LETTER TO THE EDITOR

VALIDITY OF FUNCTIONAL INDEPENDENCE MEASURE SCORES

Sir,

There is a growing and important controversy about the validity of Functional Independence Measure (FIM) scores. Ravaud et al. (1), based on the psychometric analyses of FIM data from 127 consecutive patients admitted to a medical rehabilitation unit in Paris, claim that "the total score should no longer be used to express a sole entity and that extreme caution should be applied when proposing its use as a mathematical parameter." Based on their analysis, the authors proposed the FIM instrument be reorganized to express four indicators: "self-care", "sphincter control", "social cognition and communication" and "overall body mobility". Their recommendations are supported by an earlier analysis from the United States that identified similar sub-dimensions within the motor-FIM in 14 of 20 diagnostic groups (2). However, we believe it is important to recognize that the reorganization proposed represents only one way, among many, that the FIM instrument can be expressed and that this may or may not be the best approach, depending on application.

No instrument intended to measure some global construct such as overall disability or care burden is "pure". Such instruments, by their nature, are intended to approximate some construct or series of constructs that are only inferred, as they cannot be measured directly.

Ravaud et al. (1) appear to be making the strong assertion that one can never use, for any purpose at all, an instrument that comprises anything but a set of "parallel measures" defined in the sense of classical test theory (3), quoting Silverstein et al. (4) as authoritative in this regard. However, this strong assertion ignores common scientific approaches that judge the utility of measurement procedures by their ability to predict meaningful outcomes. The economist Milton Friedman (5) asks: "What is the criterion by which to judge whether a particular departure from realism is or is not acceptable?" and then states "... a theory cannot be tested by comparing its 'assumptions' directly with 'reality.' Indeed, there is no meaningful way in which this can be done. Complete 'realism' is clearly unattainable, and the question whether a theory is realistic 'enough' can be settled only by seeing whether it yields predictions that are good enough for the purpose in hand or that are better than predictions from alternative theories."

The argument for parallel measures relates to the classical criteria applied toward the justification of summative measures (6). If a scale cannot be linear, then it becomes difficult to justify its use in parametric statistical applications such as linear regression. A summative model, according to Nunnaley (6), must meet three criteria: (a) the individual items must show a monotonic trace line (a plot of the probability of achieving certain scores on a test on the y-axis as a function of the magnitude of the latent trait measured); (b) the sum of the trace lines for the component items making up the composite score must be approximately linear; and (c) items as a whole can measure only the attribute in question. The total summed FIM measure was constructed to be related to care burden and has been shown to be correlated with minutes of care (7, 8). Nonetheless, we would not necessarily advocate the application of parametric statistical methods to the total FIM score. Such application would need to be evaluated with regard to the validity of statistical distributions and assumptions on a case-by-case basis (e.g., linear regression models with residual errors normally distributed and independent of expected values).

There is clear support for linearity of the motor and cognitive FIM dimensions that measure the attributes of physical and cognitive/communication disabilities, respectively. This support comes both from the discovery of those dimensions through Rasch analysis (9) as latent traits within the 18-item FIM instrument and through factor analytic studies designed to identify the two most important factors across 20 rehabilitation diagnostic groups. Support for the motor and cognitive domains as summative measures was obtained through multitrait analytic studies across 20 diagnostic groups using data on over 93,000 patients (10). When patients' performances on individual items are summed, the motor and cognitive domains have been shown to have comparable psychometric qualities to the Medical
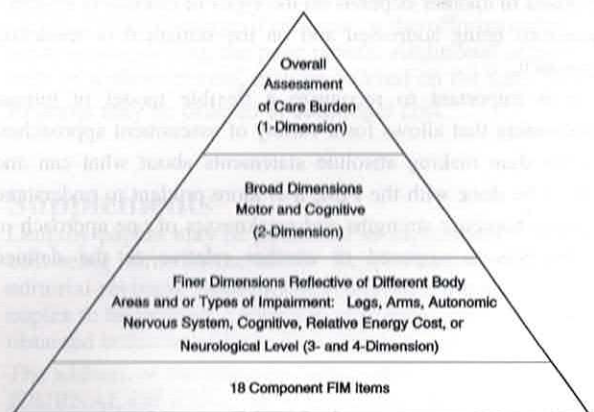

Fig. 1. The FIM hierarchy. Reprinted with permission.

Outcomes Study Short-Form 36 (11, 12), which has become a standard health status measure throughout the United States for measuring the health status of patients with chronic disease. The separate motor and cognitive FIM subscales distinguish physical from cognitive and communicative abilities.

Work by Ravaud et al. (1) provides further support for the FIM construct that has been referred to in the literature as impairment-specific dimensions (10). These subscales represent a finer resolution of disability, as they involve different functional areas of the body and form sub-dimensions within the broader motor-FIM. The functional areas include upper body, lower body and sphincter dysfunction, all of which will be affected differently by different impairments. Knowledge of disability profiles presumably measurable by the impairment-specific dimensions should be more clinically informative than motor and cognitive subscales with regard to the physiological and psychosocial mechanisms that underlie disability and even drive the rehabilitative process. Treatment decisions depend upon whether physical disability relates to problems in ADL, mobility, sphincter management or in all domains. Nevertheless, it is an empirical question as to whether this additional level of detail results in important increases in predictive or prognostic power. For example, we found that prediction of length of stay in the inpatient rehabilitation setting by a set of FIM subscales similar to those proposed by Ravaud et al. (1) resulted in a reduction in predictive power compared with the much more parsimonious representation of FIM as the two motor and cognitive subscales reflective of physical and cognitive disability (13). Although this is a compelling application, like all approaches its use, compared with other expressions, will most likely prove to have both benefits and limitations.

The FIM instrument is a multilayered and multidimensional structure that theoretically provides a number of alternative ways to express an individual's limitation in activities (Fig. 1). The multilayered approach is consistent with other instruments such as the Sickness Impact Profile (14), which can be presented either as profiles of scores or as a single aggregate score.

Each mode of FIM expression has particular applications and makes certain assumptions, rendering it valid or invalid for particular applications. The decision to use one FIM layer as opposed to another depends on the types of clinical or research questions being addressed and on the statistical or modeling approach.

It is important to recognize a flexible model of human functioning that allows for a variety of assessment approaches. Rather than making absolute statements about what can and cannot be done with the FIM, it is more prudent to understand the psychometric strengths and weaknesses of one approach or application as opposed to another relative to the defined phenomenon being measured and to the purpose for which measurements are obtained. We write to offer a challenge to the rehabilitation community and particularly to psychometricians. The challenge is to reappraise the fundamental aspects of multidimensional measurements such as the FIM instrument and how they might be applied differently, depending on the measurement needed.

REFERENCES

1. Ravaud JF, Delcey M, Yelnik A. Construct validity of the Functional Independence Measure (FIM): questioning the unidimensionality of the scale and the "value" of FIM scores. Scand J Rehabil Med 1999; 31: 31–41.
2. Stineman MG, Shea J, Jette AM, Tassoni CJ, Ottenbacher K, Fiedler RC, Granger CV. The Functional Independence Measure: Tests of scaling assumptions, structure, and reliability across 20 diverse impairment categories. Arch Phys Med Rehabil 1996; 77: 1101–1108.
3. Carmines EG, Zeller RA. Reliability and validity assessment (Series: quantitative applications in the social sciences). Newbury Park, CA: Sage Publications, 1979.
4. Silverstein B, Fisher WP, Kilgore KM, Harley JP, Harvey RF. Applying psychometric criteria to functional assessment in medical rehabilitation: II. Defining interval measures. Arch Phys Med Rehabil 1992; 73: 507–518.
5. Friedman M. The methodology of positive economics. In: Essays in positive economics. Chicago, IL: University of Chicago Press, 1953.
6. Nunnally JC. Psychometric theory. 2nd ed., New York: McGraw-Hill, 1978.
7. Disler PB, Roy CW, Smith PB. Predicting hours of care needed. Arch Phys Med Rehabil 1993; 74: 139–143.
8. Hamilton BB, Deutsch A, Russell C, Fiedler RC, Granger CV. Relation of disability costs to function: Spinal cord injury. Arch Phys Med Rehabil 1999; 80: 385–391.
9. Linacre JM, Heinemann AW, Wright BD, Granger CV, Hamilton BB. The structure and stability of the Functional Independence Measure. Arch Phys Med Rehabil 1994; 75: 127–132.
10. Stineman MG, Jette AM, Fiedler RC, Granger CV. Impairment-specific dimensions within the Functional Independence Measure. Arch Phys Med Rehabil 1997; 78: 636–643.
11. McHorney CA, Ware JE. Construction and validation of an alternative form general mental health scale for the Medical Outcomes Study Short-Form 36 Item Health Survey. Med Care 1995; 33: 15–28.
12. Stewart AL, Hays RD, Ware JE. The MOS short-form general health survey: reliability and validity in a patient population. Med Care 1988; 26: 724–732.
13. Stineman MG, Hamilton BB, Granger CV, Goin JE, Escarce JJ, Williams SV. Four methods for characterizing disability in the formation of function related groups. Arch Phys Med Rehabil 1994; 75: 1277–1283.
14. Bergner M, Bobbit RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. Med Care 1981; 19: 787–805.

Margaret G. Stineman[1] and Greg Maislin[2]
From the [1]Department of Rehabilitation Medicine, University of Pennsylvania, 101 Ralston-Penn Center, 3615 Chestnut Street, Philadelphia, PA 19104, USA, and [2]Department of Medicine, University of Pennsylvania, 3400 Spruce Street, Philadelphia, PA 19104, USA