

LAVORO ORIGINALE

Validazione delle versioni italiane del Bath Ankylosing Spondylitis Functional Index (BASFI) e del Dougados Functional Index (DFI) in pazienti con spondilite anchilosante

Validation of the Italian versions of the Bath Ankylosing Spondylitis Functional Index (BASFI) and the Dougados Functional Index (DFI) in patients with ankylosing spondylitis

F. Salaffi¹, A. Stancati¹, A. Silvestri¹, M. Carotti², W. Grassi¹

¹Cattedra di Reumatologia; ²Istituto di Radiologia, Università Politecnica delle Marche

SUMMARY

Objectives: *The Bath Ankylosing Spondylitis Functional Index (BASFI) and the Dougados Functional Index (DFI) are the most commonly used instruments to measure functioning in ankylosing spondylitis (AS). The aim of this study was to translate, adapt and validate these instruments into the Italian language.*

Methods: *The BASFI and DFI questionnaires were translated into Italian by two independent bilingual physicians who were familiar with the medical aspects of AS and by one professional translator. Two rheumatologists familiar with instrument validation, and who were aware of the purpose of the study, examined semantic, idiomatic and conceptual issues and produced by consensus unified versions of each instrument. English back-translations from the Italian were done by a professional translator unaware of the original version. Both English versions were compared, and where needed, modifications to the Italian versions were made.*

Results: *A total of 95 patients were included: 77 males, age (mean±SD) 47.9±9.3 years, and disease duration 12.4±6.6 years, and 18 females, age 45.9±8.7 years, and disease duration 11.3±8.2 years. Reliability, measured in 23 patients participating a physiotherapy program, showed an acceptable one-week test-retest intraclass correlation coefficient (ICC) - BASFI ICC: 0.91, 95% CI: 0.87-0.94 and DFI ICC: 0.86, 95% CI: 0.83-0.90. The internal consistency was 0.90 (Cronbach's alpha) for the BASFI and 0.87 for the DFI. For validity the functional indices were correlated with the Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), Bath Ankylosing Spondylitis Metrology Index (BASMI), Bath Ankylosing Spondylitis Patient Global Score (BAS-G), modified Health Assessment Questionnaire (HAQ-S), SF-36 physical component summary (SF-36 PCS), stiffness, pain, physician's assessment of disease activity, Bath AS Radiology Index-total (BASRI-t), erythrocyte sedimentation rate (ESR), and C-reactive protein (CRP). The functional indices (BASFI and DFI) were correlated with each other ($p < 0.0001$) and with activity variables. There was no significant relationship between functional indices and BASRI-t and acute phase reactants. The receiver operating characteristic (ROC) curve analysis indicated that the BASFI ranked superior compared to HAQ-S, ($p = 0.019$) and SF36 PCS ($p = 0.002$), but not respect to DFI ($p = NS$), in distinguishing between patients with high and low disease activity.*

Conclusions: *The Italian versions of the BASFI and DFI showed adequate reliability and validity in patients with AS. Because of psychometric advantages, the BASFI may be preferred in clinical trial settings. However, sensitivity to changes due to drug therapy and/or rehabilitation remains to be determined.*

Reumatismo, 2005; 57(3):161-173

Indirizzo per la corrispondenza:

Dott. Fausto Salaffi
Clinica Reumatologica
Università Politecnica delle Marche
Ospedale A. Murri, Via dei Colli, 52
60035 Jesi (AN)
E-mail: fsalaff@tin.it

INTRODUZIONE

In questi ultimi anni la codifica di criteri standardizzati per la valutazione dell'attività e della gravità di malattia, ha rappresentato uno degli obiettivi della ricerca nel campo della spondilite anchi-

losante (SA) (1-12). Fra i numerosi fattori che rendono difficoltosa la valutazione ed il monitoraggio della SA, quello principale è costituito dall'elevato grado di variabilità inter-individuale nelle diverse espressioni della malattia (interessamento assiale e/o periferico, coinvolgimento delle entesi, consistenti manifestazioni extrarticolari, quali uveiti, aritmie ipocinetiche), nella evolutività e, quindi, nella prognosi (13-18).

La valorizzazione del ruolo d'autovalutazione degli effetti del trattamento e dell'impatto della malattia sullo stato di salute ha favorito lo sviluppo di scale e di questionari orientati alla valutazione funzionale, che hanno trovato largo impiego nei trial terapeutici (7, 19-21). Fra le numerose scale specifiche di valutazione della disabilità fisica, il BASFI (Bath Ankylosing Spondylitis Functional Index) (22) ed il DFI (Dougados Functional Index) (23) sono stati recentemente adottati dall'ASAS (ASsessment in Ankylosing Spondylitis) Working Group (10) e dall'OMERACT IV (Outcome Measures in Rheumatology) (11) quali strumenti fondamentali nella valutazione del paziente spondilico in trial terapeutici (8, 24, 25). Tali strumenti si sono dimostrati affidabili, validi ed in grado di esplorare le variazioni dello stato di malattia nel tempo (26, 27). Essi hanno, inoltre, confermato il loro valore predittivo in termini di inabilità al lavoro (28-30), del grado di "consumo" di risorse sanitarie, dei costi sociali diretti ed indiretti (28-31) e dell'incidenza di mortalità (13).

Scopo di questo studio è stato quello di valutare l'affidabilità, in termini sia di riproducibilità al test-retest che di consistenza interna, e la validità (convergente e discriminante) delle versioni italiane di tali strumenti di valutazione funzionale nei pazienti con SA.

MATERIALI E METODI

Traduzione e verifica dell'equivalenza semantica dei questionari

Nella fase preliminare si è resa necessaria la traduzione italiana e la verifica quali-quantitativa dell'equivalenza semantica degli strumenti. Le versioni in lingua inglese di BASFI e DFI (22, 23) sono state, pertanto, tradotte da due medici, esperti in problematiche cliniche della SA e da un interprete professionista, specializzato in traduzione di testi medico-scientifici. Due reumatologi, con provata esperienza nella validazione di scale di valutazione ed informati circa gli obiettivi dello studio, han-

no verificato la struttura semantica, linguistica e concettuale di ciascuno strumento. La versione italiana è stata tradotta nuovamente in inglese da un traduttore professionista "in cieco" rispetto alla versione originale. Le due versioni in lingua inglese, originale e back-translated, sono state confrontate e, dove necessario, la versione italiana è stata modificata. La seconda fase del processo ha, invece, comportato la verifica dell'accettabilità e della comprensibilità degli strumenti di valutazione da parte di un ulteriore campione di 23 pazienti con SA, non coinvolti nello studio di validazione (campione di convenienza), in trattamento fisiochinesiterapico. La versione definitiva dei questionari specifici di valutazione funzionale ha dimostrato una comprensibilità ed una accettabilità da parte della stragrande maggioranza dei pazienti (BASFI: 96% e 92% dei casi, rispettivamente; DFI: 90% e 88% dei casi, rispettivamente). Nessuna variazione o adattamento culturale degli "items" si è resa necessaria per nessuno dei due questionari.

Casistica e valutazione clinico-radiologica

Lo studio è stato condotto su 95 pazienti affetti da SA, diagnosticata in accordo ai criteri Europei (32), consecutivamente reclutati presso gli ambulatori e la sezione di degenza della Clinica Reumatologica - Università Politecnica delle Marche. Tutti i pazienti sono stati sottoposti a valutazione clinico-radiologica e successivamente indagati, previo consenso informato, mediante somministrazione delle scale e dei questionari di autovalutazione. In particolare sono stati rilevati i seguenti parametri:

- *Intensità del dolore*, riferito nell'ultima settimana, mediante l'utilizzo di una scala numerica (Numerical Rating Scale - NRS) a 11 livelli (0 = dolore assente; 100 = dolore insopportabile).
- *Intensità della rigidità*, riferita nell'ultima settimana, mediante l'utilizzo di una NRS a 11 livelli (0 = rigidità assente; 100 = massima rigidità).
- *Stato globale di salute*, mediante Bath Ankylosing Spondylitis Patient Global Score (BAS-G) (33). Si tratta di un indice composito costituito da due "items" concernenti lo stato generale di salute, strutturati su una NRS; il primo quesito concerne lo stato globale di salute riferito all'ultima settimana, mentre il secondo agli ultimi sei mesi. Le possibili risposte (0 = nessun effetto; 100 = gravissimi effetti) generano un punteggio ottenuto dalla media dei singoli valori (33).
- *Valutazione della compromissione funzionale*, mediante Bath Ankylosing Spondylitis Metrology Index (BASMI) (34). Il BASMI è un indice antro-

pometrico composito, proposto per la valutazione della mobilità del rachide e dell'anca. Esso è costituito da 5 misurazioni cliniche rappresentate dalla distanza trago-parete, dalla flessione lombare ottenuta mediante il test di Shöber modificato (34, 35), dalla rotazione cervicale, dalla flessione laterale del busto e dalla distanza intermalleolare. Ad ognuna delle singole misurazioni vengono assegnati valori compresi fra 0 e 2. Il punteggio finale potrà, pertanto, variare da 0 (migliore mobilità) a 10 (peggiore mobilità).

- *Valutazione dell'attività di malattia*, mediante *Bath Ankylosing Spondylitis Disease Activity Index (BASDAI)* (5). Il BASDAI è un indice composito di attività di malattia costituito da sei "items" che indagano l'intensità del dolore a livello del collo, della schiena, delle anche e delle articolazioni periferiche, il dolore provocato dal contatto/pressione, nonché l'intensità della rigidità mattutina (5). Le risposte, riportate su NRS, generano un punteggio compreso in un range di 0-10, dove 0 corrisponde ad una "attività assente" e 10 ad una "massima attività". Il paziente impiega, mediamente, 4-5 minuti per la compilazione del questionario (5).

- *Valutazione radiologica semiquantitativa*, mediante il *Bath Ankylosing Spondylitis Radiology Index-total (BASRI-t)* (36). Il BASRI è uno score globale, che prevede la valutazione delle lesioni elementari del rachide evidenziabili alla radiologia convenzionale, quali erosioni, squaring, sclerosi, sindesmofitosi, fino alla fusione di più corpi vertebrali (36). Il BASRI spine (BASRI-s) prevede la valutazione del rachide cervicale (sulla sola proiezione latero-laterale) e lombare (valutato sulla proiezione latero-laterale e anteroposteriore) e delle articolazioni sacroiliache (sulla sola proiezione anteroposteriore). Il punteggio è compreso fra 0 e 4, sia per il rachide cervicale, che viene valutato dal margine inferiore di C1 al margine superiore di C7, che per il rachide lombare, considerato dal margine inferiore di D12 al margine superiore di S1. La valutazione è globale e lo score prevede 4 differenti stadi: 0 = normale (nessuna alterazione); 1 = dubbio (alterazioni probabili, non sicure lesioni); 2 = lieve (presenza di erosioni o squaring o sclerosi con o senza sindesmofiti in \leq due vertebre); 3 = moderato (sindesmofiti in ≥ 3 vertebre con o senza fusione di vertebre); 4 = severo (fusione interessante 3 o più vertebre). La valutazione delle sacro-iliache si basa sui criteri di New-York e prevede uno score che va da 0 a 5: stadio 0 = nessuna alterazione; stadio 1 = aspetto leggermente sfumato dei bordi articolari, pseudoallargamento o restringimento

della rima, lieve sclerosi subcondrale); stadio 2 = irregolarità dei margini articolari con immagini di erosioni, restringimento della rima, sclerosi subcondrale evidente; stadio 3 = erosioni articolari e sclerosi subcondrale ben evidenti con iniziali ponti sinostotici; stadio 4 = completa anchilosi dell'articolazione sacro-iliaca. Il BASRI-s comprende la valutazione di questi 3 score: rachide cervicale (0-4), rachide lombare (0-4) e delle articolazioni sacroiliache (0-4), pertanto lo score globale va da 2-12. Al BASRI-s è stata aggiunta la valutazione delle articolazioni coxofemorali, valutate dal radiogramma in proiezione anteroposteriore, utilizzato per lo studio delle articolazioni sacroiliache senza la necessità di eseguire radiogrammi aggiuntivi (37). Lo score per l'anca (BASRI-hip) va da 0 a 4: 0 = normale (nessuna alterazione);

1 = dubbio (riduzione focale della rima articolare);
2 = lieve (riduzione concentrica della rima articolare >2 mm);

3 = moderato (riduzione concentrica della rima articolare inferiore o uguale a 2 mm, oppure neoapposizione ossea - ponte osseo - <2 cm);

4 = severo (deformità ossea o neoapposizione - ponte osseo - >2 cm). Il BASRI-s viene sommato al BASRI-hip e si ottiene il BASRI totale (BASRI-t) il cui score è quindi compreso fra 0 e 16.

In tutti i pazienti sono state, infine, rilevate la velocità di eritrosedimentazione (VES, metodo Westergreen, mm/1° ora) e la proteina C-reattiva (PCR, mm/dl).

Scale e questionari di valutazione

BASFI. Sviluppato da un gruppo multidisciplinare di reumatologi, fisiatristi, fisioterapisti e pazienti (22), il BASFI è costituito da 10 "items" presentati su scale analogiche visive (VAS) di 100 mm (range 0-100). Il questionario, esplora il grado di difficoltà riscontrato nello svolgimento di specifiche attività della vita quotidiana, mediante domande strutturate secondo la formula: "È in grado di...?". Gli ancoraggi estremi delle VAS per le risposte ai quesiti sono strutturati come "Facilmente" (score 0) o "È impossibile" (score 100) (22). La compilazione del questionario, concepito sia per l'autosomministrazione che per l'intervista faccia a faccia, richiede 3-4 minuti. Il punteggio della scala viene calcolato come media dei punteggi dei 10 items ed è compreso, pertanto, in un range di 0-100 corrispondente, rispettivamente, al migliore ed al peggiore stato funzionale. Le recenti conferme in letteratura (38) e la nostra personale esperienza (39), riguardo la sostanziale equivalenza biometrica del-

Indice funzionale di Dougados per la spondilite anchilosante

	Si, senza difficoltà	Si, con qualche difficoltà	No, è impossibile
1. È in grado di mettersi le scarpe?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. È in grado di tirar su i pantaloni?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. È in grado di indossare un maglione?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. È in grado di entrare nella vasca da bagno?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. È in grado di rimanere in piedi per 10 minuti?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. È in grado di salire una rampa di scale?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. È in grado di correre?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. È in grado di sedersi?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. È in grado di alzarsi da una sedia?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. È in grado di entrare in automobile?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. È in grado di piegarsi per raccogliere un oggetto?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. È in grado di abbassarsi sulle gambe?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. È in grado di sdraiarsi?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. È in grado di girarsi nel letto?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. È in grado di alzarsi dal letto?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. È in grado di dormire supino (a pancia in su)?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Totale:	-----		

Bath Ankylosing Spondylitis Functional Index (BASFI)

Risponda a ciascuna domanda indicando la risposta con una crocetta sulla linea graduata da 0 a 100. Tutte le domande si riferiscono al grado di limitazione che la malattia le ha provocato nell'ultima settimana

1. È in grado di mettere i calzini o i collant senza l'aiuto di un'altra persona o di attrezzi particolari?

FACILE

0	10	20	30	40	50	60	70	80	90	100
---	----	----	----	----	----	----	----	----	----	-----

 IMPOSSIBILE

2. È in grado di piegarsi per raccogliere una matita o un oggetto dal pavimento senza l'impiego di attrezzi particolari?

FACILE

0	10	20	30	40	50	60	70	80	90	100
---	----	----	----	----	----	----	----	----	----	-----

 IMPOSSIBILE

3. È in grado di raggiungere o prendere un oggetto su una mensola posta al di sopra della testa senza l'aiuto di un'altra persona o di attrezzi particolari?

FACILE

0	10	20	30	40	50	60	70	80	90	100
---	----	----	----	----	----	----	----	----	----	-----

 IMPOSSIBILE

4. È in grado di alzarsi da una sedia senza braccioli senza aiutarsi con le braccia o senza l'aiuto di un'altra persona?

FACILE

0	10	20	30	40	50	60	70	80	90	100
---	----	----	----	----	----	----	----	----	----	-----

 IMPOSSIBILE

5. È in grado di alzarsi in piedi dalla posizione coricata senza l'aiuto di un'altra persona?

FACILE

0	10	20	30	40	50	60	70	80	90	100
---	----	----	----	----	----	----	----	----	----	-----

 IMPOSSIBILE

6. È in grado di stare in piedi per 10 minuti senza l'aiuto di un'altra persona e senza avvertire difficoltà?

FACILE

0	10	20	30	40	50	60	70	80	90	100
---	----	----	----	----	----	----	----	----	----	-----

 IMPOSSIBILE

7. È in grado di salire 12-15 gradini senza l'utilizzo del corrimano o di un bastone, mettendo un piede su ciascun gradino?

FACILE

0	10	20	30	40	50	60	70	80	90	100
---	----	----	----	----	----	----	----	----	----	-----

 IMPOSSIBILE

8. È in grado di guardare alle spalle girando il collo, senza dover girare l'intero corpo?

FACILE

0	10	20	30	40	50	60	70	80	90	100
---	----	----	----	----	----	----	----	----	----	-----

 IMPOSSIBILE

9. È in grado di effettuare esercizi fisici (es. esercizi nell'ambito della fisioterapia, attività sportiva, tempo libero)?

FACILE

0	10	20	30	40	50	60	70	80	90	100
---	----	----	----	----	----	----	----	----	----	-----

 IMPOSSIBILE

10. È in grado di svolgere le comuni attività della vita quotidiana al proprio domicilio o nel posto di lavoro?

FACILE

0	10	20	30	40	50	60	70	80	90	100
---	----	----	----	----	----	----	----	----	----	-----

 IMPOSSIBILE

Totale: _____

le scale numeriche (NRS) rispetto a quelle analogiche (VAS) e la maggiore accettabilità e comprensibilità da parte dei pazienti delle NRS (39), ci hanno indotto a sostituire, nella versione italiana del questionario, le scale analogiche con quelle numeriche a 11 livelli (appendice 1).

DFI. Sviluppato a partire da 29 items identificati da un gruppo di reumatologi con provata esperienza nel campo della SA e pubblicato per la prima volta nel 1988 (23), il DFI è costituito da 20 "items" presentati sottoforma di "Likert scale" a 3 livelli di risposta (range compreso fra 0 e 2). Il questionario, proposto inizialmente con il metodo dell'intervista e successivamente anche con l'auto-somministrazione, esplora il grado di abilità del soggetto nello svolgimento di specifiche attività della vita quotidiana. Le domande sono strutturate secondo la seguente formula: "È in grado di...?". Le opzioni di risposta, corrispondenti ad un punteggio di 0, 1 o 2, sono rispettivamente: "Sì, senza difficoltà"; "Sì, con qualche difficoltà"; "No, è impossibile" (appendice 2). La compilazione del questionario richiede 3-5 minuti. Per quanto attiene l'algoritmo di calcolo del punteggio, formalmente non esistono studi pubblicati riguardo la gestione delle risposte mancanti. Secondo le indicazioni dell'autore, tuttavia, il questionario non è da ritenersi valido in caso di mancata risposta a tre o più quesiti (23). Per un numero di "missing data" minore o uguale a due, il punteggio totale viene calcolato come media dei punteggi alle domande completate moltiplicato per venti. Tale algoritmo genera un punteggio compreso in un range di 0-40, dove 0 corrisponde alla migliore funzione e 40 alla massima compromissione funzionale possibile.

HAQ-S. L'HAQ comprende 20 "items" riguardanti altrettanti atti della vita quotidiana, suddivisi in 8 differenti categorie: lavarsi e vestirsi, alzarsi, mangiare, camminare, igiene, raggiungere oggetti, afferrare, e altre attività (40). La versione dell'HAQ modificata per le spondiloartriti (HAQ-S) include due ulteriori domini che valutano la funzionalità del collo e della schiena (41). Ogni "item" consente 4 possibili 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 singola categoria viene considerato il punteggio più alto; la somma dei punteggi diviso il nu-

mero di categorie rappresenta il punteggio finale dell'HAQ-S, che può variare, quindi, da un minimo di 0 a un massimo di 3 (41). Il questionario viene, di solito, completato in 4-5 minuti, senza particolari difficoltà di comprensione.

SF-36. L'SF-36 prevede otto scale, articolate in gruppi distinti di quesiti (da due a dieci quesiti) e da un quesito singolo sul cambiamento delle condizioni di salute, che non viene impiegata per l'attribuzione del punteggio in nessuna delle otto scale a quesiti multipli (42). Indici sintetici che descrivono globalmente lo stato di salute fisica (Physical Component Summary Scale Score - PCS) e quello mentale (Mental Component Summary Scale Score - MCS) sono stati ottenuti dall'aggregazione delle relative sottoscale (42). 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 in punteggi di scala) sono fissati da procedure matematiche standardizzate. Il questionario, nella versione italiana (43), viene solitamente completato dal paziente in 10-15 minuti. Per il nostro studio abbiamo considerato la sola dimensione fisica dell'SF-36 (SF-36 PCS).

Analisi statistica

L'analisi statistica è stata condotta mediante MedCalc (versione 7.4.2 per Windows XP, Belgio). L'analisi della distribuzione dei dati dei questionari e delle scale di valutazione, mediante Kolmogorov-Smirnov test ha dimostrato, per diversi domini e sub-domini dei questionari di valutazione, una distribuzione non-gaussiana. Sono state, pertanto, adottate anche procedure non-parametriche per l'analisi dei risultati. In particolare, i punteggi delle scale e dei questionari sono stati espressi sia come media e deviazione standard che come mediana con relativi intervalli di confidenza (I.C.) della mediana al 95%. Il confronto fra i punteggi è stato effettuato mediante procedure non-parametriche (Kruskall-Wallis test e Wilcoxon test), mentre la comparazione fra le caratteristiche demografiche è stata attuata mediante ANOVA e test chi-square. L'affidabilità, espressa sia in termini di riproducibilità al test-retest (grado di variabilità dei punteggi a seguito di misurazioni ripetute in condizioni analoghe), che di consistenza interna [grado di inter-correlazione fra i quesiti costituenti le singole

sottoscale del questionario (*item-internal consistency*) è stata valutata, rispettivamente, mediante il coefficiente di correlazione intra-classe (ICC) e il coefficiente “alpha di Cronbach”. In entrambi i casi il range dei valori è compreso fra 0 (non riproducibilità, minima consistenza interna) e 1 (massima riproducibilità, massima consistenza interna). Valori maggiori di 0.80 sono considerati accettabili in un trial clinico. Lo studio delle matrici di correlazioni (validità concorrente) fra i domini delle differenti scale e dei questionari è stato effettuato mediante lo studio delle matrici di correlazione per ranghi “rho” di Spearman. La validità discriminante degli strumenti generici e specifici è stata indagata mediante analisi delle curve operative caratteristiche (curve ROC). 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) (44, 45). 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 singolo 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 con malattia attiva vs non attiva. Come criterio esterno per dicotomizzare i pazienti attivi vs non attivi è stato impiegato, in accordo ai criteri ASAS (9, 10), un valore di BASDAI $> 0 < a 40$ mm. 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 accuratezza (44). Il confronto fra i valori medi delle aree sottese alle curve ROC è stato effettuato in accordo al metodo proposto da Hauley e McNeil (46).

RISULTATI

Casistica

Lo studio è stato condotto su 95 pazienti affetti da

SA, di cui 77 uomini con età media pari a $47,9 \pm 9,3$ anni e durata media di malattia di $12,4 \pm 6,6$ anni e 18 donne con età media pari a $45,9 \pm 8,7$ anni e durata media di malattia di $11,3 \pm 8,2$ anni. I pazienti, consecutivamente reclutati presso gli ambulatori e la sezione di degenza della Clinica Reumatologica - Università Politecnica delle Marche, sono stati sottoposti a valutazione clinica e successivamente a somministrazione delle scale e dei questionari di autovalutazione. Il tempo medio di somministrazione dei questionari specifici è risultato pari a 3,8 minuti ($\pm 1,7$) per l’HAQ-S, pari a 4,1 minuti ($\pm 2,4$) per il DFI, pari a 3,9 minuti ($\pm 2,2$) per il BASFI, mentre l’SF-36 ha comportato tempi di compilazione da parte dei pazienti assai più lunghi rispetto a tutti gli strumenti specifici ($13,8 \pm 7,4$ minuti) ($p < 0,0001$). In figura 1 vengono rappresentate le frequenze di distribuzione dei punteggi dei singoli strumenti autosomministrati. In tutti i casi la distribuzione dei dati è di natura non parametrica (Kolmogorov-Smirnov test for normal distribution).

In tabella I vengono riportati i valori medi (\pm deviazione standard), la mediana ed i relativi intervalli di confidenza al 95% (I.C. 95%) delle principali variabili demografiche, cliniche e radiologiche dei pazienti studiati.

Affidabilità

La riproducibilità al test-retest è stata calcolata in 23 pazienti, dopo risomministrazione dei questionari a distanza di una settimana dalla valutazione basale, in un contesto clinico invariato rispetto al controllo basale. La verifica della stabilità del contesto clinico è stata effettuata mediante l’ausilio di un questionario transizionale, già impiegato in precedenti studi (39). L’analisi dei dati ha consentito di dimostrare, in tutti i casi, una buona affidabilità ($> 0,80$). In particolare, il BASFI è risultato più riproducibile del DFI (BASFI: CCI = 0,91; 95% I.C. 0,87-0,94 e DFI = CCI 0,86; 95% I.C. 0,83-0,90). La verifica della consistenza interna, intesa come grado di inter-correlazione fra i quesiti costituenti le singole sottoscale del questionario, ha dimostrato valori *alpha* di Chronbach superiori al limite raccomandato di 0,80 per entrambi i questionari specifici: BASFI = 0,90 e DFI = 0,87.

Validità concorrente

Lo studio delle matrici di correlazione (test *rho* di Spearman) fra le scale di valutazione ha evidenziato correlazioni reciprocamente significative ($p < 0,0001$) tra tutti gli strumenti specifici (BASFI, DFI e HAQ-S) e con l’SF-36 PCS (Tab. II). Eleva-

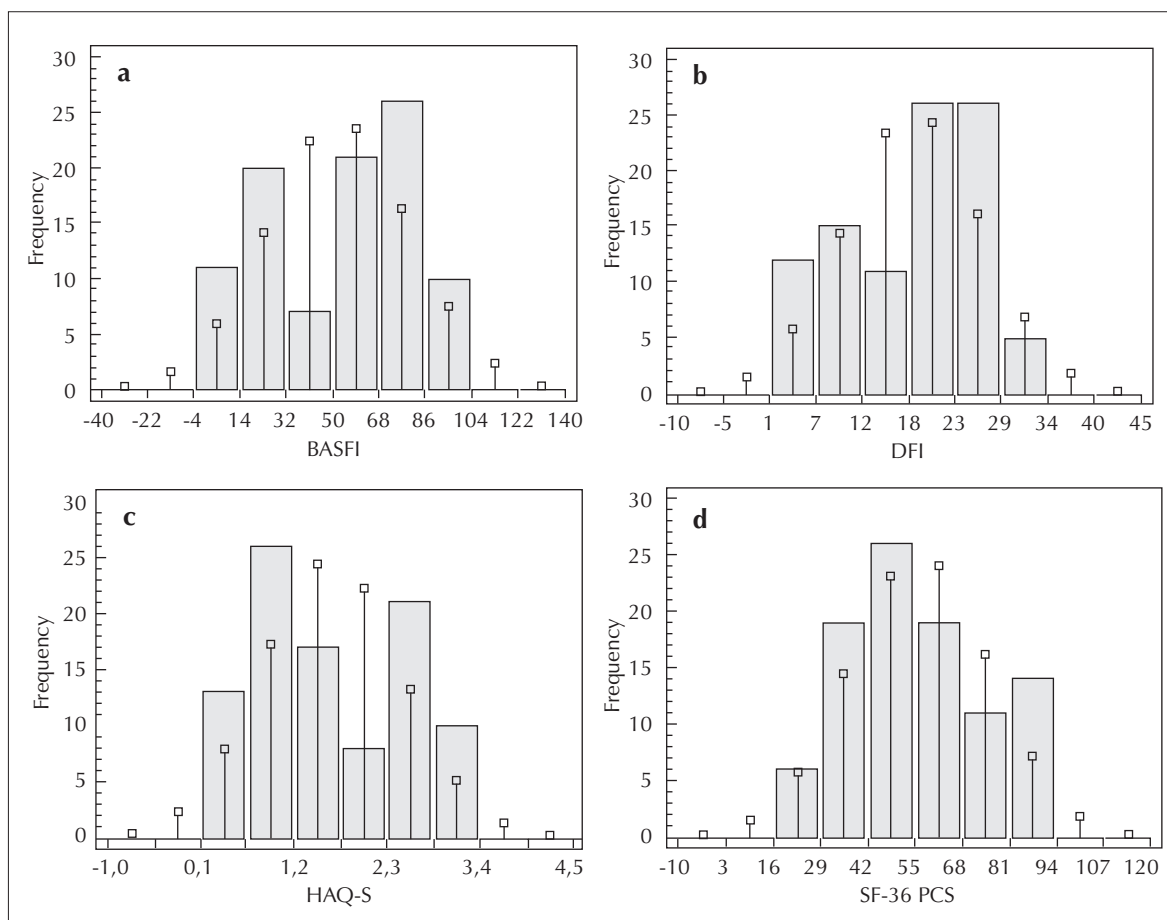


Figura 1 - Frequenze di distribuzione dei punteggi dei questionari specifici di valutazione funzionale (BASFI, DFI ed HAQ-S) e generici (SF-36 PCS). In tutti i casi la distribuzione dei dati è di natura non parametrica (Kolmogorov-Smirnov test for normal distribution).

te correlazioni sono state anche rilevate rispetto agli indici di attività di malattia (Tab. II) ed in particolare al BASDAI (Fig. 2). Il BASFI, rispetto al DFI mostra i coefficienti di correlazione più elevati (Fig. 2a). La compromissione funzionale, valutata mediante l'indice composito BASMI, è risultata correlata ($p < 0,001$) con le scale specifiche di valutazione (BASFI, DFI e HAQ-S), ma non con l'SF-36 PCS ($R_s = 0,016$; $p = \text{NS}$) (Tab. II). Nessuna significativa correlazione è emersa fra i punteggi totali di BASFI e DFI rispetto all'età, alla durata di malattia, ai test aspecifici di flogosi (VES e PCR) ed alla severità del danno strutturale (BASRI-t) (validità divergente) (Tab. II).

Validità discriminante

La validità discriminante dei questionari specifici e generici (SF-36 PCS) è stata ulteriormente indagata mediante l'analisi delle curve ROC (Fig. 3). I pazienti sono stati dicotomizzati in due gruppi (46

pazienti con basso grado di attività di malattia e 49 pazienti con malattia attiva), utilizzando, in accordo alle raccomandazioni dell'ASAS (9, 10), un valore $< 0 >$ di 40 mm del BASDAI. Il BASFI si conferma lo strumento dotato di maggiore capacità discriminante, specialmente se confrontato all'HAQ-S (differenze fra AUC = $0,115 \pm 0,05$; $p = 0,019$) ed all'SF-36 PCS (differenze fra AUC = $0,183 \pm 0,06$; $p = 0,002$). Sebbene non significativa ($p = 0,103$), l'AUC alla curva ROC è risultata maggiore anche rispetto a quella calcolata per il DFI (BASFI AUC-ROC = $0,901 \pm 0,032$ vs DFI AUC-ROC = $0,833 \pm 0,084$, rispettivamente) (Fig. 3).

DISCUSSIONE

I risultati del nostro studio, hanno dimostrato, in analogia con i dati riportati in letteratura (26, 47, 48) una buona affidabilità dei questionari specifici

Tabella I - Valori medi \pm deviazione standard (DS), mediana ed intervalli di confidenza (I.C.) al 95% delle variabili demografiche, cliniche e radiologiche dei pazienti con spondilite anchilosante (N = 95 pazienti).

	Media \pm DS	Mediana	I.C. 95% per la mediana
Età (anni)	47,7 \pm 9,2	47	44 - 49
Durata di malattia (anni)	12,2 \pm 6,8	12	10 - 14
VAS-dolore (range 0-100)	50,9 \pm 22,9	55	45 - 62
VAS-rigidità (range 0-100)	53,2 \pm 24,5	55	45,8 - 65,0
BAS-G (range 0-100)	51,7 \pm 24,9	40	40 - 64,2
BASMI (range 0-10)	5,69 \pm 1,82	6	5 - 6
BASDAI (range 0-100)	50,9 \pm 23,6	53	39,0 - 64,1
BASFI (range 0-100)	52,06 \pm 27,6	55	45,9 - 66,1
DFI (range 0-40)	18,0 \pm 8,11	20	17,0 - 22,0
HAQ-S (range 0-3)	1,63 \pm 0,82	1,5	1,47 - 1,80
SF-36 PCS (0-100)	56,1 \pm 19,3	53,5	48,7 - 62,3
PCR (mg /dl)	1,78 \pm 2,31	0,87	0,62 - 1,22
VES (mm 1 ^a ora)	29,4 \pm 16,1	26	23 - 30
BASRI-t (range 0-12)	6,0 \pm 3,1	5	5 - 6

Abbreviazioni: VAS = Visual Analogic Scale; BAS-G = Bath Ankylosing Spondylitis Patient Global Score; BASMI = Bath Ankylosing Spondylitis Metrology Index; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASRI = Bath Ankylosing Spondylitis Radiology Index; BASFI = Bath Ankylosing Spondylitis Functional Index; DFI = Disability Functional Index; HAQ-S = Health Assessment Assessment for Spondyloarthritis; SF-36 PCS = 36-Items Short Form Healthy Survey - Physical Component Summary Score (PCS); BASRI-t = Bath Ankylosing Spondylitis Radiology Index-total

Tabella II - Matrici di correlazioni rilevate fra le scale specifiche di valutazione funzionali (BASFI, DFI e HAQ-S), quelle generiche (SF-36 PCS) e le variabili cliniche e radiologiche nei pazienti affetti da spondilite anchilosante.

	BASFI	DFI	HAQ-S	SF-36 PCS
Età (anni)	0,127	0,110	0,169	-0,118
Durata di malattia (anni)	0,101	0,145	0,027	-0,112
BASMI (0-10)	0,498***	0,399**	0,343**	-0,016
BAS-G (0-100)	0,705***	0,544***	0,474***	-0,305*
VAS-dolore (0-100)	0,729***	0,531***	0,429***	-0,335**
VAS-rigidità (0-100)	0,697***	0,563***	0,438***	-0,374**
BASDAI (0-100)	0,722***	0,650***	0,523***	-0,356**
PCR (mg / dl)	0,039	0,109	0,033	-0,062
VES (mm 1 ^a ora)	0,119	0,141	0,034	-0,083
BASRI-t	0,125	0,138	0,080	-0,123

* = p < 0,05; ** = p < 0,001; *** = p < 0,0001 (correlazione per ranghi di Spearman). Abbreviazioni: VAS = Visual Analogic Scale; BAS-G = Bath Ankylosing Spondylitis Patient Global Score; BASMI = Bath Ankylosing Spondylitis Metrology Index; BASDAI = Bath Ankylosing Spondylitis Disease Activity Index; BASRI = Bath Ankylosing Spondylitis Radiology Index; BASFI = Bath Ankylosing Spondylitis Functional Index; DFI = Disability Functional Index; HAQ-S = Health Assessment Assessment for Spondyloarthritis; SF-36 PCS = 36-Items Short Form Healthy Survey - Physical Component Summary Score (PCS); BASRI-t = Bath Ankylosing Spondylitis Radiology Index-total

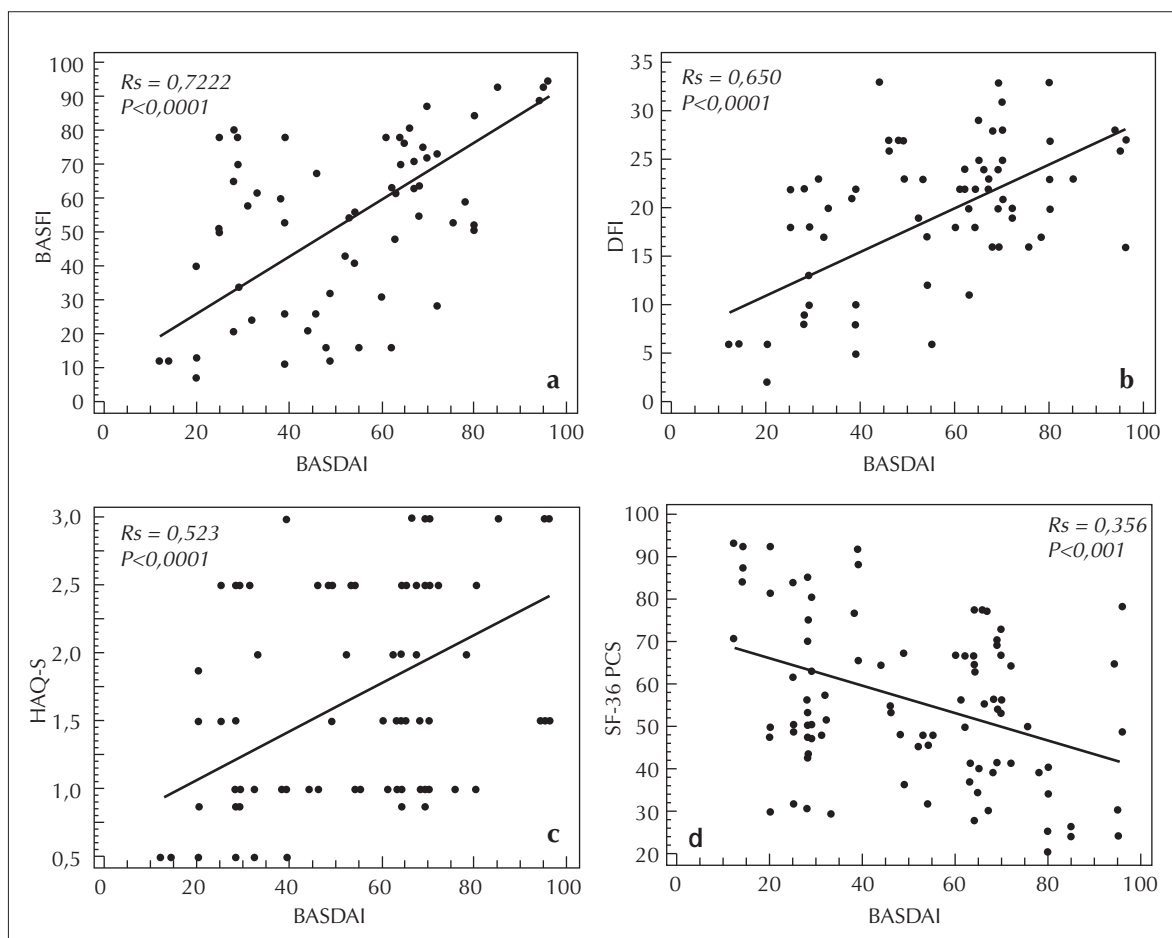


Figura 2 - Correlazione fra punteggi totali, espressi dal BASFI (Fig. 2a), dal DFI (Fig. 2b), dall'HAQ-S (Fig. 2c) e dal SF-36 PCS (Fig. 2d) e l'indice di attività di malattia (BASDAI) (test rho per ranghi di Spearman).

ci di valutazione della disabilità, sebbene il BASFI abbia mostrato la migliore concordanza al test-retest rispetto al DFI. Anche la verifica della consistenza interna ha dimostrato valori *alpha* di Cronbach superiori di 0,80 per i due questionari specifici, del tutto simili a quelli riportati per le versioni originali (22, 23, 49). In termini di validità concorrente le due scale di valutazione hanno confermato buone proprietà psicometriche. In particolare sono state osservate correlazioni reciprocamente significative tra tutti gli strumenti specifici (BASFI, DFI e HAQ-S) e con l'SF-36 PCS ed anche rispetto gli indici di attività di malattia, specie con il BASDAI. Ciò conferma, da un lato la stretta interrelazione fra la funzione fisica e le principali componenti che condizionano il grado di attività di malattia, quali il dolore e la rigidità (4, 20, 23, 29, 50, 51), dall'altro le proprietà di costrutto dei questionari specifici. Tali questionari sono, infatti, con-

cepiti allo scopo di fornire informazioni attinenti l'area di interesse nei pazienti con SA e sono incentrate su aspetti della salute tipici e sul quadro

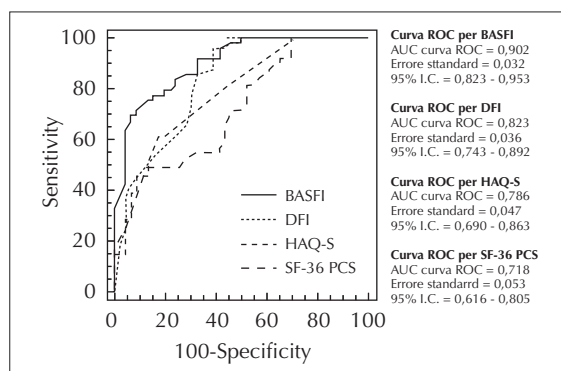


Figura 3 - Analisi della capacità discriminante dei questionari specifici (BASFI, DFI e HAQ-S) e generico (SF-36 PCS) mediante curve operative caratteristiche (curve ROC).

sintomatologico (mobilità del rachide, dolore, rigidità, attività della vita quotidiana, attività fisica, attività lavorativa). Il BASFI, tuttavia, rispetto al DFI si dimostra maggiormente correlato con gli indici di attività di malattia. Il grado di correlazione dei questionari specifici di autovalutazione della disabilità con la compromissione funzionale, valutata mediante l'indice composito BASMI, pur essendo risultata correlata, mostra valori di significatività inferiori. Una scarsa correlazione tra mobilità e disabilità è stata riportata da altri autori (4, 12, 23). Vitainen et al. (12), in particolare, avevano riportato una correlazione significativa solo con due delle cinque misure antropometriche costituenti il BASMI, ovvero con la distanza dita-suolo e con la flessione toraco-lombare. Nessuna significativa associazione è emersa, nel nostro studio, fra i punteggi del BASFI e del DFI rispetto all'età, alla durata di malattia, ai test aspecifici di flogosi ed alla severità del danno radiologico. Tali osservazioni confermano il ruolo scarsamente rilevante di VES e PCR nella valutazione dell'attività di ma-

lattia in corso di SA e quello relativo all'entità del danno radiologico (BARSII-t) nel condizionare il grado di impegno funzionale in questi pazienti (26, 50). Anche in termini di validità discriminante, il BASFI si conferma lo strumento maggiormente performante, specialmente rispetto all'HAQ-S ed all'SF-36 PCS. Calin et al. (3), comparando la capacità discriminante del BASFI rispetto al DFI in pazienti con SA in trattamento con farmaci antiinfiammatori non steroidei, ne hanno dimostrato la maggiore capacità di performance. Ruof e Stuki (26) da una revisione della letteratura, sono pervenuti ad analoghe conclusioni.

In conclusione, le versioni italiane del BASFI e del DFI hanno mostrato buona affidabilità e validità nei pazienti con SA. In virtù delle migliori proprietà psicometriche, il BASFI risulta, tuttavia, preferibile sia nel contesto di studi clinici che nella pratica clinica quotidiana. L'ulteriore tappa di validazione di tali strumenti dovrà perseguire la verifica della responsività rispetto ad interventi farmacologici e/o riabilitativi.

RIASSUNTO

Scopi: Il Bath Ankylosing Spondylitis Functional Index (BASFI) ed il Dougados Functional Index (DFI) sono gli indici di valutazione funzionale più ampiamente utilizzati nei pazienti con spondilite anchilosante (SA). Obiettivi del presente studio sono la traduzione, l'adattamento culturale e la validazione di questi strumenti per l'utilizzo nei pazienti italiani con SA.

Metodi: Le versioni in lingua inglese del BASFI e del DFI sono state tradotte da due medici bilingue esperti sulle problematiche cliniche della SA e da un interprete professionista specializzato in traduzione di testi medico-scientifici. Due reumatologi con provata esperienza nella validazione di scale di valutazione e informati circa gli obiettivi dello studio, hanno verificato la struttura semantica, linguistica e concettuale di ciascuno strumento. La versione italiana è stata tradotta nuovamente in inglese da un traduttore professionista "in cieco" rispetto alla versione originale. Le due versioni in lingua inglese, originale e back-translated, sono state confrontate e, dove necessario, la versione italiana è stata modificata.

Risultati: Sono stati studiati 95 pazienti: 77 maschi (età media 47,9±9,3 anni; durata di malattia 12,4±6,6 anni) e 18 femmine (età media 45,9±8,7 anni; durata di malattia 11,3±8,2 anni). I dati di affidabilità, ottenuti su un campione indipendente di 23 pazienti inclusi in un programma di fisioterapia, hanno mostrato un elevato coefficiente di correlazione intraclassa (ICC) rispetto ad un re-test eseguito a distanza di una settimana – BASFI ICC: 0.91, 95% IC: 0,87-0.94 e DFI ICC: 0.86, 95% IC: 0.83-0.90. La consistenza interna è risultata pari a 0,90 (Cronbach's alpha) per il BASFI e 0,87 per il DFI. Per la verifica della validità gli indici funzionali sono stati correlati con: Bath Ankylosing Spondylitis Disease Activity Index (BASDAI), Bath Ankylosing Spondylitis Metrology Index (BASMI), Bath Ankylosing Spondylitis Patient Global Score (BAS-G), Health Assessment Questionnaire modificato per le spondiloartriti (HAQ-S), componente fisica dell'SF-36 (SF-36 PCS), dolore, rigidità, VAS attività di malattia (medico), Bath AS Radiology Index-total (BASRI-t), VES e PCR. Gli indici funzionali (BASFI e DFI) hanno mostrato correlazioni significative, sia reciprocamente ($p < 0,0001$), che con le variabili di attività di malattia. Non sono emerse correlazioni significative tra indici funzionali e BASRI-t né con i test aspecifici di flogosi (VES e PCR). L'analisi delle curve operative caratteristiche (ROC) ha dimostrato la superiorità in termini di capacità discriminativa tra pazienti attivi vs non attivi, del BASFI rispetto all'HAQ-S ($p = 0,019$) ed all'SF-36 PCS ($p = 0,002$), ma non rispetto al DFI ($P = NS$).

Conclusioni: Le versioni italiane del BASFI e del DFI hanno mostrato buona affidabilità e validità nei pazienti con SA. In virtù delle caratteristiche psicometriche, il BASFI risulta preferibile agli altri indici, nel contesto di studi clinici. Sono, tuttavia, necessarie ulteriori ricerche per determinare la responsività di questi strumenti, in relazione ad interventi farmacologici e/o riabilitativi.

Parole chiave - Spondilite anchilosante, BASFI, DFI, studi di validazione, stato funzionale.

Key words - Ankylosing spondylitis, BASFI, DFI, validation studies, functional status.

BIBLIOGRAFIA

1. Anderson JJ, Baron G, Van der Heijde D, Felson DT, Dougados M. Ankylosing spondylitis assessment group preliminary definition of short term improvement in ankylosing spondylitis. *Arthritis Rheum* 2001; 44: 1867-86.
2. Braun J, Bollow M, Neure L, Seipelt E, Seyrekbasan F, Herbst H, et al. Use of immunohistologic and in situ hybridization techniques in the examination of sacroiliac joint biopsy specimens from patients with ankylosing spondylitis. *Arthritis Rheum* 1995; 38: 499-505.
3. Calin A, Nakache J, Gueguen A, Zeidler H, Mielants H, Dougados M. Outcome variables in ankylosing spondylitis: Evaluation of their relevance and discriminant capacity. *J Rheumatol* 1999; 26: 1571-8.
4. Calin A. The Dunlop-Dottridge Lecture. Ankylosing Spondylitis: defining disease status and the relationship between radiology, metrology, disease activity, function and outcome. *J Rheumatol of Rheumatology* 1995; 22: 740-4.
5. Garret S, Jenkison T, Kennedy LG, Whitelock H, Gaisford P, Calin A. A new approach to defining disease status in ankylosing spondylitis. The Bath Ankylosing Spondylitis Disease Activity. *J Rheumatol* 1994; 21: 2286-91.
6. Golder W, Bollow M, Sieper J, Van der Heijde D. Imaging and scoring in ankylosing spondylitis. *Clin Exp Rheumatol* 2002; 20 (Suppl 28): S178-84.
7. Haywood KL, Garrat AM, Jordan K, Dziedzic K, Dawes PT. Disease-specific, patient-assessed measures in health outcome in ankylosing spondylitis: reliability, validity and responsiveness. *Rheumatology* 2002; 41: 1295-302.
8. Pham T, Van der Heijde D, Calin A, Khan MA, Van der Linden S, Bellamy N, Dougados M. For the ASAS Working Group Initiation of biological agents in patients with ankylosing spondylitis: results of a Delphi study by the ASAS Group *Ann Rheum Dis* 2003; 62: 812-6.
9. Van der Heijde D, Bellamy N, Calin A, Dougados M, Khan MA, van der Linden S on behalf of the Assessment in Ankylosing Spondylitis Working Group. Preliminary core set for endpoints in ankylosing spondylitis. *J Rheumatol* 1997; 24: 2225-9.
10. Van der Heijde D, Calin A, Dougados M, Khan MA, van der Linden S, Bellamy N on behalf of the ASAS Working Group. Selection of instruments in the core set for DC-ART, SMARD, physical therapy, and clinical record keeping in ankylosing spondylitis: progress report of the ASAS Working Group. *J Rheumatol* 1999; 26: 951-4.
11. Van der Heijde D, van der Linden S, Bellamy N, Calin C, Dougados M, Khan MA. Which domains should be included in a core set for endpoints in ankylosing spondylitis: introduction to the ankylosing spondylitis module of OMERACT IV. *J Rheumatol* 1999; 26: 945-7.
12. Vittaten JV, Heikkila S, Kokko ML, Kautianien H. Clinical Assessment of Spinal Mobility Measurements in Ankylosing Spondylitis: A Compact Set for Follow-up and Trials? *Clin Rheumatol* 2000; 19: 131-7.
13. Braun J, Pincus T. Mortality, course of disease and prognosis of patients with ankylosing spondylitis. *Clin Exp Rheumatol* 2002; 20 (suppl. 28): S16-S22.
14. Braun J, Rudwaleit M, Boonen A, Zink A. Ankylosing spondylitis: an overview. *Ann Rheum Dis* 2002 a; 61: 8-18.
15. Gladman DD. Clinical aspects of the spondyloarthropathies. *Am J Med Sci* 1998; 316: 234-8.
16. Khan MA. Update on spondyloarthropathies. *Ann Intern Med* 2002; 136: 896-907.
17. Olivieri I, Salvarani C, Cantini F, Ciancio G, Padula A. Ankylosing spondylitis and undifferentiated spondyloarthropathies: a clinical review and description of a disease subset with older age at onset. *Curr Opin Rheumatol* 2001; 13: 280-4.
18. Olivieri I, Van Tubergen A, Salvarani C, Van der Linden S. Seronegative spondyloarthritides. *Best Pract Res Clin Rheumatol* 2002; 16: 723-39.
19. Abbott CA, Helliwell PS, Chamberlain MA. Functional Assessment in ankylosing spondylitis: evaluation of a new administered questionnaire and correlation with anthropometric variables. *Br J Rheumatol* 1994; 33: 1060-6.
20. Bostan EE, Borman P, Bodur H, Barça N. Functional disability and quality of life in patients with ankylosing spondylitis. *Rheumatol Int* 2003; 23: 121-6.
21. Rudwaleit M, Listing J, Brandt J, Braun J, Sieper J. Prediction of a major clinical response (BASDAI 50) to tumor necrosis factor α blockers in ankylosing spondylitis. *Ann Rheum Dis* 2004; 63: 665-70.
22. Calin A, Garrett S, Whitelock H, Kennedy LG, O'Hea J, Mallorie P, et al. A new approach to defining functional ability in ankylosing spondylitis: the development of the Bath Ankylosing Ankylosing Spondylitis Functional Index. *J Rheumatol* 1994; 21: 2281-5.
23. Dougados M, Gueguen A, Nakache J-P, Nguyen M, Mery C, Amor B. Evaluation of functional index and an articular index in ankylosing spondylitis. *J Rheumatol* 1988; 15: 302-7.
24. van Tubergen A, van der Heijde D, Anderson J, Landawé R, Dougados M, Braun J, et al. Comparison of statistically derived ASAS improvement criteria for ankylosing spondylitis with clinically relevant improvement according to an expert panel. *Ann Rheum Dis* 2003; 62: 215-21.
25. Braun J, Van der Heijde D, Dougados M, Emery P, Khan M A, Sieper J, et al. Staging of patients with ankylosing spondylitis: a preliminary proposal. *Ann Rheum Dis* 2002 c; 61 (Suppl III): iii 19-23.
26. Ruof J, Stucki G. Comparison of the Dougados Functional Index and the Bath Ankylosing Spondylitis Functional Index. A literature review. *J Rheumatol* 1999; 26: 955-60.
27. Ward MM. Health-related quality of life in ankylosing spondylitis: a survey of 175 patients. *Arthritis Care Res* 1999; 12: 247-55.
28. Boonen A, Chorus A, Miedema H, van der Heijde D,

- van der Tempel H, der Linden S. Employment, work disability, and work days lost in patients ankylosing spondylitis: a cross sectional study of Dutch patients. *Ann Rheum Dis* 2001; 60: 353-8.
29. Boonen A. Socioeconomic consequences of ankylosing spondylitis. *Clin Exp Rheumatol* 2002; 20 (Suppl 28): S23-S26.
30. Boonen A, van der Heijde D, Landewè R, Guillemin L, Rutten-van Mölken, Dougados M, et al. Direct costs of ankylosing spondylitis and its determinants: an analysis among three European countries. *Ann Rheum Dis* 2003; 62: 732-40.
31. Ward M.M. Functional disability predicts total costs in patients with ankylosing spondylitis. *Arthritis Rheum* 2002; 46: 223-31.
32. Dougados M, Van der Linden S, Juhlin R, Huitfeldt B, Amor B, Calin A, et al. The European Spondylarthropathy Study Group preliminary criteria for the classification of spondylarthropathy. *Arthritis Rheum* 1991; 34: 1218-27.
33. Jones SD, Steiner A, Garret SL, Calin A. The Bath ankylosing spondylitis patient global score (BAS-G) *Br J Rheumatol* 1996; 35: 66-71.
34. Jenkinson TR, Mallorie PA, Whitelock HC, Kennedy LG, Garrett SL, Calin A. Defining spinal mobility in ankylosing spondylitis (AS): the Bath AS Metrology Index. *J Rheumatol* 1994; 21: 1694-8.
35. Jones SD, Porter J, Garrett SL, Kennedy LG, Whitelock H, Calin A. A new scoring system for the Bath Ankylosing Spondylitis Metrology Index (BASMI). *J Rheumatol* 1995; 22: 1609.
36. Calin A, Mackay K, Santos H, Brophy S. A new dimension of outcome: application of the Bath Ankylosing Spondylitis Radiology Index. *J Rheumatol* 1999; 26: 998-1002.
37. MacKay K, Brophy S, Mack C, Doran M, Calin A. The development and validation of a radiographic grading system for the hip in ankylosing spondylitis: the Bath Ankylosing Spondylitis Radiology Hip Index. *J Rheumatol* 2000; 27: 2866-72.
38. van Tubergen A, Dabats I, Ryser L, Londono J, Burgos-Vargas R, Cardiel MH, et al. Use of a numerical rating scale as an answer modality in ankylosing spondylitis - specific questionnaire. *Arthritis Care Res* 2002; 47: 242-8.
39. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain* 2004; 8: 283-91.
40. Fries JF, Spitz PW, Kraines RG, Holman HR. Measurement of patient outcome in arthritis. *Arthritis Rheum* 1980; 23: 137-45.
41. Daltroy LH, Lasron MG, Roberts WN, Liang MH. A modification of the Health Assessment Questionnaire for the spondyloarthropathies. *J Rheumatol* 1990; 17: 946-50.
42. Ware JE Jr, Sherbourne CD. The MOS 36-item short form health survey (SF-36). Conceptual framework and item selection. *Med Care* 1992; 30: 473-81.
43. Apolone G, Mosconi P. The Italian SF-36 Health Survey: translation, validation and norming. *J Clin Epidemiol* 1998; 51: 1025-36.
44. 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.
45. Hauley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143: 29-36.
46. 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.
47. Cardiel MH, Londoño, Gutiérrez E, Pacheco-Tena C, Vázquez-Mellado J, Burgos-Vargas R. Translation, cross-cultural adaptation, and validation of the Bath ankylosing spondylitis functional index (BASFI), the Bath ankylosing spondylitis disease activity index (BASDAI) and the Dougados functional index (DFI) in a Spanish speaking population with spondyloarthropathies. *Clin Exp Rheumatol* 2003; 21: 451-8.
48. Eyres S, Tennant A, Kay L, Waxman R, Helliwell PS. Measuring disability in ankylosing spondylitis: comparison of Bath ankylosing spondylitis functional index with revised Leeds disability questionnaire. *J Rheumatol* 2002; 29: 979-86.
49. Ariza RA, Hernandez CB, Sarabia FN. Physical Function and health-related quality of life of spanish patients with ankylosing spondylitis. *Arthritis Care Res* 2003; 40: 463-7.
50. Spoorenberg A, van der Heijde D, de Klerk E, Dougados M, de Vlam K, Mielants H, et al. A comparative study of the usefulness of the Bath Ankylosing Spondylitis Functional Index and the Dougados Functional Index in the assessment of ankylosing spondylitis. *J Rheumatol* 1999; 26: 961-5.
51. Jones SD, Koh WH, Steiner A, Garrett SL, Calin A. Fatigue in ankylosing spondylitis: its prevalence and relationships with activity, sleep, and other factors. *J Rheumatol* 1996; 23: 487-90.