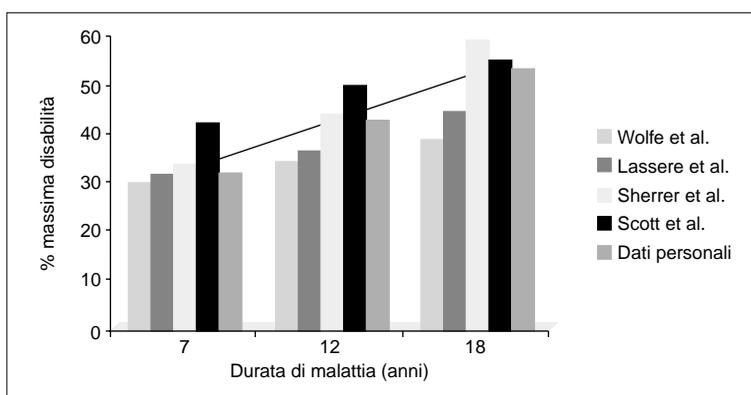
In studi precedenti al 1980, la determinazione del grado di disabilità veniva effettuata principal-

**Tabella I** – Sinossi degli studi statunitensi ed europei sulla incidenza della disabilità lavorativa (work disability) nei pazienti con AR.

STATI UNITI			EUROPA		
Autore, anno	% abbandono dell'attività lavorativa	Predittori di disabilità	Autore, anno	% abbandono dell'attività lavorativa	Predittori di disabilità
Yelin et al. 1980 (12)	62% a 10 aa.	Stadio malattia Durata malattia Attività lavorativa Stato civile	Mäkisara et al. 1982 [Finlandia] (21)	40% a 5 aa. 50% a 10 aa. 67% a 15 aa.	Grado di disabilità VES Fattore reumatoide Danno radiologico
Pincus et al. 1984 (13)	85% a 9 aa.	NR	Borg et al. 1991 [Norvegia] (22)	23% a 2 aa.	Età Attività lavorativa HAQ N° artic. tumefatte
Yelin et al. 1987 (14)	50% a 10 aa. 60% a 15 aa. 90% a 30 aa.	Attività lavorativa HAQ Impegno articolare Età	Eberhardt et al 1993 [Svezia] (23)	37% a 2 aa.	Attività lavorativa HAQ
			Doeglas et al. 1995 [Olanda]	(24) 42% a 2 aa.	HAQ Scolarità VES Durata malattia
Pincus et al. 1989 (15)	Donne: 69% Uomini: 43,9%	NR	Mau et al. 1996 [Germania] (25)	37% a 7 aa.	Età Stadio funzionale Durata malattia VES Impegno articolare
Callahan et al. 1992 (16)	43% a 5 aa.	Attività lavorativa Età Durata di malattia	Fex et al. 1998 [Svezia] (26)	37% a 8 aa.	HAQ Scolarità Età
			Salaffi et al. 1998 [Italia] (27)	41,3% a 6 aa.	HAQ Indice di Ritchie Attività lavorativa Scolarità
Reisine et al. 1995 (17)	43% a 5 aa.	Impegno articolare Stato sociale Attività lavorativa Età			
			Sokka et al. 1999 [Finlandia] (28)	19% a 2 aa. 44% a 10 aa.	Attività lavorativa Età Scolarità N° artic. tumefatte
Allaire et al. 1996 (18)	21,7% a 7 aa.	Età Bassa scolarità Durata di malattia Attività lavorativa Dolore	Albers et al. 1999 [Olanda] (29)	14% a 1 aa. 42% a 3 aa. 72% a 5 aa.	Età Sesso Stato civile Attività lavorativa
			Jäntti et al. 1999 [Finlandia] (30)	31% a 1 aa. 80% a 20 aa.	HAQ Danno radiologico
Wolfe e Hawley, 1998 (19)	23% a 5 aa. 31,5% a 10 aa. 50% a 20.9 aa.	Attività lavorativa VES HAQ Scolarità BMI N° artic. dolenti	Young et al. 1999 [Inghilterra] (31)	40% a 5 aa.	Attività lavorativa Età Stato socio-economico HAQ
			Barrett et al. 2000 [Inghilterra] (11)	14% a 1 aa. 26% a 2 aa. 33% a 5 aa. 39% a 10 aa.	HAQ Età Scolarità Attività lavorativa
De Roos et al. 1999 (20)	26,5% a 11 aa.	Età Durata di malattia Scolarità Stato civile Dolore			
			Young et al. 2000 [Inghilterra] (32)	22% a 5 aa.	Attività lavorativa HAQ
			Chorus et al. 2001 [Olanda] (33)	34% a 5 aa.	Attività di malattia Durata malattia HAQ

Attività lavorativa NR= non riportato; BMI= Body Mass Index

mente impiegando le classi funzionali di Steinbrocker (34). In particolare, veniva registrato il numero di pazienti che presentavano un grado di disabilità da moderato a severo (III e IV classe funzionale) sia nell'AR di recente insorgenza (durata di malattia <5 anni) che nelle forme inveterate (durata di malattia >15 anni). Ragan e Farrington (35) in una casistica di 409 pazienti con AR, documentarono una percentuale di pazienti con classe funzionale III/IV pari al 15% e 48% rispettivamente nei pazienti con durata di malattia inferiore a 5 e superiore a 15 anni. Nello studio di Duthie et al (36) su una casistica di 282 pazienti le percentuali documentate nei due gruppi erano rispettivamente del 25% e 38%, mentre in quella di Rasker e Cosh (37) su 100 pazienti erano rispettivamente del 5% e 33%. Pertanto, complessivamente, nelle casistiche suddette il valore medio delle percentuali di pazienti con classe funzionale III/IV può essere stimato nell'ordine del 15% nei pazienti con durata di malattia fino a 5 anni e del 40% in quelli con durata di malattia superiore a 15 anni. Negli anni successivi al 1980, il grado di di-



**Figura 2** - Progressione della disabilità funzionale, valutata mediante Health Assessment Questionnaire, in studi prospettici longitudinali.

sabilità è stato valutato mediante misure semi-quantitative, utilizzando scale e questionari orientati. Fra i numerosi indici di disabilità proposti in letteratura, quello maggiormente impiegato nell'AR è stato l'Health Assessment Questionnaire (HAQ) (38). I punteggi assoluti dell'indice di disabilità HAQ nei gruppi di pazienti aumentavano invariabilmente in rapporto alla durata di malattia. Tale evidenza è stata dimostrata in differenti modi. Quello più semplice e intuitivo è stato lo studio delle matrici di correlazione tra la durata di malattia e l'HAQ. Pincus et al (39) e Houssein et al (40)

**Tabella II** – Progressione annuale dell'indice di disabilità Health Assessment Questionnaire (HAQ) in pazienti con artrite reumatoide.

Autore	Anno	Pazienti	Disegno dello studio	Durata di malattia	Incremento annuale unità/HAQ
Sherrer et al. (43)	1986	681	Trasversale/longitudinale	Variabile	0.072
Wolfe et al. (47)	1991	561	Trasversale/longitudinale	Variabile	0.020
Leigh et al. (46)	1992	209	Prospettico a 8 anni	Variabile	0.018
Ward et al. (51)	1993	282	Prospettico a 10 anni	Variabile	0.008 – 0.020
Gardiner et al. (52)	1993	175	Prospettico a 5 anni	Variabile	0.030
Guillemin et al. (53)	1994	337	Prospettico a 2 anni	< 5 anni	0.016
Eberhardt et al. (58)	1995	63	Prospettico a 5 anni	Recente esordio	0.020
Lassere et al. (44)	1995	353	Trasversale	Variabile	0.045
Callahan et al. (54)	1997	100	Prospettico a 5 anni	Variabile	- 0.006
Leymarie et al. (55)	1997	370	Prospettico a 2 anni	< 5 anni	0
Ward et al. (56)	1998	182	Prospettico a 10 anni	Variabile	0.015 – 0.019
Fex et al. (26)	1998	106	Prospettico a 8 anni	< 2 anni	0.062
Munro et al. (57)	1998	440	Prospettico a 5 anni	Variabile	0.119
Truro (41)	1998	33	Prospettico a 5 anni	Recente esordio	0.006
Whipps (41)	1998	46	Prospettico a 4 anni	Variabile	0.023
Greenwood et al. (48)	2001	207	Prospettico a 1 anno	Variabile	0.031
Wolfe et al. (59)	2001	12.949	Prospettico a 9 anni	Variabile	0.013
Landewé et al. (60)	2002	148	Prospettico a 1 anno	< 1 anno	0.010
Salaffi et al.	2002	92	Prospettico a 5 anni	Variabile	0.030
(casistica personale)					
<b>Valore medio di progressione annuale dell'HAQ</b>					<b>0.028</b>

su una casistica rispettivamente di 200 e 259 pazienti con AR, con un ampio range di durata di malattia, hanno rilevato una correlazione di poco superiore a 0.30. In un'ulteriore metanalisi di 725 casi, Scott et al. (41) hanno trovato coefficienti di correlazione compresi fra 0.27 e 0.30. Esistono in letteratura pochi studi longitudinali a lungo termine sulla progressione della disabilità nel tempo, in quanto l'utilizzo dell'HAQ su vasta scala risale a poco più di due decenni. Diversi contributi derivanti da studi trasversali hanno dimostrato l'andamento nel tempo dei punteggi dell'HAQ. L'analisi dei dati relativi a quattro studi (41-44), condotti su un numero variabile di pazienti da 264 a 725, con durata di malattia compresa fra 1 anno e più di 25 anni, ha rilevato punteggi medi di HAQ pari a 1, 1.25 e 1.50 in pazienti con durata di malattia rispettivamente di 7, 12 e 18 anni. Tali risultati sono sostanzialmente comparabili a quelli ottenuti nella nostra casistica (Fig. 2). Analoghe conclusioni sono derivate anche dalla tecnica della regressione del quantile e delle relative curve di riferimento dei centili per i punteggi di HAQ (45). Queste curve mostrano una mediana dei punteggi di HAQ di circa 0.9 a 9 anni, con progressivi incrementi fino a 1.4 a 18 anni. Un approccio alternativo allo studio della progressione della disabilità nell'AR è rappresentato dal calcolo dell'incremento annuo dei punteggi di HAQ. Leigh et al. (46), in uno studio su 209 pazienti seguiti tra il 1981 e il 1989, hanno riscontrato un incremento annuo dell'HAQ pari a 0.018 unità. Questa metodica può essere applicata alla trasformazione dei dati provenienti sia da studi trasversali che longitudinali. L'analisi degli studi pubblicati su questo argomento (26, 41, 43, 44, 47-60) e la nostra personale esperienza inducono a prospettare un incremento medio del grado di disabilità di 0.028 unità HAQ/anno (Tab. II). Ciò significherebbe che in 25 anni l'HAQ aumenterebbe meno di una unità. È tuttavia documentabile una forte variabilità inter-individuale nei punteggi dell'indice di disabilità, specie nelle prime fasi della malattia. Eberhardt e Fex (58), in uno studio prospettico a 5 anni su 63 pazienti con AR in fase "early", hanno documentato una mediana dei punteggi di HAQ assestata, a 5 anni, su valori di 0.7 nei maschi e 1.1 nelle femmine, con un "rate di progressione" annuo complessivo pari a 0.020. Tuttavia, la variazione inter-individuale era particolarmente elevata (58). Risultati sovrapponibili sono emersi dallo studio di Wiles et al (61), su una casistica di 433 pazienti con AR, in cui il valore dell'HAQ della maggior parte dei pazienti devia

rispetto alla curva dei centili ottenuta dallo studio di Lassere (44, 45). A 5 anni, solo il 19% dei pazienti rimaneva nello stesso quartile. Dal quinto anno in poi, la percentuale aumentava al 65% dei casi (61).

In ultima analisi, i numerosi contributi della letteratura dimostrano come il "rate di progressione" della disabilità resti relativamente costante nel tempo e come tale percentuale corrisponda a un valore inferiore all'1% della massima disabilità possibile (41). Questa tendenza al graduale incremento della disabilità si sovrappone, nei singoli pazienti, a brusche e transitorie modifiche dello stato funzionale, che possono determinare variazioni medie anche superiori al 30% nell'indice di disabilità (62, 63). Ciò rende di difficile lettura il dato relativo alla progressione della disabilità nel breve termine, specie se essa viene considerata una misura di outcome funzionale. A tale scopo vengono utilizzate due metodologie di approccio che consentono, in definitiva, di trasformare un end-point parziale (come il punteggio di un indice di disabilità) in una misura di outcome: il metodo delle aree sottese alla curva (*Area Under the Curve*) e il metodo del livello mantenuto (*Sustained Level Method*) (41, 50). Poiché l'outcome rappresenta essenzialmente il risultato di tutto ciò che avviene nel corso della malattia, teoricamente il calcolo dell'area sottesa alla curva rappresentativa dell'andamento delle misure di processo nel tempo dovrebbe soddisfare la necessità di un'appropriata misura di outcome. In tal modo, la trasformazione delle misure seriali rende possibile l'utilizzo di misure "surrogate", quali gli "end points parziali" per la valutazione degli "outcomes" di

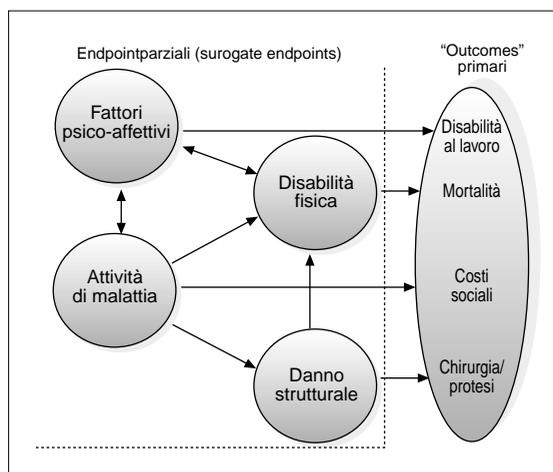
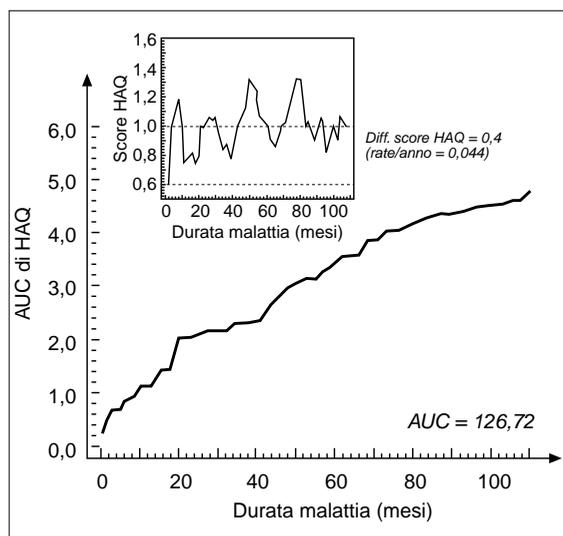


Figura 3 - "Endpoint" parziali ed "outcomes" primari identificati nell'artrite reumatoide.

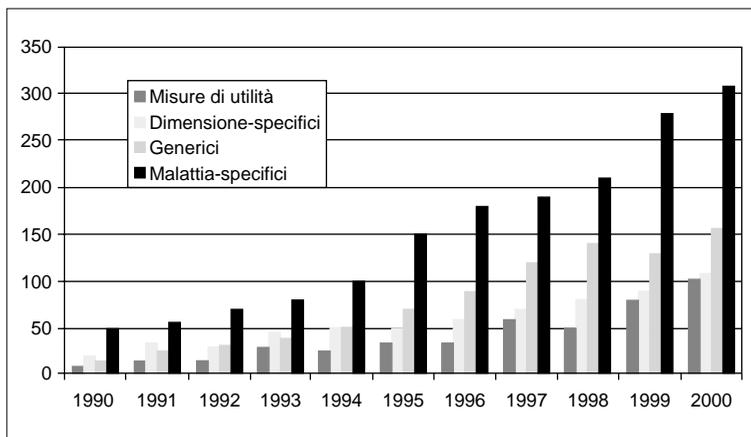


**Figura 4** - Grafico esemplificativo di trasformazione di misure seriali mediante il calcolo dell'area sotto la curva (Area under the Curve). Nel riquadro si noti l'ampia variazione (>30%) dell'indice di disabilità Health Assessment Questionnaire, nel breve periodo.

malattia (Fig. 3). Questa metodologia di analisi delle misure seriali, come proposto da Matthews (64), rappresenta un metodo efficace sia per valutare la risposta al trattamento in un trial clinico, sia per monitorare la malattia nel tempo. Nella figura 4 viene rappresentato un esempio di trasformazione di misure seriali dell'HAQ in un paziente con AR seguito in follow up di oltre 10 anni. Il metodo del livello mantenuto, meno impiegato del precedente, consiste invece nello stabilire un punteggio minimo di HAQ che debba necessariamente essere mantenuto per tutta la durata dell'osservazione: in sostanza, si stabilisce per il gruppo o il paziente in esame una soglia di disabilità (ad esempio un HAQ = 2 per almeno due anni) (41, 50).

Appare evidente, pertanto, che riuscire a individuare i livelli di inibizione funzionale determinati dalla malattia sullo stato globale di salute può risultare di aiuto nell'approntare una serie di programmi (schemi farmacologici, tecniche di economia articolare, strategie cognitive nell'affrontare i problemi ecc.), volti a un maggiore coinvolgimento dei pazienti nella gestione della propria salute. In questo senso, le linee guida internazionali tengono conto di

questi diversi aspetti e includono, fra le misure raccomandate nella valutazione del decorso a lungo termine dell'AR, anche quelle concernenti la valutazione della disabilità e della QdV (65-67). Sebbene il termine di QdV appaia di comprensione immediata, esso non è di facile definizione e determinazione. In un recente editoriale, Testa e Simonson (68), partendo dalla definizione di salute proposta dall'Organizzazione Mondiale della Sanità (OMS) (8), definiscono la QdV come gli aspetti (o domini) della salute fisico-funzionale, psicologico-emotiva e sociale che sono influenzati dalle esperienze, credenze, obiettivi e aspettative degli individui. La QdV riguarda, quindi, un concetto ad ampio spettro, che è modificabile in maniera complessa dalla percezione della propria salute fisica e psicologico-emotiva, dal livello di indipendenza, dalle relazioni sociali e dalla interazione con il proprio specifico contesto ambientale. In accordo con questa definizione, la valutazione della QdV prevede la quantificazione della componente obiettiva dello stato di salute filtrata dalla percezione del paziente. Il riconoscimento di tale capacità di autopercezione degli effetti degli interventi di assistenza sanitaria e dell'impatto della malattia cronica sullo stato di salute ha favorito lo sviluppo di scale e di questionari orientati come l'HAQ, che hanno trovato crescente impiego nei trial terapeutici. Nella figura 5 viene riportata la crescita esponenziale dei contributi scientifici su trial clinici e studi osservazionali, pubblicati nel decennio 1990-2000 in un ambito di 30 differenti specialità mediche, in cui figurano impiegate le varie scale e i numerosi strumenti di valutazione dello stato globale di salute (scale generiche, dimensione-specifici-



**Figura 5** - Tasso di incremento nell'impiego delle scale di valutazione dello stato di salute (generiche e specifiche) in ambito scientifico e negli studi clinici per il decennio 1990-2000.

che, malattia-specifiche e misure di utilità) (69). Dal confronto emerge che la reumatologia risulta la specialità con la più elevata quota di contributi scientifici (pari al 13%) riguardanti tale tipo di valutazione, seguita dall'oncologia (11%) e dalla geriatria (9%) (69).

### 3. STRUTTURA CONCETTUALE DELLE SCALE E DEI QUESTIONARI DI VALUTAZIONE

Un questionario di valutazione della disabilità o della QdV consiste in una o più serie di quesiti, ognuno dei quali esplora un componente del costrutto da analizzare o parte di esso (68, 70-73). Indici sintetici che descrivono globalmente la dimensione (fisica, psicologica, sociale ecc.) possono essere derivati dall'aggregazione dei punteggi di diverse scale e componenti del questionario. I quesiti e le scale dei questionari di valutazione sono, in genere, concepiti per l'auto-somministrazione, la somministrazione telefonica o quella condotta attraverso un colloquio faccia a faccia. Un sistema standardizzato di attribuzione

dei punteggi consente l'interpretazione del questionario e rende possibile il confronto tra i risultati di studi diversi. La possibilità di disporre di diverse traduzioni fa sì che tali questionari possano essere utilizzati in paesi, culture e lingue differenti (validità trans-culturale). A questo riguardo, le tappe di traduzione e valutazione qualitativa, tese ad assicurare l'equivalenza semantica dei quesiti in ogni paese, l'ordinalità e le proprietà interne di traduzione delle modalità di risposta, l'accettabilità e la chiarezza del linguaggio, devono essere attuate nel più rigoroso rispetto delle procedure metodologiche di validazione (70-74). Le fasi di attribuzione dei punteggi (ricodifica dei valori delle risposte date e di quelle mancanti, inversioni dei punteggi e/o ricalibrazione dei punteggi dei quesiti, calcolo dei punteggi grezzi di scala, trasformazione dei punteggi "grezzi" e costruzione di indici sintetici basati su standard di riferimento ricavati da popolazioni di controllo) vengono, di norma, eseguiti con l'ausilio di un computer, utilizzando gli algoritmi e le griglie di calcolo suggeriti dai vari manuali d'uso ai questionari o mediante programmi "software", altrimenti disponibili.

**Tabella III** – Contenuto dei principali questionari generici e specifici di valutazione della disabilità e della qualità della vita maggiormente impiegate nell'artrite reumatoide

Strumento	Dimensione Fisica		Dimensione Psico-Sociale			Attività di gruppo nel tempo libero	Salute psicologica
	Mobilità/attività fisica, sintomi	Auto-assistenza	Attività di ruolo	Comunicazione	Interazione sociale		
<b>Generici</b>							
SF-36	+	+	+	0	+	+	+
GHQ	+	+	+	0	+	+	+
SIP	+	+	+	+	+	+	+
NHP	+	+	0	0	+	0	+
MHIQ	+	+	+	0	+	+	+
QWB	+	0	0	0	+	+	+
EUROQoL	+	+	0	0	+	+	+
<b>Specifici</b>							
AIMS	+	+	+	+	+	+	+
AIMS2	+	+	+	+	+	+	+
MACTAR/PET	+	+	+	+	+	+	+
FSI	+	+	+	0	+	+	0
HAQ	+	+	+	0	0	0	0
DASH	+	+	+	0	+	0	0

+ = caratteristica esaminata; 0 = caratteristica non esaminata. Abbreviazioni: GHQ = General Health Questionnaire; (MOS) SF-36 = Medical Outcomes Study (MOS) 36-Items Short-Form Healthy Survey; SIP = Sickness Impact Profile; NHP = Nottingham Health Profile; MHIQ = McMaster Health Index Questionnaire; QWB = Quality of Well-Being; EUROQoL: European Quality of Life Questionnaire; AIMS = Arthritis Impact Measurement Scales; MACTAR/PET = McMaster-Toronto Arthritis Patient Preference Disability Questionnaire/Problem Elicitation Technique; FSI = Functional Status Index; HAQ = Health Assessment Questionnaire; DASH = Disabilities of Arm, Shoulder and Hand Questionnaire.

#### 4. QUESTIONARI GENERICI E ARTRITE-SPECIFICI

Uno strumento per la raccolta dei dati clinici nasce da due esigenze fondamentali. La prima è quella di poter disporre di una serie di parametri che permettano di caratterizzare una certa malattia nei pazienti e fra i pazienti che ne sono affetti (finalità discriminativa). La seconda è quella di poter acquisire informazioni utili per seguire il paziente in relazione all'evolversi della sua malattia (finalità valutativa). Il costrutto dei questionari e la standardizzazione del contenuto sono generalmente diversi per queste due finalità. Guyatt et al. (71,73) hanno proposto uno schema di classificazione delle scale di valutazione al fine di delineare la gamma di applicazione e il contenuto delle stesse. Le due principali categorie del sistema sono gli strumenti *generici* e gli strumenti *specifici*.

##### 4.1. Gli strumenti generici

Gli strumenti generici, in particolare i profili dello stato di salute, sono stati messi a punto allo scopo di fornire una descrizione delle condizioni sanitarie in una vasta gamma di popolazioni, in relazione alle diverse dimensioni dello stato di salute basato su molteplici attributi (70-74). Tale approccio valutativo si fonda sul principio che lo stato di salute può essere definito da un certo numero di concetti, componenti e dimensioni, possibilmente collegati in maniera gerarchica. In teoria, il sistema di attributi può essere tanto ampio e profondo, quanto è richiesto dal grado di dettaglio che si intende perseguire nello studio. Alcuni di tali strumenti forniscono un punteggio per ciascuna scala o componente o dimensione del profilo, mentre altri forniscono solo un punteggio globale, il cui impiego richiede cautela.

Come riportato nella tabella III, esistono diversi strumenti nell'ambito di tale categoria. Qui di seguito vengono riepilogate le caratteristiche principali degli strumenti più rappresentativi.

**4.1.1 I profili sanitari.** I profili dello stato di salute che più di altri hanno trovato impiego nell'AR includono: il Medical Outcomes Study (MOS) 36-Items Short-Form Healthy Survey (SF-36) (75), il General Health Questionnaire (GHQ) (76), il Sickness Im-

pact Profile (SIP) (77), il Nottingham Health Profile (NHP) (78), il McMaster Health Index Questionnaire (MHIQ) (79) e l'European Quality of Life Questionnaire (EUROQoL) (80).

L'SF-36 è un questionario sviluppato nel corso del progetto del Medical Outcomes Study (MOS) per valutare i principali "concetti" riguardanti lo stato di salute. Concepito per l'auto-somministrazione, la somministrazione telefonica o quella condotta attraverso un colloquio faccia a faccia, l'SF-36 è stato validato originariamente negli Stati Uniti (75) e successivamente tradotto e adattato in diversi contesti linguistici. I dati raccolti nel Progetto dell'International Quality of Life Assessment (IQUOLA) dimostrano che il questionario è valido e affidabile anche nella versione italiana (81). Il confronto transnazionale di campioni rappresentativi di soggetti di controllo stratificati per fasce di età, fornisce, sostanzialmente, analoghi punteggi classi-dimensionali (Fig. 6) (82). Il questionario prevede otto scale a quesito multiplo, ciascuna delle quali contiene da due a dieci quesiti e una valutazione a singolo quesito sul cambiamento delle condizioni di salute, che non viene impiegata per l'attribuzione del punteggio in nessuna delle otto scale a quesiti multipli (83, 84). Indici sintetici che descrivono globalmente lo stato di salute fisica (attività fisica, ruolo e salute fisica, salute in generale, dolore) e quello mentale (vitalità, ruolo e stato emotivo, salute mentale, attività sociali) sono stati ottenuti dall'aggregazione delle diverse sottoscale (85). Tali indici sintetici dimostrano una maggiore responsabilità nel follow-up (86-88). I quesiti e le sottoscale dell'SF-36 sono organizzati in modo tale che tanto più è elevato il punteggio tanto migliore è lo stato di salute del soggetto. Gli algoritmi relativi alle fasi di attribuzione dei punteggi (ricodifica dei valori, inversione e/o ricalibrazione dei punteggi, calcolo dei punteggi grezzi, trasformazione

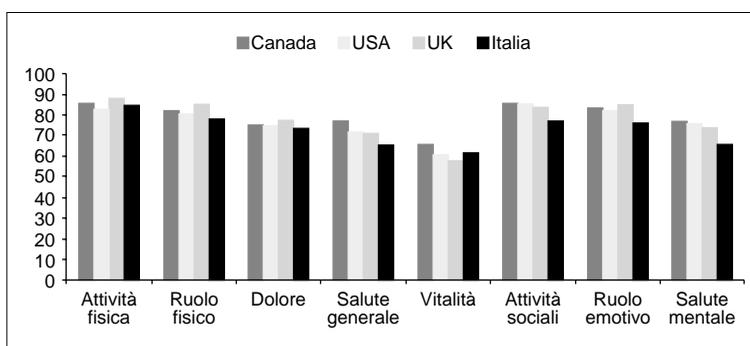


Figura 6 - Dati normativi: confronto transnazionale delle classi-dimensioni del SF-36.

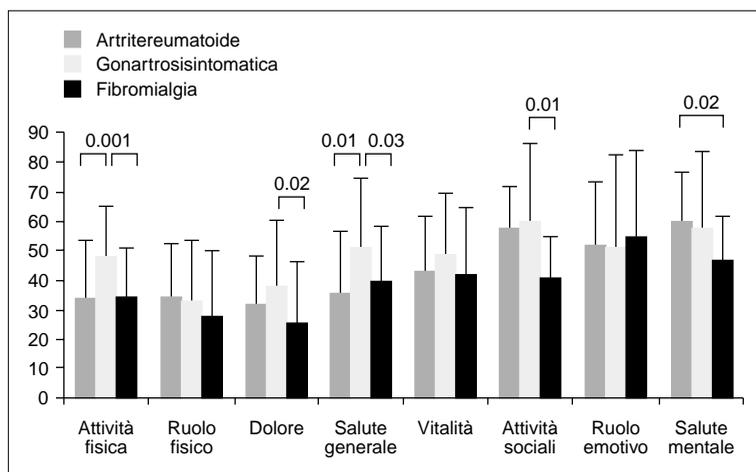


Figura 7 - Confronto fra le dimensioni dell'SF-36 nell'artrite reumatoide, gonorrosi sintomatica e fibromialgia.

in punteggi di scala) sono fissati da procedure matematiche standardizzate. Il questionario viene solitamente completato dal paziente in 5-8 minuti. Fra gli strumenti generici, l'SF-36 è quello maggiormente impiegato nella valutazione delle malattie muscolo-scheletriche. L'impiego di questa scala nei pazienti con AR dimostra, soprattutto se comparata alla gonorrosi sintomatica e alla fibromialgia, come le aree (o domini) di salute più marcatamente coinvolte siano quelle concernenti la disabilità fisica (in particolare l'attività fisica e la salute in generale) (Fig. 7).

Il GHQ è stato inizialmente proposto in una versione estesa costituita da 140 quesiti e ha subito negli anni numerose e ulteriori verifiche in vari contesti di epidemiologia psichiatrica. L'attuale e più recente versione, validata nel contesto del progetto *EURIDISS (European Research on Incapacitating Diseases and Social Support)* (76) e maggiormente utilizzata in ambito reumatologico, comprende 28 quesiti (GHQ-28), strutturati secondo una "Likert scale", attinenti a 6 principali aree dello stato globale di salute: depressione, ansietà, qualità del sonno, inadeguatezza sociale, ipocondria e sintomi somatici (76). L'intero questionario autosomministrato viene, di norma, completato in 10-15 minuti.

Il SIP, sviluppato negli USA negli anni '70, si basa su un modello concettuale di malattia intesa come manifestazione di cambiamenti comportamentali. La focalizzazione sui comportamenti anomali di malattia presenta numerosi vantaggi rispetto alla registrazione delle sensazioni: i comportamenti sono osservabili e accessibili a osservazioni ester-

ne, mentre le sensazioni sono variabili, soggettive e difficilmente misurabili. Concepito all'origine come un indice dell'efficienza funzionale del singolo paziente, è stato successivamente utilizzato come mezzo di valutazione dello stato globale di salute e dei risultati conseguiti nei programmi terapeutici, anche in ambito reumatologico. La versione finale del questionario comprende 136 quesiti, attinenti a disfunzioni correlate alla malattia relative a 12 domini (mobilità, deambulazione, igiene personale e movimento, attività domestiche, capacità lavorativa, interazioni sociali, comunicazione, comportamento in sta-

to di vigilanza, sonno e riposo, mangiare, attività ricreative e divertimento) (77). Il punteggio delle scale è compreso fra 0 e 100. Studi di validazione hanno suggerito la possibilità di aggregare tali concetti di salute in due principali indici sintetici comprendenti la funzione fisica (cammino, mobilità, cura di sé) e lo stato psico-sociale (interazioni sociali, vigilanza, comportamento emotivo) (77, 89). L'intero questionario, autosomministrato o proposto per intervista, viene di norma, completato in 20-30 minuti. Più recentemente è stata proposta una versione ridotta del questionario (SIP68), che ha mostrato analoghe caratteristiche psicometriche rispetto alla versione originale.

Il NHP è un questionario autosomministrato sviluppato con la finalità di indagare sulla domanda di prestazioni terapeutiche in ambito di medicina generale. L'attuale versione del NHP è suddivisa in due parti: la prima comprende 38 quesiti attinenti a 6 differenti dimensioni funzionali: mobilità fisica (8 quesiti), dolore (8 quesiti), qualità del sonno (5 quesiti), reazioni emotive (9 quesiti), isolamento sociale (5 quesiti) e livello di energia (3 quesiti); la seconda parte è strutturata in ulteriori 7 item concernenti le ripercussioni sul lavoro, la vita sociale, i rapporti familiari, le attività ricreative ecc. in seguito ai cambiamenti dello stato di salute (78). Le risposte sono di tipo dicotomico, i punteggi di ogni scala sono compresi fra 0 (nessun problema) e 100 (presenza di tutti i problemi). Nell'analisi viene calcolato un punteggio finale compreso tra 0 e 100, che considera la gravità, oltre che la semplice presenza/assenza del fenomeno. Un punteggio più elevato indica una maggiore compromissione del pa-

ziente. Il questionario è stato impiegato anche in ambito reumatologico, dimostrando una buona affidabilità e validità (40, 90). La sua compilazione richiede 10-15 minuti.

Il MHIQ è stato concepito principalmente per l'autovalutazione dello stato di salute globale di pazienti ambulatoriali e di coloro che vivono in comunità. Esso comprende 59 item concernenti il benessere fisico (mobilità, cura personale, altre attività fisiche), sociale (benessere generale, svolgimento del proprio ruolo, partecipazione alla vita familiare, relazioni con gli amici) ed emotivo (autostima, pensieri sul futuro, avvenimenti della vita) (79). Tutti i quesiti sono riferiti al presente e riguardano non tanto la capacità di effettuare qualcosa, quanto il reale svolgimento di tale attività. I punteggi delle scale sono stati standardizzati in "valori indice" compresi fra 0 (funzione estremamente scadente) e 4 (funzione estremamente buona). Il questionario richiede 20-30 minuti per la compilazione.

L'EUROQoL è un questionario autosomministrato, proposto nel contesto del progetto dell'*European Quality of Life Group* (80) e impiegato in ambito reumatologico (91, 92). Si tratta di uno strumento di facile e rapida compilazione (7-10 minuti), ma di non altrettanto agevole computazione. Esso è costituito di due parti distinte. La prima esplora 5 aree di interesse: mobilità, igiene personale, attività sociali, dolore e ansia/depressione. Ogni singola area prevede, a sua volta, tre livelli di gravità (nessun problema, problema di qualche entità, problema di estrema gravità): la combinazione dei vari livelli di gravità per le 5 aree consente, in tal modo, di classificare 243 differenti stati di salute (80). La seconda sezione è costituita da una scala analogica visiva (VAS) di 20 cm sulla quale il paziente indica il migliore (punteggio = 0) o il peggiore (punteggio = 100) stato di salute possibile. La possibilità di computare l'intero questionario anche su di una scala compresa fra valori di 0 e 1 ha indotto diversi autori a inserire l'EUROQoL fra le scale di misura delle utilità (92, 93).

**4.1.2 Le scale di misura dell'utilità.** Impiegate in ambito sanitario per la valutazione dell'impatto della malattia sullo stato globale di salute, sono derivate dalle concettualizzazioni teoriche della scienza economica e decisionale. Tali scale consentono una stima quantitativa del valore o della preferenza che i pazienti attribuiscono al loro stato di salute, se paragonato a uno stato di benessere completo (punteggio 1) o alla morte (punteggio 0) (93,

94). Esistono fondamentalmente due sistemi di approccio per la valutazione delle utilità. Il primo è quello di classificare i pazienti in categorie basate sulle loro stesse risposte a una serie di quesiti riguardanti l'efficienza funzionale; il secondo approccio, basato sull'utilità associata allo stato di salute, è quello di chiedere ai pazienti di esprimere un'unica valutazione concernente gli aspetti della qualità della loro vita. Il Quality of Well-Being (QWB), noto anche come Index of Well-Being (IWB), e l'Health Status Index (HSI) utilizzano il primo tipo di approccio (93-95). L'indice QWB, in particolare, è suddiviso in quattro sezioni: complesso sintomi-problemi, mobilità, funzione fisica e ruolo sociale. Le risposte ottenute in ciascuna sezione generano un punteggio la cui somma fornisce un punteggio globale di benessere. L'intervallo è compreso tra 0 (morte) e 1 (stato di perfetta salute) (95). A tali scale si ricorre per calcolare, ad esempio, la sopravvivenza in anni, opportunamente corretta in funzione della QdV (*Quality-adjusted life years - QALYs*), che può essere assicurata al paziente da programmi sanitari o interventi terapeutici mirati (93, 94). Nella ponderazione dei QALYs sono stati utilizzati diversi altri metodi (93-97). Questi metodi, che costituiscono il secondo tipo di approccio alla valutazione dell'utilità, comprendono principalmente le scale di valutazione (*rating scales*), il metodo del rischio standard o metodo delle scommesse (*standard gamble*) e il metodo delle alternative temporali (*time trade-off*) (96, 97). Una tipica *rating scale* è costituita da una linea orizzontale o verticale tracciata con due punti estremi di ancoraggio chiaramente definiti. Lo stato di salute maggiormente preferito è posto a un estremo della scala, mentre quello meno preferito è posto all'altro estremo (0 = peggiore stato di salute immaginabile; 1 = migliore stato di salute immaginabile). Gli intervalli esistenti tra le varie posizioni corrispondono alle differenze di preferenza, come percepite dal paziente. Una possibile variante alla classica *rating scale* include l'impiego di aiuti visivi, come quello del "termometro" graduato da 0 a 100 (96, 97) che riporta l'indicazione degli stati di salute. Il metodo dello "standard gamble" o del rischio standard è il metodo classico delle preferenze cardinali. Esso prevede che al soggetto vengano offerte due alternative. La prima alternativa consiste in un trattamento con due possibili risultati: o il soggetto ritorna al suo normale stato di salute e vive per altri  $t$  anni (con probabilità  $p$ ), oppure il paziente muore immediatamente (probabilità  $1-p$ ). La seconda alternativa presenta come

unico risultato il perdurare dello stato da cui è affetto il paziente per il resto della sua vita ( $t$  anni). Alla probabilità di morire con la prima alternativa terapeutica vengono attribuiti valori diversi fino al punto in cui l'intervistato è indifferente fra le due alternative (96, 97). La tecnica "time trade-off" è più semplice dello "standard gamble", pur trattandosi di un confronto abbinato in cui il soggetto deve scegliere tra due alternative terapeutiche: continuare a vivere un certo numero di anni ( $t$ ) con la malattia da cui è affetto oppure una condizione di buona salute per un tempo  $x$  inferiore a  $t$  seguita dalla morte. Il valore di  $x$  al quale l'intervistato risulta indifferente tra le due alternative viene utilizzato per calcolare il valore di preferenza relativo a quello stato (96, 97). Sebbene, in generale, le scale di valutazione delle utilità trovino potenziale applicazione in reumatologia per acquisire informazioni circa il rischio accettabile di un trattamento farmacologico o chirurgico, tali metodi non sono né di semplice utilizzo né di facile interpretazione e il loro impiego richiede la consultazione di un economista esperto in questa metodologia.

#### 4.2 Gli strumenti specifici

Gli strumenti specifici sono stati concepiti in modo tale da fornire informazioni più attinenti all'area di interesse (71-73, 98). Le valutazioni specifiche in campo reumatologico, e in particolare nell'AR, sono incentrate su aspetti della salute tipici e sul quadro sintomatologico (mobilità, destrezza, attività fisica, ruolo sociale e familiare, dolore, attività della vita quotidiana, dolore e stato psico-emotivo). Gli strumenti specifici più utilizzati a questo proposito sono rappresentati da: l'Arthritis Impact Measurement Scales (AIMS) (99) e la sua versione recentemente rivisitata (AIMS2) (100), il McMaster Toronto Arthritis Patient Preference Questionnaire/Problem Elicitation Techniques (MACTAR/PET) (101), il Functional Status Index (FSI) (102) e l'Health Assessment Questionnaire (HAQ) (38). Tali strumenti, testati in studi internazionali su popolazioni miste di pazienti e in studi comparativi, si sono dimostrati di rilevante utilità nella valutazione e nel monitoraggio dello stato di salute del paziente con AR in un'ampia gamma di interventi sanitari. Per quanto attiene ad altre scale specifiche, proposte per la valutazione della compromissione funzionale e dell'handicap a carico di singoli distretti articolari, quali il ginocchio (scala di Lysholm, scala Cincinnati o Knee Injury and Osteoarthritis Outcome Score), l'anca (Harris Hip Score, Hip score della Mayo Clinic o scala di Mer-

le D'Aubigné) o il rachide (Roland-Morris Disability Questionnaire, Oswestry Disability Questionnaire, Waddell Disability Index, Dallas Pain Questionnaire, Neck Pain Questionnaire, Aberdeen Back Pain Scale) (70), per quanto validate in letteratura, sono state raramente impiegate nell'AR. Fra questi strumenti, meritano, comunque, menzione le scale di valutazione dell'arto superiore, in particolare della spalla e del gomito (Disabilities of Arm, Shoulder and Hand Questionnaire) (103, 104) e della mano (indice algofunzionale di Dreiser) (105-107) per la loro potenziale utilità in ambito valutativo.

L'AIMS è un questionario autosomministrato, composto da 46 quesiti distribuiti in 9 sottoscale (99). Lo strumento è finalizzato a esplorare le seguenti componenti principali dello stato di salute: mobilità (4 quesiti), attività fisica (5 quesiti), destrezza (5 quesiti), ruolo sociale (7 quesiti), attività sociale (5 quesiti), attività quotidiane (4 quesiti), dolore (4 quesiti), depressione (6 quesiti), ansia (6 quesiti). I quesiti sono disposti nel questionario con scalarità decrescente, cosicché il punteggio più elevato è quello che stabilisce anche il punteggio della scala. La versione italiana dell'AIMS (108) mostra, analogamente alla versione originale (99), buone caratteristiche di validità e di affidabilità. Recentemente, Meenan e Mason (100) hanno proposto e validato una versione rivisitata dell'AIMS, l'AIMS2. Rispetto alla versione precedente, sono stati ricalibrati e riformulati alcuni quesiti in ordine alle caratteristiche di scalarità, è stato ampliato il loro numero (78 quesiti) e sono state previste tre ulteriori dimensioni concernenti il lavoro, la funzionalità delle braccia e il supporto della famiglia e degli amici. Inoltre, sono state introdotte tre nuove sezioni (soddisfazione in ciascuna area della propria salute, impatto globale dell'artrite in ogni area, aree di speranza di miglioramento) di rilevante interesse in chiave valutativa. La soddisfazione determinata dalla capacità di eseguire determinati compiti e l'importanza percepita dalle proprie capacità (priorità funzionale del paziente) rappresentano due fattori determinati nel processo di autovalutazione. Il questionario prevede, di norma, 20 minuti per la compilazione. La validità e l'affidabilità dell'AIMS2 hanno trovato conferme in ambito reumatologico (109) ed è attualmente disponibile la versione italiana validata (110).

Il MACTAR è un questionario specifico di valutazione, costituito da scale a quesito multiplo mediante le quali il paziente classifica, per ordine di importanza, le proprie attività funzionali in aree

diverse (cura personale, mobilità, attività di ruolo, attività ricreative, attività sociali, stato emotivo, comunicazione e qualità del sonno) e da una valutazione a singolo quesito sulle modificazioni dello stato globale di salute. Di recente è stata proposta una variante del MACTAR, il Problem Elicitation Technique (PET) (101), che, analogamente alla versione rivisitata AIMS2 (100), enfatizza il concetto di “priorità funzionali del paziente”. Il grado di difficoltà percepito nell’ eseguire determinati compiti, la severità o la frequenza di ogni singolo problema indicato vengono, infatti, valutati dal paziente mediante l’ utilizzo di scale semiquantitative a differenti livelli di gravità (da 0 = nessuna difficoltà, a 7 = attività impossibile). Il punteggio pesato della PET è determinato sommando i punteggi relativi ai 5 problemi più significativi. Il questionario può essere utilizzato solo da un intervistatore esperto, appositamente addestrato, e l’ intervista necessita di 15-20 minuti. In pazienti con artrite reumatoide indagini comparative con misure cliniche tradizionali e altri strumenti di valutazione della QdV hanno dimostrato che il MACTAR/PET è dotato di buona capacità discriminante e di discreta sensibilità (111).

Il FSI esplora il grado di dipendenza, di dolore e di difficoltà ad adempiere alle comuni attività quotidiane. La versione originale del FSI è costituita da 45 quesiti e necessita di 60-90 minuti per la

compilazione (102). La revisione del questionario ha conseguito la semplificazione del costrutto, migliorandone le caratteristiche psicometriche, oltre che la praticità d’ uso. Tale versione, ridotta a 18 quesiti, raggruppati in 5 aree di interesse (mobilità, funzione della mano, cura personale, attività domestiche e attività/ruolo sociale), richiede 20-30 minuti per la compilazione e necessita di un intervistatore. In studi comparativi il FSI ha confermato le caratteristiche di validità, riproducibilità e responsabilità.

L’ HAQ comprende 20 quesiti riguardanti altrettanti atti della vita quotidiana, suddivisi in 8 differenti categorie: lavarsi e vestirsi, alzarsi, mangiare, camminare, igiene, raggiungere oggetti, afferrare, e altre attività (38). Ogni quesito consente 4 risposte, in relazione al grado di difficoltà che comporta l’ azione richiesta: 0 = senza difficoltà; 1 = con qualche difficoltà; 2 = con molta difficoltà; 3 = non possibile. Per ogni categoria viene considerato il punteggio più alto; la somma dei punteggi (da 0 a 24) diviso 8 rappresenta il punteggio finale dell’ HAQ, che può variare, quindi, da un minimo di 0 a un massimo di 3 (38). Il questionario viene, di solito, completato in 5-8 minuti, senza alcuna difficoltà di comprensione. Il questionario è validato anche nella versione italiana (112,113). Nel tentativo di semplificare ulteriormente tale scala sono state proposte due versioni ridotte a soli 8

**Tabella IV** – Set di quesiti (items) costituenti il questionario HAQ, e le sue modificazioni (MHAQ, RA-HAQ).

MHAQ	RA-HAQ	HAQ	Quesito	Sottoscala
*	*	*	Vestirsi da solo, allacciarsi le scarpe e abbottonarsi gli abiti?	Lavarsi e vestirsi
		*	Lavarsi i capelli?	
*	*	*	Alzarsi da una sedia senza braccioli?	Alzarsi
		*	Salire e scendere dal letto?	
		*	Tagliare la carne?	Mangiare
*	*	*	Portare alla bocca un bicchiere o una tazza piena?	
		*	Aprire una confezione di latte?	
*		*	Passeggiare su un terreno in piano?	Camminare
	*	*	Salire cinque gradini?	
*	*	*	Lavare ed asciugare tutto il corpo?	Igiene
		*	Farsi un bagno nella vasca?	
		*	Utilizzare i servizi igienici da solo?	
		*	Raggiungere e prendere un oggetto del peso di circa due chili (per esempio un sacchetto di zucchero posto sopra la Sua testa)?	Raggiungere oggetti
*	*	*	Chinarsi per raccogliere un indumento caduto a terra?	Afferrare
	*	*	Aprire la portiera della macchina?	
*		*	Aprire un barattolo già aperto in precedenza?	
		*	Aprire e chiudere un rubinetto?	Altre attività
	*	*	Fare delle commissioni e fare spesa?	
*		*	Salire e scendere dalla macchina?	
		*	Svolgere lavori quali passare l’ aspirapolvere o pulire un cortile?	

quesiti: l'HAQ modificato da Pincus (MHAQ) (114) e l'HAQ derivato da una complessa procedura di analisi statistica (Rasch analyses) in un ampio campione di pazienti con AR (RA-HAQ) (115). In tabella IV vengono riportati gli item relativi a ognuna delle versioni del questionario. L'HAQ è lo strumento maggiormente utilizzato in ambito reumatologico e viene raccomandato dall'American College of Rheumatology (ACR) (116), dall'European League Against Rheumatism (EULAR) (66) e dalle recenti conferenze internazionali "Outcome Measures in Rheumatoid Arthritis Clinical Trials" (OMERACT) (67) fra le misure di valutazione di attività e di outcome dell'artrite reumatoide. Sebbene nell'AR l'HAQ si sia dimostrato affidabile, valido e in grado di rispondere ai mutamenti dello stato di malattia nel tempo, confermando il suo valore predittivo in termini di disabilità al lavoro, del grado di utilizzo dei presidi sanitari, del rischio di gastrotossicità del trattamento antinfiammatorio (117) e dell'incidenza di mortalità (1, 13, 39, 118, 119), esso ha varie limitazioni (120). In primo luogo esplora elettivamente la disabilità funzionale e, pertanto, la sua applicazione dovrebbe essere integrata da altri strumenti multidimensionali, quali i questionari generici (ad esempio l'SF-36) o le scale di valutazione dell'impatto psico-sociale (ad esempio le scale della depressione, dell'ansia e delle attività sociali dell'AIMS2). In secondo luogo, i quesiti del questionario sono eccessivamente sbilanciati sul versante della valutazione dell'impegno delle grosse articolazioni (anche, ginocchia e spalle) e sono relativamente poco sensibili a monitorare, per esempio, la funzione della mano.

Ciò rende ragione delle evidenti discrepanze, rilevate in letteratura, nei dati concernenti l'interrelazione fra disabilità fisica valutata con l'HAQ (che enfatizza l'impegno delle grandi articolazioni) e la progressione del danno radiologico (valutato a carico di mani e polsi con metodi di scoring quantitativo) e della necessità di affiancare, per tali finalità, scale specifiche di valutazione della compromissione funzionale di singoli distretti articolari (es. il DASH o l'indice di Dreiser).

Il DASH (103, 104) è un questionario di valutazione della disabilità dell'arto superiore, costituito da 30 quesiti, finalizzati a esplorare la funzione fisica (21 quesiti attinenti la capacità di scrivere, di girare una chiave nella serratura ecc.), i sintomi (6 quesiti orientati in particolare all'autovalutazione del dolore), il ruolo e la funzione sociale (3 quesiti), nel corso dell'ultima settimana. Sono previste

anche quattro voci aggiuntive rispettivamente per sportivi o musicisti, e per i lavoratori manuali. Si tratta di uno strumento auto-somministrato, in cui ogni singolo quesito consente 5 risposte (Likert scale), in relazione al grado di difficoltà che comporta l'azione richiesta (0 = senza difficoltà; 5 = estrema difficoltà). Il punteggio finale del DASH, calcolato dopo la normalizzazione dei punteggi di scala, è compreso fra un minimo di 0 e un massimo di 100. Il test viene, di solito, completato in 10-15 minuti. La validità del DASH è stata recentemente comprovata in pazienti affetti da artrite psoriasica (121).

L'Indice algofunzionale di Dreiser è il test più utilizzato per la valutazione della compromissione funzionale della mano (105-107). Lo strumento analizza la capacità di eseguire 10 gesti di vita quotidiana mediante una scala ordinale a quattro livelli (da 0 = possibile senza difficoltà a 3 = impossibile). Il punteggio totale è ottenuto dalla somma delle singole voci ed è compreso fra 0 e 30. La scala è stata validata nel 1995 su pazienti con osteoartrosi della mano e di recente ne è stata confermata la responsività nella stessa patologia (106, 107). È attualmente in corso la validazione multicentrica italiana del questionario. Altri indici meno diffusi per la valutazione della compromissione funzionale della mano sono:

- a) la scala di disabilità funzionale di Duruoz, sviluppata e validata nell'artrite reumatoide (ma non nella osteoartrosi delle dita), che prevede 18 domande poste da un esaminatore con 6 livelli ordinali di risposta, relative alle difficoltà funzionali provate durante il mese precedente (122);
- b) il test funzionale per la mano proposto da Backman, piuttosto lungo di esecuzione e che richiede materiale specifico (123);
- c) l'indice AUSCAN, un questionario autosomministrato composto da 15 voci (5 sul dolore, 1 sulla rigidità e 9 sulla funzione) (124,125).

## 5. CARATTERISTICHE PSICOMETRICHE DEGLI STRUMENTI DI VALUTAZIONE DELLA DISABILITÀ E DELLA QUALITÀ DELLA VITA

Una revisione critica della letteratura in questo settore fa emergere la pleora di sistemi di valutazione a disposizione e, al tempo stesso, la scarsa uniformità dei dati raccolti con le diverse scale, relativamente alle variabili valutate e ai differenti algoritmi utilizzati.

Da ciò deriva l'impossibilità (o l'estrema difficoltà) nella comparazione dei risultati. In una materia complessa e ancora mal definita, qual è la valutazione della disabilità e della QdV, viene pertanto richiesto il rispetto delle procedure metodologiche di validazione degli strumenti impiegati (72-74). I percorsi di verifica delle caratteristiche biometriche e psicometriche, quali l'affidabilità, la validità e la sensibilità, devono essere attuati con procedure corrette, tenendo conto delle indicazioni standardizzate della scienza quantitativa (72-74, 126).

Per validità (*validity*) di uno strumento si deve intendere la capacità di misurare l'entità per il quale l'indice viene proposto. Essa viene tradizionalmente distinta in validità di contenuto (*content validity*), di criterio (*criterion-related validity*) e di costrutto (*construct validity*). La validità di contenuto (o comprensività) verifica quanto gli items compresi nel test sono adeguati a misurare l'obiettivo per il quale la scala è stata costruita. La validità di criterio si riferisce al grado con cui una misurazione si correla con altri parametri di riferimento.

In assenza di misure considerate come un "gold standard", la validità di criterio di una scala può essere stabilita sulla base della credibilità scientifica, verificabile attraverso "focus groups" con esperti e "panel" di pazienti. La validità di costrutto del questionario è la capacità di porre in rilievo tutte le componenti coinvolte nel concetto teorico da misurare. La corretta validazione del costrutto di un questionario dovrebbe, a tal proposito, prevedere la verifica, sia della capacità di rilevare le differenze clinicamente significative tra i pazienti (validità discriminante), che il grado di correlazione esistente tra le dimensioni dello strumento e altre misure cliniche (validità convergente) (72-74, 126).

Il secondo requisito che una scala di misura deve soddisfare è rappresentato dall'affidabilità (*reliability*). Uno strumento viene definito tanto più affidabile quanto meno la misura è sensibile all'errore casuale. L'errore casuale è associato sia alla definizione di ripetibilità, ovvero di concordanza tra misure ripetute sullo stesso campione, che di riproducibilità, ovvero di concordanza tra misure ottenute con lo stesso metodo, ma in condizioni diverse. L'affidabilità viene espressa in termini di consistenza interna [grado di inter-correlazione fra i quesiti costituenti le singole sottoscale del questionario (*item-internal consistency*) e fra i singoli quesiti e le diverse sottoscale (*item-discriminant validity*)], di riproducibilità inter-esaminatore (grado al

quale i punteggi ottenuti da un esaminatore sono sovrapponibili a quelli registrati da un secondo esaminatore), di riproducibilità intra-esaminatore (grado al quale i punteggi ottenuti da un esaminatore in una determinata circostanza sono sovrapponibili con quelli ottenuti quando lo stesso esaminatore riapplica la stessa misura in un tempo successivo) e di riproducibilità test-retest (grado di variabilità dei punteggi a seguito di misurazioni ripetute in condizioni analoghe). La consistenza interna viene, di norma, calcolata mediante il coefficiente "alpha di Cronbach" (che può variare da un minimo di 0 a un massimo di 1: valori maggiori di 0.80 sono considerati significativi) o la item-to-total correlation (correlazione tra il valore di ciascuna voce e la somma di tutte le altre della scala). Per quantificare la correlazione fra misure costituite da variabili continue o razionali, distribuite normalmente, viene comunemente impiegato il coefficiente di correlazione  $r$  di Pearson; se le variabili non mostrano una distribuzione "normale", si preferisce ricorrere a un test non parametrico quale il coefficiente di correlazione per ranghi di Spearman.

La riproducibilità per dati ordinali viene comunemente quantificata dal coefficiente kappa ( $k$ ) (che può essere pesato in base al grado di discordanza rilevato,  $k_w$ ): il suo massimo valore è 1, quando l'accordo fra misure ripetute è perfetto. Un valore pari a 0 significa non accordo al di là della casualità. Anche se le dimensioni di kappa dipendono da altri fattori oltre la riproducibilità, un kappa inferiore a 0,40 è considerato basso (scarsa riproducibilità), superiore a 0,80 è considerato molto buono. Un metodo alternativo di valutare la riproducibilità di punteggi su scale ordinali (a patto che essi si avvicinino agli assunti di "normalità" nella distribuzione) è la stima del coefficiente di correlazione intra-classi (ICC), mediante il calcolo della varianza. Anche in questo caso il range dei valori è compreso fra 0 (non riproducibilità) e 1 (varianza dovuta interamente alla variabilità fra soggetti) (72-74).

Infine, per uno strumento con finalità valutative, designato cioè a monitorare nel paziente le modifiche dello stato di salute è fondamentale la verifica della *sensibilità al cambiamento* (nota anche come responsività), che rappresenta la proprietà psicometrica più rilevante di uno strumento di misura, anche per le dirette implicazioni inerenti la determinazione della numerosità campionaria (errore di tipo II) (128-130). Operativamente, la responsività (*responsiveness*) di un questionario o di una dimensione da esso esplorata, può essere defi-

nita come la probabilità di individuare un cambiamento clinicamente rilevante nelle condizioni di salute del paziente o per la risposta al trattamento o per effetto indesiderato della terapia. Sebbene non vi sia unanime accordo circa l'approccio ottimale alla misurazione della responsività, il calcolo dello *Standardized Response Mean* (SRM), la stima dell'*Effect Size* (ES) e l'applicazione del metodo delle Curve Operative Caratteristiche (*Receiver Operating Characteristic-ROC curves*) raccolgono i maggiori consensi (129-130). I primi due metodi sarebbero applicabili solo nel caso in cui le condizioni di intervallarietà delle misure siano verificate. Elevati valori di SRM e ES indicano una maggiore sensibilità nel cogliere le modifiche dello stato di salute. In particolare, valori di 0.2, 0.5, e 0.8 o superiori corrispondono, rispettivamente, a una bassa, moderata o elevata responsività (131). Le curve ROC sono la rappresentazione grafica delle caratteristiche di un test, intese in termini di sensibilità (% di veri positivi rispetto al totale dei positivi) e di specificità (% di veri negativi rispetto al totale dei negativi) (132). Con le curve ROC si suddivide il campo di variabili di un test in piccoli intervalli nei quali si analizza la sensibilità e la specificità del test in esame; tali dati vengono poi riportati rispettivamente sull'asse delle ordinate e delle ascisse in un sistema di assi cartesiani. I punti così ottenuti, per ogni intervallo di frequenza, vengono congiunti da un'unica linea spezzata; si costruisce in tal modo la ROC del test. Il test la cui ROC sottende l'area maggiore viene considerato come il più affidabile per discriminare i pazienti le cui condizioni cliniche risultano migliorate da quelli in cui invece non lo sono. Se l'area sottesa alla curva ROC non supera il valore della bisettrice degli assi cartesiani (cioè il 50% dell'area massima ottenibile, e cioè 1) si può dedurre che il test non è applicabile per le finalità per le quali è stato utilizzato. Un'area sottesa compresa tra 0.50 e 0.70 è indice di "scarsa" accuratezza del test, valori compresi tra 0.70 e 0.90 comprovano invece l'utilità del test, per un suo utilizzo nell'impiego clinico, mentre valori superiori a 0.90 esprimono una elevata ac-

curatezza (133). Il confronto fra i valori medi delle aree sottese alle curve ROC viene solitamente effettuato in accordo al metodo proposto da Hauley e McNeil (134).

## 6. CONCLUSIONI

Fra i numerosi elementi che rendono complesso il problema della valutazione dello stato di salute del paziente con AR, quello principale è rappresentato dall'elevato grado di variabilità inter-individuale delle diverse espressioni di malattia. In questo senso, la misura della disabilità introduce una moderna prospettiva nell'approccio del medico alla gestione del paziente, in quanto i benefici e i rischi possono essere valutati non solo sulla base delle osservazioni e delle percezioni del medico, ma anche e soprattutto sulla base dei valori personali del singolo paziente, che in tale modo può avere un ruolo assai significativo nell'orientare le scelte terapeutiche nei propri confronti. Tale aspetto riafferma il principio secondo il quale una valutazione non è soltanto un fatto tecnico ma anche, e inevitabilmente, una presa di posizione socio-culturale. I questionari di autovalutazione della disabilità e della QdV rappresentano un modo pratico e oggi ben validato sul piano internazionale per acquisire dati clinici di buona qualità, con minimo dispendio di tempo da parte del medico e con scarso impiego di risorse economiche. La loro importanza non si esaurisce, tuttavia, nel fatto di semplificare le procedure di raccolta delle informazioni cliniche o di eliminare alcuni difetti relativi alla parzialità e alla variabilità con cui il medico formula le domande e registra le risposte. Tali strumenti hanno rivelato un elevato grado di affidabilità e validità in merito all'identificazione dei bisogni dei pazienti, nello stabilire le priorità degli interventi terapeutici e riabilitativi e nella pianificazione strategica degli interventi sanitari, e sono da ritenersi parte integrante nel necessario processo di miglioramento dello stile di lavoro assistenziale e di ricerca nel campo delle malattie reumatiche.

### RIASSUNTO

La disabilità e la qualità della vita (QdV) nell'artrite reumatoide (AR) sono considerati elementi di fondamentale importanza in chiave prognostica, che condizionano la richiesta assistenziale da parte del paziente e ne determinano la compliance e il grado di soddisfazione nei confronti del trattamento. Il riconoscimento della capacità di auto-percezione degli effetti degli interventi di assistenza sanitaria e dell'impatto della malattia cronica sullo stato di salute ha favorito, negli anni più recenti, lo sviluppo di scale e di questionari orientati per la valutazione della disabilità e della

qualità della vita, che hanno trovato crescente impiego in ambito reumatologico. Esistono, a tale proposito, numerosi strumenti, generici e specifici, validati nei pazienti con AR. Gli strumenti generici sono stati messi a punto allo scopo di fornire una descrizione delle condizioni sanitarie, in relazione alle diverse dimensioni e forniscono un quadro generale dello stato di salute in un gruppo eterogeneo di patologie, mentre gli strumenti specifici sono stati concepiti in modo tale da fornire informazioni più attinenti all'area di interesse, nella patologia in esame, e sono pertanto più sensibili nel valutare i cambiamenti clinicamente significativi. Come in altri campi della medicina, tali strumenti focalizzano differenti aspetti: a) i segni clinici e i sintomi (biologici e psicologici); b) la funzione fisica e/o cognitiva; c) lo stato di benessere e la funzione emotiva; d) la funzione sociale; e) il grado di soddisfazione nei confronti della terapia; f) la qualità della vita in rapporto alla malattia. Gli strumenti di valutazione della disabilità e della QdV rappresentano un modo pratico e oggi ben validato sul piano internazionale per acquisire dati clinici di buona qualità, con minimo dispendio di tempo da parte del medico e con scarso impiego di risorse economiche. Tali strumenti hanno rivelato un elevato grado di affidabilità e validità in merito all'identificazione dei bisogni dei pazienti, nello stabilire le priorità degli interventi terapeutici e riabilitativi e nella pianificazione strategica degli interventi sanitari, e sono da ritenersi parte integrante nel necessario processo di miglioramento dello stile di lavoro assistenziale e di ricerca nel campo delle malattie reumatiche. La presente rassegna offre una disamina dei principali strumenti di valutazione della compromissione fisica, della disabilità e della qualità della vita nei pazienti con AR.

**Parole chiave** - Artrite reumatoide, disabilità, qualità della vita, outcome, strumenti generici e specifici.

**Key words** - *Rheumatoid arthritis, disability, quality of life, outcome, generic and disease-specific instruments.*

## BIBLIOGRAFIA

1. Wong JB, Ramey DR, Singh G. Long-term morbidity, mortality, and economics of rheumatoid arthritis. *Arthritis Rheum* 2001; 44: 2746-9.
2. Gordon P, West J, Jones H, Gibson T. A 10 year prospective follow-up of patients with rheumatoid arthritis 1986-96. *J Rheumatol* 2001; 28: 2409-15.
3. Ward MM, Lubeck D, Leigh JP. Long-term health outcomes of patients with rheumatoid arthritis treated in managed care and fee-for-service practice settings. *J Rheumatol* 1998; 25: 641-9.
4. Leardini G, Salaffi F, Montanelli R, Gertzel S, Canesi B. A multicentric cost-of-illness study on rheumatoid arthritis in Italy. *Clin Exp Rheumatol* 2002; 20: 505-15.
5. Jääntti JK, Kaarela K, Belt EA, Kautiainen HJ. Incidence of severe outcome in rheumatoid arthritis during 20 years. *J Rheumatol* 2002; 29: 688-92.
6. Medeiros MMC, Ferraz MB, Quaresma MR. The effect of rheumatoid arthritis on the quality of life of primary caregivers. *J Rheumatol* 2000; 27: 76-83.
7. Cooper NJ. Economic burden of rheumatoid arthritis: a systematic review. *Rheumatology* 2000; 39: 28-33.
8. World Health Organization. International classification of impairments, disabilities and handicaps (ICIDH): a manual of classification relating to the consequences of disease. Geneva: WHO, 1980.
9. Devins GM, Edworthy SM, Seland TP, Klein GM, Leendert CP, Mandin H. Differences in illness intrusiveness across rheumatoid arthritis, end-stage renal disease and multiple sclerosis. *J Nerv Ment Dis* 1993; 181: 377-81.
10. Lorish CD, Abraham N, Austin J, Bradley LA, Alarcon GS. Disease and psychosocial factors related to physical functioning in rheumatoid arthritis. *J Rheumatol* 1991; 18: 1150-7.
11. Barrett EM, Scott DG, Wiles NJ, Symmons DP. The impact of rheumatoid arthritis on employment status in the early years of disease: a UK community-based study. *Rheumatology (Oxford)* 2000; 39: 1403-9.
12. Yelin E, Meenan R, Nevitt M, Epstein W. Work disability in rheumatoid arthritis: effects of disease, social, and work factors. *Ann Rheum Dis* 1980; 95: 551-6.
13. Pincus T, Callahan LF, Sale WG, Brooks AL, Payne LE, Vaughn WK. Severe functional declines, work disability, and increased mortality in seventy-five rheumatoid arthritis patients studied over nine years. *Arthritis Rheum* 1984; 27: 864-72.
14. Yelin E, Henke C, Epstein W. The work dynamics of the person with rheumatoid arthritis. *Arthritis Rheum* 1987; 30: 507-12.
15. Pincus T, Mitchell JM, Burkhauser RV. Substantial work disability and earnings losses in individuals less than age 65 with osteoarthritis: comparisons with rheumatoid arthritis. *J Clin Epidemiol* 1989; 42: 449-57.
16. Callahan LF, Bloch DA, Pincus T. Identification of work disability in rheumatoid arthritis: physical, radiographic and laboratory variables do not add explanatory power to demographic and functional variables. *J Clin Epidemiol* 1992; 45: 127-38.
17. Reisine S, McQuillan J, Fifield J. Predictors of work disability in rheumatoid arthritis patients. A five-year followup. *Arthritis Rheum* 1995; 38: 1630-7.
18. Allaire SH, Anderson JJ, Meenan RF. Reducing work disability associated with rheumatoid arthritis: identification of additional risk factors and persons likely to benefit from intervention. *Arthritis Care Res* 1996; 9: 349-57.
19. Wolfe F, Hawley DJ. The longterm outcomes of rheumatoid arthritis: Work disability: a prospective 18 year study of 823 patients. *J Rheumatol* 1998; 25: 2108-17.
20. De Roos AJ, Callahan LF. Differences by sex in cor-

- relates of work status in rheumatoid arthritis patients. *Arthritis Care Res* 1999; 12: 381-91.
21. Makisara GL, Makisara P. Prognosis of functional capacity and work capacity in rheumatoid arthritis. *Clin Rheumatol* 1982; 1: 117-25.
  22. Borg G, Allander E, Berg E, Brodin U, From A, Trang L. Auranofin treatment in early rheumatoid arthritis may postpone early retirement. Results from a 2-year double blind trial. *J Rheumatol* 1991; 18: 1015-20.
  23. Eberhardt K, Larsson BM, Nived K. Early rheumatoid arthritis-some social, economical, and psychological aspects. *Scand J Rheumatol* 1993; 22: 119-23.
  24. Doeglas D, Suurmeijer T, Krol B, Sanderman R, van Leeuwen M, van Rijswijk M. Work disability in early rheumatoid arthritis. *Ann Rheum Dis* 1995; 54: 455-60.
  25. Mau W, Bornmann M, Weber H, Weidemann HF, Hecker H, Raspe HH. Prediction of permanent work disability in a follow-up study of early rheumatoid arthritis: results of a tree structured analysis using RECPAM. *Br J Rheumatol* 1996; 35: 652-9.
  26. Fex E, Larsson BM, Nived K, Eberhardt K. Effect of rheumatoid arthritis on work status and social and leisure time activities in patients followed 8 years from onset. *J Rheumatol* 1998; 25: 44-50.
  27. Salaffi F, Ricciardi MR, Carotti M, Peroni M, Cervini C. La disabilità al lavoro nell'artrite reumatoide all'esordio: Valutazione prospettica di 6 anni. *Reumatismo* 1998; 50 (Suppl 3): abstract 248.
  28. Sokka T, Kautiainen H, Mottonen T, Hannonen P. Work disability in rheumatoid arthritis 10 years after the diagnosis. *J Rheumatol* 1999; 26: 1681-5.
  29. Albers JM, Kuper HH, van Riel PL, Prevoe ML, van 't Hof MA, van Gestel AM, et al. Socio-economic consequences of rheumatoid arthritis in the first years of the disease. *Rheumatology* 1999; 38: 423-30.
  30. Jäntti J, Aho K, Kaarela K, Kautiainen H. Work disability in an inception cohort of patients with seropositive rheumatoid arthritis: a 20 year study. *Rheumatology* 1999; 38: 1138-41.
  31. Young A, Biddle P, Cox N. Predicting work disability in early rheumatoid arthritis. Results from a 5 years inception cohort of 677 patients (abstract). *Ann Rheum Dis* 1999; 58(suppl.): 92.
  32. Young A, Dixey J, Cox N, Davies P, Devlin J, Emery P, et al. How does functional disability in early rheumatoid arthritis (RA) affect patients and their lives? Results of 5 years of follow-up in 732 patients from the Early RA Study (ERAS). *Rheumatology* 2000; 39: 603-11.
  33. Chorus AMJ, Miedema HS, Wevers CWJ, van der Linden. Work factors and behavioural coping in relation to withdrawal from the labour force in patients with rheumatoid arthritis. *Ann Rheum Dis* 2001; 60: 1025-32.
  34. Steinbrocker O, Traeger CH, Batterman RC. Therapeutic criteria in rheumatoid arthritis. *JAMA* 1949; 140: 659-662.
  35. Ragan C, Farrington E. The clinical features of rheumatoid arthritis. *J Am Med Assoc* 1962; 181: 663-7.
  36. Duthie JJR, Brown PE, Truelove LH, Baragar FD, Lawrie AJ. Course and prognosis in rheumatoid arthritis. A further report. *Ann Rheum Dis* 1964; 23: 193-202.
  37. Rasker JJ, Cosh JA. The natural history of rheumatoid arthritis: a fifteen year follow up study. *Clin Rheumatol* 1984; 3: 11-20.
  38. Fries JF, Spitz PW, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23: 137-45.
  39. Pincus T, Callahan LF, Brooks RH, Fuchs HA, Olsen NJ, Kaye JI. Self-report questionnaire scores in rheumatoid arthritis compared with traditional physical, radiographic and laboratory measures. *Ann Intern Med* 1989; 110: 259-66.
  40. Houssein DA, McKenna SP, Scott DL. The Nottingham Health Profile as a measure of disease activity and outcome in rheumatoid arthritis. *Br J Rheumatol* 1997; 36: 69-73.
  41. Scott DL, Pugner K, Kaarela K, Doyle DV, Woolf A, Holmes J, et al. The links between joint damage and disability in rheumatoid arthritis. *Rheumatology* 2000; 39: 122-32.
  42. Wolfe F, Kleinheksel SM, Cathay MA, Hawley D, Spitz PW, Fries JF. The clinical value of the Stanford Health Assessment Questionnaire Functional Disability Index in patients with rheumatoid arthritis. *J Rheumatol* 1988; 15: 1480-8.
  43. Sherrer YS, Bloch DA, Mitchell DM, Young DY, Fries JF. The development of disability in rheumatoid arthritis. *Arthritis Rheum* 1986; 29: 494-500.
  44. Lassere M, Wells G, Tugwell P, Edmonds J. Percentile curve reference charts of physical function: rheumatoid arthritis population. *J Rheumatol* 1995; 22: 1241-6.
  45. Lassere M, Houssein D, Scott D, Edmonds J. Reference curves of radiographic damage in patients with RA: application of quantile regression and fractional polynomials. *J Rheumatol* 1997; 24: 1288-94.
  46. Leigh JP, Fries JF, Parikh N. Severity of disability and duration of disease in rheumatoid arthritis. *J Rheumatol* 1992; 19: 1906-11.
  47. Wolfe F, Hawley DJ, Cathay MA. Clinical and health status measures over time: prognosis and outcome assessment in rheumatoid arthritis. *J Rheumatol* 1991; 18: 190-7.
  48. Greenwood M, Scott DL, Carr AJ, Deodhar A, Dolan L, Doyle DV, et al. Pain and disability in rheumatoid arthritis. *Ann Rheum Dis* 1999; 58(Suppl): 110.
  49. Brekke M, Hjortdahl P, Kvien TK. Self-efficacy and health status in rheumatoid arthritis : a two-year longitudinal observational study. *Rheumatology* 2001; 40: 387-92.
  50. Wiles N, Barrett E, Silman A, Symmons D. Disability in patients with early inflammatory polyarthritis cannot be "tracked" from year to year: an examination of the hypothesis underlying percentile reference charts. *J Rheumatol* 1999; 26: 800-4.
  51. Ward MM, Leigh JP, Fries JF. Progression of functional disability in patients with rheumatoid arthritis.

- Associations with rheumatology subspecialty care. *Arch Intern Med* 1993; 153: 2229-37.
52. Gardiner PV, Sykes HR, Hassey GA, Walker DJ. An evaluation of the health assessment questionnaire in the long-term longitudinal follow-up of disability in rheumatoid arthritis. *Br J Rheumatol* 1993; 32: 724-8.
  53. Guillemin F, Suurmeijer T, Krol B, Bombardier C, Briancon S, Doeglas D, et al. Functional disability in early rheumatoid arthritis: description and risk factors. *J Rheumatol* 1994; 21: 1051-5.
  54. Callahan LF, Pincus T, Huston JW, Brooks RH, Nance EP, Kaye JJ. Measures of activity and damage in rheumatoid arthritis: depiction of changes and prediction of mortality over five years. *Arthritis Care Res* 1997; 10: 381-94.
  55. Leymarie F, Jolly D, Sanderman R, Briancon S, Marchand AC, Guillemin F et al. Life events and disability in rheumatoid arthritis: a European cohort. *Br J Rheumatol* 1997; 36: 1106-12.
  56. Ward MM, Lubeck D, Leigh JP. Longterm health outcomes of patients with rheumatoid arthritis treated in managed care and fee-for-service practice settings. *J Rheumatol* 1998; 25: 641-9.
  57. Munro R, Hampson R, McEntegart A, Thomson EA, Madhok R, Capell H. Improved functional outcome in patients with early rheumatoid arthritis treated with intramuscular gold: results of a five year prospective study. *Ann Rheum Dis* 1998; 57: 88-93.
  58. Eberhardt KB, Fex E. Functional impairment and disability in early rheumatoid arthritis. Development over 5 years. *J Rheumatol* 1995; 22: 1037-42.
  59. Wolfe F, Choi HK. Benchmarking and the percentile assessment of RA: adding a new dimension to rheumatic disease measurement. *Ann Rheum Dis* 2001; 60: 994-6.
  60. Landewé RBM, Boers M, Verhoeven AC, Westhovens R, van de Laar MAFJ, Markusse HM, et al. COBRA combination therapy in patients with early rheumatoid arthritis. *Arthritis Rheum* 2002; 46: 347-56.
  61. Wiles NJ, Symmons DPM, Barrett EM, Silman AJ. Is it possible to construct percentile reference charts of disability for patients with early inflammatory polyarthritis? *Br J Rheumatol* 1998; 37: 136.
  62. Greenwood MC, Doyle DV, Ensor M. Does the Stanford Health Assessment Questionnaire have potential as a monitoring tool for subjects with rheumatoid arthritis? *Ann Rheum Dis* 2001; 60: 344-8.
  63. Wolfe F, Pincus T, Fries JF: Usefulness of the HAQ in the clinic [letter]. *Ann Rheum Dis* 2001; 60: 811.
  64. Matthews JNS, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *BMJ* 1990; 300: 230-5.
  65. Boers M, Tugwell P, Felson DT, van Riel PL, Kirwan JR, Edmonds JP, et al. World Health Organisation and International League of Association for Rheumatology core endpoints for symptom modifying antirheumatic drugs in rheumatoid arthritis clinical trials. *J Rheumatol* 1994; 41 (suppl): 86-9.
  66. Guidelines for the clinical investigation of drugs used in rheumatic diseases. World Health Organization Regional Office for Europe, European League Against Rheumatism 1983; 1-33.
  67. Tugwell P, Boers M for the OMERACT Committee. Developing consensus on preliminary core efficacy endpoints for rheumatoid arthritis clinical trials. *J Rheumatol* 1993; 20: 555-6.
  68. Testa MA, Simonson DC. Assessment of quality of life outcomes. *N Engl J Med* 1996; 28: 835-40.
  69. Garratt A, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *BMJ* 2002; 324: 1417-22.
  70. Salaffi F, Stancati A. Valutazione della qualità della vita nelle malattie reumatiche: strumenti e metodi. In: Scale di valutazione e malattie reumatiche. Edizioni. Mattioli 1885, pag. 3-20, 2001.
  71. Guyatt GH, Feeny DH, Patrick DL. Measuring health-related quality of life. *Ann Intern Med* 1993; 118: 622-9.
  72. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991; 44: 1271-8.
  73. Guyatt GH, van Zanten SJOV, Feeny DH, Patrick DI. Measurement quality of life in clinical trials: a taxonomy and review. *Can Med Assoc* 1989; 140: 1441-8.
  74. Deyo RA, Inui TS. Toward clinical applications of health status measures: sensitivity of scales to clinically important changes. *Health Serv Res* 1984; 19: 277-89.
  75. Ware JE Jr, Sherbourne CD. The MOS 36-item short form health survey (SF-36). 1. Conceptual framework and item selection. *Med Care* 1992; 30: 473-81.
  76. Sanderman R, Stewart R. The assessment of psychological distress: psychometric properties of the General Health Questionnaire (GHQ). *Int J Health Sciences* 1990; 1: 195-202.
  77. Bergner N, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981; 19: 787-805.
  78. Hunt SM, McEven J, McKenna SP. Measuring health status: a new tool for clinicians and epidemiologists. *J Roy Coll Gen Pract* 1985; 35: 185-8.
  79. Chambers LW. The McMaster Health Index Questionnaire: assessment of quality of life in clinical trials of cardiovascular therapies. Edited by NK Wenger, ME Mattson, J Elison. New York. Lejaco Publishing, 1984.
  80. Hurst NP, Jobanputra P, Hunter M, and the Economic and Health Outcomes Research Group. Validity of EuroQoL. A generic health status instrument in patient with rheumatoid arthritis. *Br J Rheumatol* 1994; 33: 655-62.
  81. Apolone G, Mosconi P. The Italian SF-36 Health Survey: translation, validation and norming. *J Clin Epidemiol* 1998; 51: 1025-36.
  82. Hopman WM, Towheed T, Anastassiades T, Tenenhouse A, Poliquin S, Berger C, et al. Canadian nor-

- mative data for the SF-36 health survey. *CMAJ* 2000; 163: 265-71.
83. Apolone G, Mosconi P. Come usare il questionario sullo stato di salute SF-36 (versione italiana). Progetto IQOLA. Istituto di Ricerche Farmacologiche "Mario Negri", 1996.
  84. Ware JE, Kosinski M, Keller SD. SF-36 physical and mental health summary scales: a user's manual. Boston: The Health Institute, New England Medical Centre; 1994.
  85. Ware J, Kosinski M, Bayliss M, Rogers WH, Raczec A. Comparison of methods for the scoring and statistical analysis of SF-36 Health Profile and summary measures: summary of results from the Medical Outcomes Study. *Med Care* 1995; 4: AS264-79.
  86. Brazier JE, Harper R, Munro J, Walters SJ, Snaith ML. Generic and condition-specific outcome measures for people with osteoarthritis of the knee. *Rheumatology* 1999; 38: 870-7.
  87. Fransen M, Edmonds J. Reliability and validity of the EuroQol in patients with osteoarthritis of the knee. *Rheumatology* 1999; 38: 807-13.
  88. Ruta DA, Hurst NP, Kind P, Hunter M, Stubbings A. Measuring health status in British patients with rheumatoid arthritis: reliability, validity and responsiveness of the short form 36-item health survey (SF-36). *Br J Rheumatol* 1998; 37: 425-36.
  89. Sullivan M, Ahlmèn M, Bjelle A. Health status assessment in rheumatoid arthritis. I. Further work on the validity of the Sickness Impact Profile. *J Rheumatol* 1990; 17: 439-47.
  90. Chung YL, Mitchell HL, Houssien DA, Al-Mahrouki H, Carr AJ, Scott DL. A comparative study of outcome in myositis and other musculoskeletal disorders assessed using the Nottingham health profile. *Clin Exp Rheumatol*. 2001; 19: 447-50.
  91. Hurst NP, Kind P, Ruta D, Hunter M, Stubbings A. Measuring health-related quality of life in rheumatoid arthritis: validity, responsiveness and reliability of EUROQOL (EQ-5D). *Br J Rheumatol* 1997; 36: 551-9.
  92. Wolfe F, Hawley DJ. Measurement of the quality of life in rheumatic disorders using the EuroQol. *Br J Rheumatol* 1997; 36: 786-93.
  93. Torrance G. Measurement of health status utilities for economic appraisal. *J Health Econ*. 1986; 5: 1-30.
  94. Torrance G. Utility approach to measuring health related quality of life. *J Chronic Dis* 1987; 40: 593-600.
  95. Balaban D, Sagi PV, Godfarb NI, Nettle S. Weights for scoring the QWB instrument among RA patients: a comparison to general population weights. *Med Care* 1986; 24: 973-80.
  96. Bleichrodt H, Johannesson M. Standard gamble, time trade-off and rating scale: experimental results on the ranking properties of QALYs. *J Health Econ* 1997; 16: 155-75.
  97. Bennett KJ, Torrance GW. Measuring health state preferences and utilities: rating scales, time trade-off and standard gamble techniques. In Spilker B ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*, 2nd ed. Philadelphia: Lippincott-Raven 1996: 253-65.
  98. Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health status measurement: a clarification. *J Clin Epidemiol* 1989; 42: 403-8.
  99. Meenan RF, Gertman MP, Mason JH. Measuring health status in arthritis. *The Arthritis Impact Measurement Scales*. *Arthritis Rheum* 1980; 23: 146-52.
  100. Meenan RF, Mason JH, Anderson JJ, Guccione AA, Kazis LE. AIMS2. The content and properties of a revised and expanded Arthritis Impact Measurement Scales health status questionnaire. *Arthritis Rheum* 1992; 35: 1-10.
  101. Tugwell P, Bombardier C, Buchanan W, Goldsmith CH, Grace E, Hanna B. The MACTAR Patient Preference Disability Questionnaire. An individualized functional priority approach for assessing improvement in physical disability in clinical trials in rheumatoid arthritis. *J Rheumatol* 1987; 14: 446-51.
  102. Jette AM, Deniston OL. Inter-observer reliability of a functional status assessment instrument. *J Chron Dis* 1978; 31: 573-80.
  103. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder, and hand). *Am J Industrial Med* 1996; 29: 602-8.
  104. Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *J Hand Ther* 2001; 14: 128-46.
  105. Hochberg MC, Vignon E, Maheu E. Session 2: clinical aspects. Clinical assessment of hand OA. *Osteoarthritis Cartilage* 2000; 8 Suppl A: S38-40.
  106. Dreiser RL, Maheu E, Guillou GB, Caspard H, Grouin JM. Validation of an algofunctional index for osteoarthritis of the hand. *Rev Rhum (Engl ed)* 1995; 62 (suppl 1): 43S-53S.
  107. Dreiser RL, Maheu E, Guillou GB. Sensitivity to change of the functional index for hand osteoarthritis. *Osteoarthritis Cartilage*. 2000; 8 Suppl A: S25-8.
  108. Salaffi F, Ferraccioli GF, Troise Rioda W, Carotti M, Sacchini G, Cervini C. Validità ed affidabilità della versione italiana dell'Arthritis Impact Measurement Scales in pazienti con artrite reumatoide. *Rec Prog Med* 1992; 83: 7-11.
  109. Haavardsholm EA, Kvien TK, Uhlig T, Smedstad LM, Guillemin F. A comparison of agreement and sensitivity to change between AIMS2 and a short form of AIMS2 (AIMS2-SF) in more than 1,000 rheumatoid arthritis patients. *J Rheumatol* 2000; 27: 2810-6.
  110. Salaffi F, Piva S, Barreca C, Cacace E, Ciancio G, Leardini G, et al. Validation of an Italian version of the arthritis impact measurement knee. Gonarthrosis and Quality of Life Assessment (GOQOLA) Study Group. *Rheumatology* 2000; 39: 720-7.
  111. Verhoeven AC, Boers M, van der Linden S. Validity

- of the MACTAR questionnaire as a functional index in a rheumatoid arthritis clinical trial. *J Rheumatol* 2000; 27: 2801-9.
112. Ranza R, Marchesoni A, Calori G et al. The Italian version of the functional disability index of the Health Assessment Questionnaire. A reliable instrument for multicenter studies on rheumatoid arthritis. *Clin Exp Rheumatol* 1993; 11: 123-8.
  113. Salaffi F, Carotti M, Cervini C. Health Assessment Questionnaire: la versione italiana di uno strumento di valutazione della qualità della vita nel paziente con artrite reumatoide. *Adria Med* 1997; 19: 5-11.
  114. Pincus T, Summey JA, Soraci SA, Jr., Wallston KA, Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum* 1983; 26: 1346-53.
  115. Wolfe F. Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): analyses in 2491 rheumatoid arthritis patients following leflunomide initiation. *J Rheumatol* 2001; 28: 982-9.
  116. Felton DT, Anderson JJ, Boers A, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis Rheum* 1993; 36: 729-40.
  117. Aha AS, Morran C, Sturrock RD, Russell RI. The health assessment questionnaire as a predictor of non-steroidal peptic ulceration. *Br J Rheumatol* 1993; 32: 135-8.
  118. Taha AS, Morran C, Sturrock RD, Russell RI. The Health Assessment Questionnaire as a predictor of non-steroidal peptic ulceration. *Br J Rheumatol* 1993; 32: 135-8.
  119. Kazis EL, Anderson JJ, Meenan R. Health status as a predictor of mortality in rheumatoid arthritis: a five-year study. *J Rheumatol* 1990; 17: 609-13.
  120. van der Heijde. Impact of rheumatoid arthritis on physical function during the first five years. No longer a question mark? *Rheumatology* 2000; 39: 579-80.
  121. Navsarikar A, Gladman DD, Husted JA, Cook RJ. Validity assessment of the disabilities of arm, shoulder, and hand questionnaire (DASH) for patients with psoriatic arthritis. *J Rheumatol* 1999; 26: 2191-4.
  122. Duruoz MT, Poiradeau S, Fermanian J, Menkes CJ, Amor B, Dougados M, et al. Development and validation of a rheumatoid hand functional disability scale that assesses functional handicap. *J Rheumatol* 1996; 23: 1167-72.
  123. Backman C, Maheu E, Guillou GB, Caspard H, Grouin JM. Arthritis hand function test: inter-rater reliability among self-trained raters. *Arthritis Care Res* 1995; 8: 10-5.
  124. Hochberg MC, Vignon E, Maheu E. Session 2: clinical aspects. *Clinical assessment of hand OA. Osteoarthritis Cartilage* 2000; 8 Suppl A: S38-40.
  125. Jones G, Cooley HM, Bellamy N. A cross-sectional study of the association between Heberden's nodes, radiographic osteoarthritis of the hands, grip strength, disability and pain. *Osteoarthritis Cartilage* 2001; 9: 606-11.
  126. Guyatt GH, Kirshner B, Jaeschke R. Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol* 1992; 45: 1341-5.
  127. Stucki G, Daltroy L, Katz JN, Johannesson M, Liang MH. Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. *J Clin Epidemiol* 1996; 49: 711-7.
  128. Salaffi F, Stancati A, Carotti M. Responsiveness of Health Status Measures and Utility-Based Methods in Patients with Rheumatoid Arthritis. *Clin Rheumatol* 2002; 21: 478-87.
  129. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Controlled Clin Trials* 1991; 12: 142S-158S.
  130. Deyo RA, Centor RM. Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. *J Chron Dis* 1986; 39: 897-906.
  131. Kazis L, Anderson J, Meenan R. Effect sizes for interpreting changes in health status. *Med Care* 1989; 27: S178-9.
  132. Hauley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143: 29-36.
  133. Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988; 240: 1285-93.
  134. Hauley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983; 148: 839-43.