



# Ricerca in Riabilitazione



Periodico di aggiornamento scientifico  
Anno 8° - N° 1 - Dicembre 1999

## EDITORIALE

### Sette anni di sistema FIM-Functional Independence Measure™ in Italia: non soltanto una scala di misura.

di Luigi Tesio

Questo numero di Ricerca in Riabilitazione vuole essere un rapido strumento informativo per coloro che già conoscono la FIM™-Functional Independence Measure, uno strumento standard internazionale per la misura della disabilità nei settori riabilitativo ed assistenziale. La FIM non è

soltanto un questionario che registra dei numeri. Essa fa parte di un vero e proprio sistema integrato di servizi che comprende formazione e accreditamento degli operatori, misura delle condizioni funzionali del paziente, analisi di efficacia ed efficienza dell'assistenza in protocolli clinico-gestionali e controllo di qualità delle misure stesse.

La FIM e i relativi corsi di formazione sono gli elementi più noti di questo sistema. Soltanto in Italia sono oltre duemila gli operatori che hanno frequentato, fra il 1993 e il 1999, il Corso di accreditamento condotto da So.Ge.Com-Ricerca in Riabilitazione su licenza statunitense e quindi secondo lo standard internazionale. Meno note sono le altre potenzialità del sistema. A queste ultime So.Ge.Com accenna molto brevemente durante i corsi, per poi riservare ulteriori dettagli a coloro che chiedono di instaurare collaborazioni su progetti applicativi o di ricerca. Il gradimento riscontrato dalle iniziative didattiche e il prestigio scientifico conquistato dalla FIM

nella ricerca internazionale hanno fatto crescere progressivamente il numero di richieste di collaborazione. Queste provengono ormai da istituzioni pubbliche nazionali, regionali e locali, da associazioni di malati, da fondazioni scientifiche, da aziende ospedaliere, da industrie farmaceutiche e da singoli clinici e ricercatori. Per facilitare la comunicazione con questi utenti So.Ge.Com ha affidato a Ricerca in Riabilitazione il compito di produrre un riassunto tecnico-scientifico delle caratteristiche d'insieme del sistema FIM, con particolare attenzione al controllo di qualità dei dati con analisi di Rasch. Questo ultimo metodo è un punto di forza del sistema ma è anche l'aspetto meno intuitivo. Quello che segue è un riassunto rivolto a coloro che sono già entrati in contatto con il sistema FIM, senza che per questo debbano avere particolari competenze statistiche.

Come sempre Ricerca in Riabilitazione resta a disposizione per ogni ulteriore chiarimento. ■

## Il sistema FIM™

### CHE COSA È IL SISTEMA FIM.

di Ferdinando Bonfanti

Amministratore Unico, SO.GE.COM srl Milano

La Functional Independence Measure-FIM™ è una scala di misura della disabilità. Essa ha l'aspetto di un questionario e misura l'autosufficienza in 18 attività della vita quotidiana ritenute indispensabili e di regola anche sufficienti per la vita autonoma a domicilio. Per ogni attività il punteggio può variare fra 1 (completa dipendenza) e 7 (completa autosufficienza). Di conseguenza il punteggio totale può variare fra 18 e 126. La FIM fornisce una misura affidabile, valida e sensibile della disabilità complessiva dei pazienti con le più varie menomazioni neuromotorie: dall'emiparesi post-ictale alla protesi articolare, dalla sclerosi multipla alla poli-patologia invalidante geriatrica, fino alle patologie croniche psichiatrico-comportamentali. La scala FIM è ormai uno standard internazionale. Essa si è affermata come valido indice di appropriatezza e di efficacia dell'assistenza riabilitativa. Eppure, più che di

una scala bisognerebbe parlare di un vero e proprio "sistema FIM". Infatti il sistema comprende:

- a) un manuale e rigorose procedure di accreditamento dei rilevatori (è richiesto un Corso della durata di un giorno, concluso da un esame);
- b) una scheda socio-sanitaria con un insieme standard di dati. La scheda consente di correlare le misure FIM alle caratteristiche cliniche e socio-demografiche del paziente;
- c) una banca-dati centrale che analizza le prestazioni di migliaia di unità operative di riabilitazione negli Stati Uniti. La banca-dati fornisce servizi statistici che consistono soprattutto nel controllo di qualità delle rilevazioni e nel confronto fra i dati del singolo utente e i dati normativi generali.

### NASCITA DELLO STRUMENTO

La FIM è nata negli Stati Uniti fra il 1984 e il 1987. Un gruppo di lavoro diretto da Carl V Granger (cui si accennerà fra poco) ha risposto al mandato espresso congiuntamente da molte Associazioni Scientifiche: fisiatriche, infermieristiche, di fisioterapisti ecc. L'obiettivo era quello di

sviluppare una scala di misura dell'autosufficienza che avesse i seguenti requisiti:

- a) buone caratteristiche psicometriche (validità, riproducibilità, sensibilità ecc.);
- b) la proprietà di essere "disease and discipline free", ovvero applicabile indipendentemente dalla patologia che avesse causato la disabilità e dalle competenze specialistiche coinvolte nella misura e nel trattamento;
- c) accettazione da parte di tutto il settore riabilitativo, afflitto — allora come adesso — dalla continua proliferazione di scale di misura dell'autosufficienza tutte diverse fra loro.

### ASPETTI GIURIDICI E ORGANIZZATIVI

La titolare giuridica del sistema FIM è la UB Foundation Inc., un ente non-profit che è una filiazione della State University of New York (SUNY).

La SUNY, fondata nel 1846, è fra le più antiche e prestigiose Università pubbliche degli Stati Uniti. Essa ha diverse sedi nello Stato di New York, fra le quali una nella città di Buffalo ove si trova un famoso Policlinico con un Dipartimento di Medicina Riabilitativa. Qui la FIM è nata e

## Il controllo di qualità delle rilevazioni FIM: dai punteggi sul questionario a misure valide e affidabili

di Anna Simone

Tecnico Statistico

Luigi Tesio

Primario Fisiatra, Direttore dell'Unità di Ricerca, Valutazione Funzionale e Verifica di Qualità in Riabilitazione, IRCCS Fondazione Salvatore Maugeri, Pavia

### IL PROBLEMA: CONVALIDARE PUNTEGGI SOGGETTIVI

Le rilevazioni su questionari — tra le quali rientra la misura di autosufficienza — vengono definite "soggettive" perché non hanno un corrispettivo in dati strumentali quali una radiografia, un elettrocardiogramma, un dato di laboratorio. D'altro canto non esiste altro modo pratico di misurare l'autosufficienza che non sia quello di "dare un voto" all'autosufficienza stessa. Evidentemente la mancanza di questo tipo di oggettività rappresenta un difetto e un ostacolo per chiunque voglia utilizzare queste misure a fini di controllo di qualità, o addirittura a fini di controllo economico dell'assistenza. Una possibilità di "oggettivazione" di queste misure, tuttavia, esiste. La strategia più affermata consiste nell'analisi di verosimiglianza (*likelihood*), secondo tecniche statistiche definite nel loro complesso *item-response analysis-IRT*, laddove il termine *item* indica qualsiasi voce di un questionario. L'analisi di Rasch è la versione più avanzata di IRT per applicazioni all'area della Riabilitazione. Si tenterà ora una sintesi delle proprietà e delle applicazioni del modello, con richiami esemplificativi alle sue applicazioni alla scala FIM.

### UN NUOVO ORIZZONTE METODOLOGICO: IL MODELLO DI GEORG RASCH

Il modello proposto negli anni 1960/1980 dal matematico danese Georg Rasch si sta affermando in Riabilitazione sia per la costruzione, sia per la validazione di scale comportamentali item/risposta. Inoltre si stanno affermando anche le sue applicazioni al controllo di qualità dei dati. Infatti, una volta che si sia stabilita la validità della scala il modello si presta molto bene al controllo di verosimiglianza delle risposte ai questionari.

Il modello prescrive che la probabilità di risposta categorica (0 invece che 1, no/sì ecc.) a un certo item (per esempio: alimentazione, locomozione, comprensione ecc.) sia dettata da due parametri soltanto: la difficoltà intrinseca dell'item e la abilità intrinseca del soggetto, lungo una stessa variabile unidimensionale che rappresenta un costrutto concettuale continuo "da meno a più".

A partire da una matrice di risposte ordi-

nali grezze (diversi items con risposte no/sì = 0/1, oppure no/lieve/medio = 0/1/2 ecc.) il modello stima in un quadro probabilistico la "massima verosimiglianza" dei due parametri (difficoltà e abilità, per ciascun item e per ciascun soggetto), ovvero i valori che più avvicinano i dati osservati a dati modello-compatibili. Soggetti più abili avranno più probabilità di superare uno qualsiasi degli items, rispetto a soggetti meno abili. Items più difficili avranno maggiori probabilità di "resistere" ad uno qualsiasi dei soggetti, rispetto ad items più facili. La relazione che lega fra loro la probabilità di successo (per un soggetto rispetto a un item, o viceversa) e questi due parametri è dettata da un'equazione relativamente semplice che rappresenta il cuore stesso del modello di Rasch. Appositi algoritmi producono la stima "massimamente verosimile" di abilità dei soggetti e di difficoltà degli items. Inoltre, anche se la spiegazione va oltre gli scopi di questo articolo è importante sapere che il metodo risolve classici problemi psicometrici quali:

- la discontinuità e la non proporzionalità dei punteggi ordinali: che cosa impedisce di definire i livelli "dipendente/autonomo con ausilio/indipendente" con livelli 0/4/7/9 invece che 0/1/2? Che cosa impone di credere che la differenza fra 2 e 1 rappresenti veramente (e non soltanto quanto a simboli numerici) lo stesso intervallo quantitativo che separa 1 da 0?
- la stima di eventuali risposte mancanti. Che cosa ci impone di ritenere che due soggetti che rispondano a 9 items su 10 abbiano davvero conseguito la sessa misura (9/9 o, se si preferisce, 9/10), se ciascuno omette un diverso item con difficoltà diversa?

### LA CONVALIDA DEI PUNTEGGI: CORRISPONDENZA FRA PREVISIONI DEL MODELLO E DATI OSSERVATI

Poiché il modello di Rasch si muove in un orizzonte probabilistico si calcola anche il margine di errore delle stime di abilità e di difficoltà. Viene quindi misurata l'aderenza ("fit") delle risposte osservate (per esempio le 18 risposte al questionario FIM) alle attese del modello.

Se le attese sono rispettate i punteggi di un certo soggetto o di un certo item vengono in qualche modo convalidati (si parla di *objective measurement*). Viceversa, soggetti o items che accumulino troppe risposte inattese divengono oggetto di una vera e propria procedura diagnostica molto sofisticata e sensibile che è volta a definire le cause di questa non-verosimiglianza.

Per esempio è dimostrato che nella FIM l'autosufficienza nell'alimentazione risulta una prestazione più probabile, in qualsiasi patologia, rispetto all'autosufficienza nel salire le scale. Che cosa si deve pensare di un questionario FIM dal quale il soggetto appaia autosufficiente nel fare le scale ma non nell'alimentarsi? Vi è stato un errore di codifica? Oppure il paziente presenta caratteristiche

cliniche molto atipiche? E se invece vi fossero molti soggetti con lo stesso schema di punteggio? In questo caso vi potrebbe essere un difetto nella omogeneità (*unidimensionality*) degli stessi items. La domanda diviene: alimentazione e scale rappresentano una stessa variabile, così da fornire punteggi cumulabili, oppure rappresentano variabili con scarsa o nulla relazione fra loro, il che spiega il loro rapporto erratico nei diversi soggetti? Oppure ancora il rilevatore compie un errore sistematico nell'attribuire i punteggi ad uno dei due items: per esempio, egli sopravvaluta le difficoltà incontrate nell'alimentazione?

### LA TRASFORMAZIONE DEI PUNTEGGI GREZZI IN MISURE INTERVALLARI

La unidimensionalità è il primo requisito non soltanto perché i punteggi dei diversi items siano sommabili, ma anche perché essi siano riproducibili. Se il punteggio di un item varia in modo erratico rispetto a quello degli altri items, molto probabilmente anche il punteggio totale varierà in modo erratico da una rilevazione all'altra.

Con l'analisi di Rasch l'osservazione comportamentale diviene correlabile appropriatamente a classiche misure intervallari (per esempio fisiche o econometriche). Il termine *intervallare* ricorda che i simboli numerici non sono soltanto ordinati in sequenza come le quantità che rappresentano (per esempio, 2 chilogrammi vuol dire "più peso" di 1 chilogrammo), ma che essi sono anche proporzionali al "vero intervallo" che separa le grandezze che essi rappresentano. Per esempio, l'intervallo di peso che separa la misura di 1 chilogrammo dalla misura di 2 chilogrammi è pari alla metà dell'intervallo di peso che separa la misura di 5 chilogrammi dalla misura di 7 chilogrammi. Al contrario, non è affatto noto — a meno, appunto, di trasformazioni come quelle consentite dall'analisi di Rasch — quale differenza di autosufficienza rappresenti il passaggio da un punteggio FIM di 1 ad un punteggio di 2, rispetto al passaggio da un punteggio di 5 ad un punteggio di 7.

La stima di misure veramente intervallari della abilità dei soggetti rende più valide le procedure statistiche che presuppongano variabili intervallari (ad esempio, test sulle medie invece che sulle mediane, correlazioni, analisi della varianza, regressioni ecc.).

### APPLICAZIONI DELLA STIMA DI "VEROSIMIGLIANZA"

**a) Applicazione a soggetti e rilevatori:** identificazione dei casi atipici, per il controllo delle rilevazioni su questionari

Se si pensa ad un utilizzo di queste misure per scopi gestionali e di controllo di qualità diviene particolarmente interessante la possibilità di stimare la coerenza di una serie di risposte rispetto alle attese del modello. Il punteggio totale non basta: evidentemente, anche risposte date completamente a caso potrebbero produrre un punteggio cumulativo verosimile. La disponibilità di una

continua — per così dire — ad avere la propria residenza ufficiale.

Lo sviluppo pratico del sistema è affidato a *Uniform Data System for Medical Rehabilitation-UDSMr*, un'agenzia tecnico-scientifica che afferisce anch'essa alla SUNY di Buffalo. UDSMr tiene i corsi FIM in tutti gli Stati Uniti, gestisce la banca dati centrale e offre assistenza tecnica per studi internazionali. Il braccio scientifico di UDSMr è il *Center for Functional Assessment Research-CFAR* che afferisce al Department of Rehabilitation Medicine della stessa Università. Il CFAR non si occupa soltanto di studi che comprendano la scala FIM, ma svolge anche ricerche a tutto campo nel settore dello sviluppo, della validazione e dell'applicazione clinica, gestionale ed econometrica di misure psico-comportamentali in area riabilitativa. Direttore sia di UDSMr, sia del CFAR, sia del Dipartimento di Medicina Riabilitativa presso la SUNY è il prof. Carl V. Granger, medico fisiatra che da decenni è forse la più autorevole voce a livello internazionale in tema di misure funzionali in Riabilitazione.

#### UN'APPLICAZIONE GESTIONALE IMPORTANTE: LA CLASSIFICAZIONE PER GRUPPI "ISO-RISORSE" IN RIABILITAZIONE

La banca-dati FIM ha consentito di costruire un sistema classificativo dei ricoveri riabilitativi. All'atto del ricovero il caso viene inserito in una categoria diagnostica molto generale (ictus, condizione ortopedica, ecc.). Quindi, sulla base del punteggio FIM e — per alcune categorie diagnostiche — sulla base dell'età, è possibile inserire il caso in un "gruppo funzionale" ("Function related group", FRG) per il quale si può prevedere con notevole precisio-

ne il tempo di degenza, così come il punteggio FIM di dimissione. Il principio di identificare gruppi "isorisorse" è lo stesso che informa il sistema DRG (Diagnosis Related Groups) in base al quale sono retribuiti, negli Stati Uniti come in Italia, gli ospedali per pazienti acuti. Lo sviluppo di "FRG" basati sulla scala FIM è stato guidato da Margaret G. Stineman, medico fisiatra, professore di Medicina Riabilitativa presso la Pennsylvania State University (abbreviata in Penn State) di Filadelfia, MD-USA. Anche la Stineman è un ricercatore di grandissimo rilievo internazionale. Il sistema di "grouping" statunitense è ora proprietà della Penn State, che ha accordi di licenza con la SUNY. Quindi il sistema FIM-FRG americano nel suo complesso è protetto da un doppio Copyright.

Perché sviluppare un sistema FRG visto che esiste già un sistema DRG? Entrambi i sistemi mirano ad una classificazione che consenta di prevedere il consumo di risorse. La classificazione secondo DRG, tuttavia, si basa sulla diagnosi di patologia (quella principale, ed eventuali comorbidità o complicazioni), e predice circa 30% della varianza dei costi sostenuti dall'ospedale per quel dato tipo di pazienti. Quindi il sistema DRG non considera direttamente le condizioni funzionali del paziente. Per esempio la codifica DRG può dirci se un paziente presenta emiplegia ed anche se ha presentato gravi complicazioni durante il ricovero in fase acuta. Questa informazione diviene insufficiente se lo scopo del ricovero è un progetto di riabilitazione. Infatti non si può sapere in quali condizioni funzionali (motorie, sfinteriche, cognitive) si trovasse il paziente al momento del ricovero. Di conseguenza, come è facilmente intuibile il sistema DRG non è efficace nel prevedere il consumo di risorse imposto dall'assistenza riabilitativa (circa il 12% di varianza spiegata). Il consumo, infatti, dipende in gran parte da costi fissi "alberghieri" molto simili per tutti i pazienti. Anche per la quota variabile il costo giornaliero è molto simile per tutti i pazienti (salvo il caso di craniolesi e mielolesi in fase post-acuta, che impegnano più risorse). Ciò che differenzia il costo del singolo caso è dunque la durata del ricovero, questa sì molto variabile. È intuitivo che la durata tenda ad essere prevedibile sulla base della disabilità all'ingresso. Un sistema tipo FIM-FRG, infatti, predice circa il 30% della varianza del tempo di degenza (e quindi, indirettamente, del consumo di risorse): questa è la stessa prestazione offerta dal sistema DRG sui ricoveri per pazienti acuti.

Inoltre il punteggio FIM costituisce anche a) un indice di appropriatezza: pazienti troppo o troppo poco disabili non sono idonei ad un progetto di riabilitazione

intensiva e b) di efficacia/efficienza: per un dato FRG mi attendo un certo tempo di degenza, ma anche un certo incremento di autosufficienza.

Infine, ragionare per FRG e non per diagnosi di malattia consente anche confronti multicentrici su una base omogenea: si paragoneranno, per esempio, non due casistiche di ictus cerebrale, ma due casistiche di pazienti emiparetici fra loro omogenei per livello di disabilità all'ingresso.

#### IL SISTEMA FIM NELLA RICERCA INTERNAZIONALE E ITALIANA

La scala FIM ormai è molto presente nella letteratura internazionale e sta diventando lo standard di misura più diffuso all'interno di protocolli di misura di appropriatezza e di efficacia/efficienza dei ricoveri riabilitativi e, più in generale, dei processi di cura e di assistenza rivolti a pazienti non autosufficienti.

Sta diventando molto difficile monitorare le moltissime ricerche che hanno utilizzato la scala FIM. Si è assunto questo compito lo stesso UDSMr, che mantiene una robusta rubrica bibliometrica sul proprio sito web [www.udsm.org](http://www.udsm.org). Il sito è facilmente visitabile e contiene anche molte informazioni sullo sviluppo del sistema FIM negli Stati Uniti.

La FIM è presente in Italia dal 1993, attraverso una società licenziataria in esclusiva (So.Ge.Com Editrice-Ricerca in Riabilitazione srl-Milano; e-mail: [sogecom@tin.it](mailto:sogecom@tin.it)). La Società conduce i corsi di accreditamento, definisce accordi di licenza sotto Copyright e offre servizi epidemiologico-statistici correlati alla scala FIM. In questi anni sono stati pubblicati moltissimi lavori relativi ad applicazioni FIM su casistica nazionale. La FIM si è dimostrata molto bene applicabile alla realtà sanitaria italiana nei suoi vari livelli assistenziali: ospedale per acuti, riabilitazione intensiva ed estensiva, casa di riposo, RSA, assistenza domiciliare.

In particolare sono stati realizzati o sono in corso di svolgimento anche studi di applicazione gestionale, compreso lo sviluppo di sistemi classificativi per gruppi funzionali. So.Ge.Com è soltanto il nodo italiano di una rete di licenziatari internazionali. Le licenze hanno lo scopo di assicurare l'uniformità dello strumento, in termini sia di validità trans-culturale della traduzione, sia di procedure di formazione degli operatori e quindi, in ultima analisi, di attribuzione dei punteggi. Questa forma di protezione dello strumento è stata la chiave per lo sviluppo della FIM come misura standard, e quindi come strumento ideale per confronti internazionali e per studi multicentrici nazionali. ■

banca-dati di riferimento (come, per la FIM, la banca-dati italiana e ancor più quella americana) consente di fornire "valori di ancoraggio" di difficoltà dei singoli items, ovvero stime pre-definite che costituiscano standard di riferimento. Questi standard sono molto precisi: di conseguenza diviene molto precisa anche la stima di una eventuale "incoerenza" dello schema di risposta di ogni soggetto. Il responsabile di uno studio che utilizzi questionari ha quindi a disposizione un valido strumento di controllo di qualità dei dati che gli pervengono. Misure "incoerenti" impongono procedure diagnostiche e correttive: per esempio una critica delle procedure di punteggio eseguite da un certo rilevatore. Talvolta, invece, la non-verosimiglianza corrisponde a procedure assistenziali atipiche che privilegiano o penalizzano certe attività del paziente rispetto ad altre.

**b) Applicazione alla scala: stima generalizzabile della riproducibilità delle sue misure**

Le misure che risultino convalidate con analisi di Rasch hanno una riproducibilità intrinseca notevole. Non è necessario confermarla con misure ripetute quali, per esempio, rilevazioni contemporanee ad opera di più osservatori oppure misure successive eseguite — assumendo come stabili le condizioni cliniche — alla dimissione da un reparto ed al contemporaneo ingresso in un altro. Infatti il modello di Rasch presenta in se stesso una grande generalizzabilità: se il paziente presenta punteggi ben centrati nei limiti di confidenza predetti per ciascun item, verosimilmente egli riprodurrà coerentemente lo stesso schema di risposta anche in qualsiasi altra circostanza.

Nel metodo di Rasch le misure ripetute sono comunque molto utili, ma vengono eseguite prevalentemente nella fase di costruzione e di validazione di una scala di misura e non — come nella psicometria classica — in ogni fase di applicazione per studiare la riproducibilità di un certo campione di soggetti e di osservatori. Infatti nel metodo di Rasch le misure ripetute servono a studiare una forma più sofisticata di riproducibilità: la stabilità intrinseca degli items. Con questo termine si intende la invarianza degli intervalli di difficoltà reciproca degli items, rispetto alle più diverse variabili indipendenti: tempo 0 vs. tempo 1, maschi vs. femmine, emiplegici verso paraplegici, questionari italiani verso questionari stranieri ecc. Il requisito che viene verificato è quello che la scala di misura mantenga sempre lo stesso significato: così come i numeri segnati lungo un righello devono restare allineati in un certo ordine e con certe distanze reciproche, quale che sia l'oggetto da misurare e quale che sia la sua lunghezza.

Per esempio, si possono confrontare due osservatori che attribuiscono ad un paziente un punteggio FIM totale identico ma composto in modo completamente diverso? Supponiamo che esso sia comprensivo di un 7 nell'alimentazione e di un 2 nel salire le scale in

un caso, e viceversa nell'altro caso. Analogamente: che cosa si può dire di una scala che presenti gerarchie di difficoltà sistematicamente diverse in pazienti emiplegici ed in pazienti paraplegici? Essa sta ancora misurando la stessa variabile nelle due categorie di pazienti? Questo tipo di analisi esplora quello che in gergo si chiama *item bias*, ovvero la suscettibilità degli items a interferenze sistematiche nel loro valore di difficoltà reciproca.

Infine, va ricordato che i test classici misurano l'accordo:

- a) fra i punteggi di singoli items o su punteggi cumulativi comunque questi ultimi siano composti, e non sull'intero schema di risposte
- b) su punteggi non intervallari. Per quanto essi siano ordinati gerarchicamente, questi punteggi rappresentano intervalli ignoti, e in più
- c) si assume anche che essi siano privi di errore (cioè che "2" sia veramente "2" e non qualche cosa che sarebbe potuto essere "1" o "3" in altre circostanze). Accordo e disaccordo, dunque, riguardano quantità definite in modo piuttosto arbitrario. Infine
- d) i test riguardano i singoli campioni di misure in esame (osservatore 1 verso osservatore 2, per esempio) così che rimane aperto il problema della generalizzabilità (L'accordo fra questi due osservatori non garantisce l'accordo fra altri due in altre circostanze).

L'analisi di Rasch, invece, mette a confronto:

- (a) la verosimiglianza di un intero schema di risposta
- (b) in termini di misure intervallari che — come si è già detto — sono anche
- (c) provviste di una stima di errore. Il tutto, viene posto a confronto con
- (d) le attese molto generalizzabili del modello.

**LA SCALA FIM:**

**RIPRODUCIBILITÀ E STABILITÀ**

Per quanto attiene la FIM esiste ormai una vastissima letteratura che conferma:

- a) riproducibilità "classica" (riproducibilità dei punteggi grezzi intra-inter osservatori con Cohen's K o ICCs)
- b) stabilità intesa come riproducibilità dei punteggi totali grezzi *cross-modality*: telefonica vs. intervista diretta; anamnestica vs. osservazione. La Letteratura si riferisce prevalentemente a dati americani (si veda al sito [www.ud-smr.org](http://www.ud-smr.org)) ma è stato concluso con esito del tutto favorevole anche uno studio italiano di riproducibilità intra-inter osservatori.
- c) riproducibilità e stabilità secondo analisi di Rasch.

In sintesi, dati americani e dati italiani concordano nel suggerire che la scala FIM mantenga sostanzialmente lo stesso significato se applicata, per

esempio, a pazienti emiplegici invece che protesizzati d'anca, oppure a pazienti in degenza riabilitativa ospedaliera invece che in casa di riposo. Procedure diagnostiche apposite consentono di identificare instabilità non dovute a errori di misura ma che indichino, in generale, particolari procedure assistenziali. Per esempio in casa di riposo l'autonomia sferica tende ad essere meno elevata rispetto all'autonomia locomotoria. Un motivo può essere rappresentato dall'obiettivo di rendere i pazienti liberamente deambulanti grazie al ricorso a presidi per l'incontinenza. Il contrario si osserva in riabilitazione ospedaliera, ove il paziente deambulante di solito non è o non è più incontinente. Una volta che sia stato scoperto, questo bias suggerisce di calcolare valori di riferimento specifici che consentano controlli di qualità specifici per la struttura cronico-geriatrica o per quella ospedaliera. La letteratura rivela anche che il rapporto fra punteggio FIM e minuti assistenziali è comunque simile in qualsiasi tipo di contesto assistenziale. Di conseguenza il confronto fra punteggi rimane possibile purché li si interpreti come indicatori di carico assistenziale complessivo, indipendentemente dal tipo di assistenza che ciascun punteggio comporta.

È facile comprendere come questo tipo di ragionamenti sia supportato bene dall'analisi di Rasch, e molto meno bene da misure ottenute con tecniche convenzionali.

**BIBLIOGRAFIA**

*Modello di Rasch e scale funzionali in riabilitazione*

- 1) WRIGHT BD, STONE MH  
*Best test design*. MESA Press, Chicago-IL 1979
- 2) ANDRICH D  
*Rasch models for measurement*. Sage University Papers Series on Quantitative Applications in the Social Sciences. Sage, Beverly Hills, CA 1988
- 3) TESIO L  
*Quality assessment of the FIM — Functional Independence Measure — ratings through Rasch analysis*. Eur Med Phys 1997; 33: 69-78
- 4) TESIO L, GRANGER CV, FIEDLER RC  
*A unidimensional pain/disability measure for low-back pain syndromes*. Pain 1997; 69: 269-278
- 5) PENTA M, THONNARD JL, TESIO L  
*ABILHAND: a Rasch-built measure of manual ability*. Arch Phys Med Rehabil 1998; 79: 1038-1042