

Clinimetrics: the science of clinical measurements

G. A. Fava,^{1,2} E. Tomba,¹ N. Sonino^{2,3}

SUMMARY

'Clinimetrics' is the term introduced by Alvan R. Feinstein in the early 1980s to indicate a domain concerned with indexes, rating scales and other expressions that are used to describe or measure symptoms, physical signs and other clinical phenomena. Clinimetrics has a set of rules that govern the structure of indexes, the choice of component variables, the evaluation of consistency, validity and responsiveness. This review illustrates how clinimetrics may help expanding the narrow range of information that is currently used in clinical science. It will focus on characteristics and types of clinimetric indexes and their current use. The clinimetric perspective provides an intellectual home for clinical judgment, whose implementation is likely to improve outcomes both in clinical research and practice.

The term 'clinimetrics' was introduced by Alvan R. Feinstein in 1982 to indicate a domain concerned with indexes, rating scales and other expressions that are used to describe or measure symptoms, physical signs and other distinctly clinical phenomena (1,2). The customary taxonomy does not include patterns of symptoms, severity of illness, effects of comorbid conditions, timing of phenomena, rate of progression of illness, functional capacity and other clinical features that demarcate major prognostic and therapeutic differences among patients who otherwise seem deceptively similar, because they have the same diagnosis and laboratory results (2). Well-known examples of clinimetric indexes are Jones criteria for rheumatic fever (1), the New York Heart Association Functional Classification (3) and Apgar's method of scoring the newborn's condition (4). Clinimetrics has a set of rules that govern the structure of indexes, the choice of component variables, the evaluation of consistency and validity (5). There has been an enormous redundancy in instruments measuring clinical variables in the past two decades (6). It is thus important to know the different characteristics of the available indexes to determine which are the most suitable and reliable methods for specific purposes.

Recent awareness of the inadequacy of disease as the primary focus of medical care makes the clinimetric challenge of the utmost importance. Tinetti

and Fried (7) have observed that, when disease became the focus of medicine in the past two centuries, the average life expectation was 47 years and most clinical encounters were for acute illness. The changed spectrum of health conditions (shifted towards ageing and chronicity) and the interindividual variabilities in health priorities suggest that the aim of treatment should refer to personal goals, that may range from attainment of cure to prevention of recurrence, from removal of functional impairment to alleviation of symptoms. This requires identification of all modifiable biological and non-biological factors involved (7). Disease-specific guidelines provide very limited indicators for patients with multiple conditions (8). Furthermore, there are medically unexplained symptoms that occur in up to 30–40% of medical patients and increase medical utilisation costs (9). With this regard, it has been suggested that it is not that certain disorders lack an organic explanation; it is our assessment that is inadequate in most of the clinical encounters (9). In fields such as neurology and geriatrics, the main challenge is to classify disability rather than disease. Accordingly, a number of clinimetric instruments have been developed for measuring pain (10), muscle functioning (11), gait and balance (12), dyskinesia (13) and delirium (14). Measurement of symptoms (15) and a multiaxial nosography (DSM-IV) (16) have become the mainstay of psychiatric assessment.

Review Criteria

The Authors performed a Medline search of the English language articles concerned with the keyword 'clinimetric' from 1982 to July 2011. It was supplemented by a manual search of the literature. The most representative contributions in terms of clinical implications were selected.

Message for the Clinic

Clinimetrics may help expanding the narrow range of information that is currently used in clinical science and provides an intellectual home for clinical judgment, whose implementation is likely to improve outcomes both in clinical research and practice.

¹Laboratory of Psychosomatics and Clinimetrics, Department of Psychology, University of Bologna, Bologna, Italy
²Department of Psychiatry, State University of New York at Buffalo, Buffalo, New York
³Department of Statistical Sciences, University of Padova, Padova, Italy

Correspondence to:

G. A. Fava, MD,
 Department of Psychology,
 University of Bologna, Viale
 Berti Pichat 5, 40127 Bologna,
 Italy
 Tel.: ++ 390512091339
 Fax: ++ 39051243086
 Email: giovanniandrea.
 fava@unibo.it

Disclosure

The authors have no financial interests to disclose.

The emerging needs for a more complete and satisfactory assessment underscore the importance of clinimetrics as a unifying perspective. In clinical medicine, there is the tendency to rely exclusively on 'hard data', preferably expressed in the dimensional numbers of laboratory measurements, excluding 'soft information' such as impairment, distress and well-being (17). This soft information, however, may now be reliably assessed by clinimetric methods. The aim of this review was to illustrate how clinimetrics may help expanding the narrow range of information that is currently used in clinical science, with main focus on characteristics and types of clinimetric indexes, their current use and their role in adjusting the clinical assessment to the 21st century medicine needs.

Characteristics of clinimetric indexes

Clinimetrics is concerned with the quality of clinical measurements, that includes both standardisation and sensibility (5,6). For standardisation, we want an index to have reliability and validity. Reliability often has an external part (observer variability in using the index, such as inter-rater agreement) and an internal part (consistency). Agreement between raters on categorical variables is often reported using kappa, that represents agreement corrected for chance (5). The problem of observer variability occurs in virtually all medical fields. For instance, considerable variation was found in interpretation of mammograms (18), assessment of physical signs in respiratory disease (19), histopathological diagnosis of melanoma (20) and ultrasound estimation of fetal weight (21). Validity, on the other hand, reflects the accuracy with which the phenomenon under observation is measured with a standard reference procedure, or gold standard (5). An index is valid if it measures what it is supposed to measure.

Feinstein (5) formulated the concept of sensibility for evaluating an index, with the specifications outlined in Table 1. An essential requisite for sensibility of an index is its discrimination properties (responsiveness/sensitivity), which means that it should be able to detect clinically relevant changes in health status over time (22,23). Sensibility, thus appears an important concept for developing and selecting indexes to monitor changes in clinical trials. As important is the clinimetric concept of incremental validity, that refers to the unique contribution (or incremental increase) in predictive power associated with a particular assessment procedure in the clinical decision process (24). Accordingly, each distinct aspect of measurement should deliver an unique increase in information to qualify for inclusion. In clinical research, several scales are often used under

Table 1 Features determining the sensibility of an index

- Purpose and framework (clinical functions, justification and applicability)
- Comprehensibility
- Replicability (e.g. clarity of instructions)
- Discrimination properties (responsiveness/sensitivity)
- Face validity (biological and clinical coherence of components)
- Content validity (comprehensiveness, weighting of components, quality of basic data)
- Ease of practical application

the misguided assumption that nothing will be missed. On the contrary, violation of the concept of incremental validity leads to conflicting results (23). For instance, in psychometrics, the same properties that give a scale a high score for homogeneity may obscure its ability to detect change, and redundant scale items may increase homogeneity, but decrease its sensitivity (25). A high correlation is often regarded as evidence that the two scales measure the same factor. However, a high correlation does not indicate similar sensitivity: a common content of two scales may ensure a high positive correlation between them, but the items they do not share may be important in determining their sensitivity (23). In clinimetrics, homogeneity of components is not requested and single items may be weighed in different ways: what matters is the capacity of an index to discriminate between different groups of subjects and to reflect changes in experimental settings, such as drug trials (15,23,26).

Clinimetric principles should guide the selection of methods to be used for a specific assessment and the modalities in which the assessment unfolds.

Clinimetric methods

Clinimetric indexes are arbitrary ratings for the various clinical phenomena that are observed or experienced and cannot be expressed in dimensional numbers (5). The structure, choice of variables, and organisation of output scales may greatly vary. The use that indexes may have in clinical medicine may range from diagnostic criteria of disease (determination of status) and instructions for decisions that lead to diagnostic and therapeutic actions (guidelines) to description of alterations recorded in repeated ratings (monitoring of change) and prognostic estimation (prediction). The same index may have more than one role. For instance a staging system for cancer may be used to rate the condition of the patient as well as to indicate a prognostic estimation (5).

A unique feature of clinimetric tools is to provide a broad global rating of clinical phenomena. Although the sensitivity of these methods is acknowledged in drug trials, where they often yield the most sensitive discrimination between drug and placebo effects (15), their value in clinical practice is currently underestimated. Another key characteristic is the collaboration of the patient. This is particularly true in the psychosocial domain, where many observer and self-rating scales have been developed. Although observer-rated methods make full use of the clinical experience and comparison potential of the interviewer, self-rating methods allow a more direct assessment of the patient subjective perceptions, such as quality of life.

Engel (27) criticised the attitude that tends to restrict what is categorised as disease to what the physician does understand and recognise, and he/she believes can be helped by his/her intervention. The fact that physicians arbitrarily exclude certain categories of complaints or signs as not pertinent reflects their social and institutional roles, that may vary with time and circumstances (27,28). According to Feinstein (2), this restricted frame of mind and the paucity of attention to inherently clinical phenomena stand in sharp contrast with the advanced state of methods (e.g. imaging, laboratory measurements), that do not derive from clinical examination and interaction with the patient. Engel identified the key characteristic of clinical science in its explicit attention to humanness, where observation (outer-viewing), introspection (inner-viewing) and dialogue (inter-viewing) are the basic methodological triad for clinical assessment and for making patient data scientific (28). Clinimetrics addresses also the various components of the patients–physician relationship (29), that are essential for shared decision (30) and self-management (31).

Clinimetric domains

Clinimetric indexes can be divided into ailment-oriented and general (5). Ailment-oriented indexes refer to specific diseases, states and clinical manifestations. They can produce diagnostic criteria for a particular disease or describe symptoms, temporal occurrence of sign and symptoms and of associated disturbances (comorbidity). General indexes refer to general health and functional states that are not distinctive for a particular disease or condition (5). The amount of clinical information that general indexes may yield is considerable, with particular reference to the psychosocial aspects of medical care (32). Some important domains are illustrated here.

- Cumulative exposures to stressful experiences, often subsumed under the term of ‘allostatic load’, may increase vulnerability to disease onset (33,34). Clinimetric instruments based on structured methods of data collection have been developed and tested in controlled studies: they include life events inventories for adults (35) and children (36), clinical measures of allostatic load (37,38) and questionnaires for the evaluation of workplace environmental stress (39).

- An increasing body of evidence links the initiation and progression of several medical disorders, such as diabetes, obesity and cardiovascular illness to lifestyle and behaviour (40,41). The benefits of modifying lifestyles were demonstrated in several randomised controlled trials, that measured health habits such as physical activity, exercise and food intake (42). Clinimetric tools for assessing alcohol consumption (43), nicotine dependence (44) and sleep quality (45) have also been developed and validated.

- Positive health is often regarded as the absence of illness, despite the fact that half a century ago the World Health Organization defined health as a ‘state of complete physical, mental and social well-being and not merely the absence of disease or infirmity’ (46). Several studies have suggested that psychological well-being plays a buffering role in coping with stress and has a favourable impact on disease course (47–49). For measuring psychological well-being, there are several clinimetric scales available (49–52).

- Once the symptoms of a medical disease are perceived by a person, or he/she has been told by a doctor that he/she is ill even if symptoms are absent, then this disease related information does influence the patient’s experience and behaviour, as well as the course and therapeutic outcome of a given illness episode (53). The study of illness behaviour (the ways in which individuals experience, perceive, evaluate and respond to their own health status) has yielded important information in medical patients (54). There are many manifestations of abnormal illness behaviour, that range from hypochondriasis and disease phobia to illness denial and lack of compliance (55). Different clinimetric strategies for assessing illness behaviour have been developed (55–57).

- Medical disorders are frequently associated with psychological symptoms, such as depression and anxiety (32,53). There is a substantial body of evidence that also psychological symptoms which do not reach the threshold of a psychiatric disorder may affect quality of life and social functioning and lead to increased health care utilisation (32). Depressed affect may have an impact on morbidity, mortality, compliance and increase susceptibility to medical illness (58). Similarly, the relationship between anxiety disorders and medical illness has been found to

entail important clinical implications (59). Other psychological correlates of clinical relevance are irritable mood (particularly as a risk factor in cardiovascular medicine) and demoralisation (55). Many observer and self-rated methods of assessment of affective states have been validated by clinimetric criteria (26,37,50,55,60–62). It should be noted that in many cases the inter-observer agreement obtained with these scales is far higher than that achieved by other forms of medical assessment (63).

Other domains that may be explored by general indexes include quality of life (64), physical activity (65), functional disability (66,67), comorbidity (68) and recovery (69).

Remodelling the clinical assessment

The need to include clinimetric consideration of function in daily life, performance of social roles, intellectual capacity, emotional stability, well-being, lifestyle and stress, has emerged as a crucial part of clinical investigation and patient care (7,9,70). Feinstein urged clinicians to develop a 'basic science' of their own 'to study clinical phenomena directly, to specify the importance of different types of clinical data, to improve the scientific quality of the data, to identify (or create) appropriate systems of taxonomy for classifying the information, and to develop intellectual models and pragmatic methods that would articulate the clinical process, recapitulate it, and use the results for quantified analyses' (70, p. 800). Yet, such 'basic' scientific challenges have been generally overlooked in the past decades, where emphasis has been placed on models derived from basic sciences (17). Psychosocial aspects have become particularly important in the setting of chronic diseases, where cure cannot take place, and also extend over family caregivers of chronically ill patients and health pro-

viders. Both in primary care and speciality fields, the traditional morphologic constraints (e.g. cardiology, gastroenterology) appear to be more and more inadequate in dealing with symptoms and problems that cut across organ system subdivisions. These inadequacies appear to be particularly pronounced in the large proportion of patients who lack an organic explanation for their disturbances. Furthermore, there is increasing need of expanding the concept of recovery merely based on laboratory and/or imaging parameters to consideration of the level of functioning in daily life. Patients, for instance, may display a normalisation of hormone parameters and yet feel impairment and distress (9,71). These issues are well known in rehabilitation medicine (72). Once again clinimetric methods are crucial in the full assessment of recovery. Finally, there is recent emphasis on patient-reported outcomes, any report coming directly from patients, without interpretation of physicians or others, about how they function or feel in relation to a health condition or its therapy (73,74). This area of research may benefit from the insights gained in clinimetric research (75), with special reference to psychosocial domains.

Conclusion

The clinimetric perspective not only allows clinicians to make full use of the potential of the clinical information that is already available; it also answers the increasing needs of expanding the content of customary clinical information, by including evaluation of variables such as stress, lifestyle, well-being, illness behaviour, psychological symptoms and perceptions. Clinimetrics provides an intellectual home for clinical judgment, whose implementation is likely to improve outcomes both in clinical research and practice.

References

- 1 Feinstein AR. The Jones criteria and the challenge of clinimetrics. *Circulation* 1982; **66**: 1–5.
- 2 Feinstein AR. An additional science for clinical medicine: IV. The development of clinimetrics. *Ann Intern Med* 1983; **99**: 843–8.
- 3 The Criteria Committee of the New York Heart Association. *Disease of the Heart and Blood Vessels*, 6th edn. Boston: Little Brown, 1964.
- 4 Feinstein AR. Multi-item 'instruments' vs. Virginia Apgar's principles of clinimetrics. *Arch Intern Med* 1999; **159**: 125–8.
- 5 Feinstein AR. *Clinimetrics*. New Haven, CT: Yale University Press, 1987.
- 6 De Vet HCW, Terwee CB, Bouter LM. Current challenges in clinimetrics. *J Clin Epidemiol* 2003; **56**: 1137–41.
- 7 Tinetti ME, Fried T. The end of the disease era. *Am J Med* 2004; **116**: 179–85.
- 8 Tinetti ME, Bogardus ST, Agostini JV. Potential pitfalls of disease specific guidelines for patients with multiple conditions. *N Engl J Med* 2004; **351**: 2870–4.
- 9 Fava GA, Sonino N. Psychosomatic assessment. *Psychother Psychosom* 2009; **78**: 333–41.
- 10 da Cunha Meneres Cost L, Maher CG, McAuley JH, Pena Costa LO. Systematic review of cross-cultural adaptations of McGill Pain Questionnaire reveals a paucity of clinimetric testing. *J Clin Epidemiol* 2009; **62**: 934–43.
- 11 de Koning CH, van der Henvel SP, Staal JB, Smits-Engelsman BC, Hendricks EJ. Clinimetric evaluation of methods to measure muscle functioning in patients with non-specific neck pain. *BMC Musculoskel Disord* 2008; **9**: 142.
- 12 Barker AL, Nitz JC, Low Choy NL, Haines T. Measuring fall risk and predicting who will fall. *J Gerontology A* 2005; **64**: 916–24.
- 13 Goets CG, Nutt JG, Stebbing GT. The Unified Dyskinesia Rating Scale: presentation and clinimetric profile. *Mov Disord* 2008; **23**: 2398–403.
- 14 Wong CL, Holroyd-Leduc J, Simel DL, Straus SE. Does this patient have delirium? Value of bedside instruments. *JAMA* 2010; **304**: 779–86.
- 15 Fava GA, Rafanelli C, Tomba E. The clinical process in psychiatry. A clinimetric approach. *J Clin Psychiatry* 2011; doi:10.4088/JCP.10r06444.

- 16 American Psychiatric Association. *Diagnostic Statistical Manual of Mental Disorders (DSM-IV)*, 4th edn. Washington: APA, 1994.
- 17 Feinstein AR. The intellectual crisis in clinical science. *Persp Biol Med* 1987; **30**: 215–30.
- 18 Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability in radiologists' interpretation of mammograms. *N Engl J Med* 1994; **331**: 1493–9.
- 19 Smyllie HC, Blendis LM, Armitage P. The observer disagreement in physical signs of the respiratory system. *Lancet* 1965; **2**: 412–3.
- 20 Shoo BA, Sagebiel RW, Kashani-Sabet M. Discordance in the histopathologic diagnosis of melanoma at a melanoma referral centre. *J Am Acad Dermatol* 2010; **62**: 751–6.
- 21 Dudley NJ. A systematic review in the ultrasound estimation of fetal weight. *Ultrasound Ob Gynecol* 2005; **25**: 80–9.
- 22 Husted JA, Cook RJ, Farewell VT, Gladmann DD. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000; **53**: 459–68.
- 23 Fava GA, Belaise C. A discussion on the role of clinimetrics and the misleading effects of psychometric theory. *J Clin Epidemiol* 2005; **58**: 753–6.
- 24 Sechrest L. Incremental validity: a recommendation. *Educ Psychol Meas* 1963; **23**: 153–8.
- 25 Wright JG, Feinstein AR. A comparative contrast of clinimetrics and psychometric methods for constructing indexes and rating scales. *J Clin Epidemiol* 1992; **45**: 1201–18.
- 26 Bech P. Fifty years with the Hamilton Scales for anxiety and depression. *Psychother Psychosom* 2009; **78**: 202–11.
- 27 Engel GL. The biopsychosocial model and medical education. *N Engl J Med* 1982; **306**: 802–5.
- 28 Engel GL. From biomedical to biopsychosocial. *Psychother Psychosom* 1997; **66**: 57–62.
- 29 Sledge WH, Feinstein AR. A clinimetric approach to the components of the patient–physician relationship. *JAMA* 1997; **278**: 2043–8.
- 30 Joosten EA, De Fuentes-Merillas L, de Weert GH, Sensky T, van der Staak CP, de Jong CA. Systemic review of the effects of shared decision-making on patient satisfaction, treatment adherence and health status. *Psychother Psychosom* 2008; **77**: 219–26.
- 31 Bodenheimer T, Lorig K, Holman H, Grumbach K. Patient self-management of chronic disease in primary care. *JAMA* 2002; **288**: 2469–75.
- 32 Fava GA, Sonino N. Psychosomatic medicine. *Int J Clin Practice* 2010; **64**: 1155–61.
- 33 McEwen BS. Protective and damaging effects of stress mediators. *N Engl J Med* 1998; **338**: 171–9.
- 34 Shonkoff JP, Boyce WT, McEwen BS. Neuroscience, molecular biology, and the childhood roots of health disparities. *JAMA* 2009; **301**: 2252–9.
- 35 Paykel ES. The interview for recent life events. *Psychol Med* 1997; **27**: 301–10.
- 36 Coddington RD. The significance of life events as etiological factors in the diseases of children, I: a survey of professional workers. *J Psychosom Res* 1972; **16**: 7–18.
- 37 Sonino N, Fava GA. A simple instrument for assessing stress in clinical practice. *Postgrad Med J* 1998; **74**: 408–10.
- 38 Fava GA, Guidi J, Semprini F, Tomba E, Sonino N. Clinical assessment of allostatic load and clinimetric criteria. *Psychother Psychosom* 2010; **79**: 280–4.
- 39 Karasek RA, Brisson C, Kawakami N, Houtman I, Bongers B, Amick B. The Job Content Questionnaire (JCQ): an instrument for internationally comparative assessments of psychosocial job characteristics. *J Occup Health Psychol* 1998; **3**: 332–55.
- 40 Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual causes of death in the United States, 2000. *JAMA* 2004; **291**: 1238–45.
- 41 Roger VL. Lifestyle and cardiovascular health. *JAMA* 2009; **302**: 437–9.
- 42 Tomba E. Assessment of lifestyle in relation to health. *Adv Psychosom Med* 2011; **32**: 72–96.
- 43 Barbor TF, Bohn MJ, Kranzler HR. The alcohol use disorders identification test (AUDIT): validation of a screening instrument for use in medical settings. *J Stud Alcohol* 1995; **56**: 423–32.
- 44 Fagerström KO, Schneider NG. Measuring nicotine dependence: a review of the Fagerstrom Tolerance Questionnaire. *J Behav Med* 1989; **12**: 159–82.
- 45 Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res* 1989; **28**: 193–213.
- 46 World Health Organization. *World Health Organization Constitution*. Geneva: WHO, 1948, p. 28.
- 47 Pressman SD, Cohen S. Does positive affect influence health? *Psychol Bull* 2005; **131**: 925–71.
- 48 Chida Y, Steptoe A. Positive psychological well-being and mortality. *Psychosom Med* 2008; **70**: 741–56.
- 49 Fava GA, Tomba E. Increasing psychological well-being and resilience by psychotherapeutic methods. *J Personality* 2009; **77**: 1903–34.
- 50 Kellner R. A symptom questionnaire. *J Clin Psychiatry* 1987; **48**: 269–74.
- 51 Antonovsky A. The structure and properties of the sense of coherence scale. *Social Sci Med* 1993; **36**: 725–33.
- 52 Ryff CD. Happiness is everything, or is it? *J Personal Soc Psychol* 1989; **6**: 1069–81.
- 53 Lipowski ZJ. Physical illness and psychopathology. *Int J Psychiat Med* 1974; **5**: 483–97.
- 54 Mechanic D. Social psychological factors affecting the presentation of bodily complaints. *N Engl J Med* 1972; **286**: 1132–9.
- 55 Porcelli P, Sonino N eds. *Psychological Factors Affecting Medical Conditions*. Basel: Karger, 2007.
- 56 Sirri L, Grandi S, Fava GA. The Illness Attitude Scales. A clinimetric index for assessing hypochondriac fears and beliefs. *Psychother Psychosom* 2008; **77**: 337–50.
- 57 Pilowsky I. *Abnormal Illness Behaviour*. Chichester: Wiley, 1997.
- 58 Katon WJ. Clinical and health service relationships between major depression, depressive symptoms, and general medical illness. *Biol Psychiatry* 2003; **54**: 216–26.
- 59 Fava GA, Porcelli P, Rafanelli C, Mangelli L, Grandi S. The spectrum of anxiety disorders in the medically ill. *J Clin Psychiatry* 2010; **71**: 910–4.
- 60 Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins symptom checklist (HSCL). *Mod Probl Pharmacopsychiatry* 1974; **7**: 79–110.
- 61 Zigmond AS, Snaith RP. The Hospital Anxiety and Depression scale. *Acta Psychiatr Scand* 1983; **67**: 361–70.
- 62 Guidi J, Fava GA, Bech P, Paykel ES. The Clinical Interview for Depression: a comprehensive review of studies and clinimetric properties. *Psychother Psychosom* 2011; **80**: 10–27.
- 63 Rush AJ, First MB, Blacker D. *Handbook of Psychiatric Measures*, 2nd edn. Washington, DC: American Psychiatric Press, 2008.
- 64 Testa MA, Simonson DC. Assessment of quality of life outcomes. *N Engl J Med* 1996; **334**: 835–40.
- 65 Gebruers N, Vauroy C, Truijen S, Engelborgh S, de Deyn PP. Monitoring physical activity after stroke. *Arch Phys Med Rehab* 2010; **91**: 288–97.
- 66 Feinstein AR, Josephy DR, Wells CK. Scientific and clinical problems in indexes of functional disabilities. *Ann Intern Med* 1986; **105**: 413–20.
- 67 Grill E, Stucki G. International classification of functioning disability and health. *J Clin Epidemiol* 2009; **62**: 891–8.
- 68 de Groot V, Beckerman H, Tankhorst GJ, Bouter LM. How to measure comorbidity: a critical review of available methods. *J Clin Epidemiol* 2003; **56**: 221–9.
- 69 Kluivers KB, Hendricks JC, Mol BW et al. Clinimetric perspectives of three instruments measuring post-operative recovery in a gynecological surgical population. *Surgery* 2008; **144**: 12–21.
- 70 Feinstein AR. Clinical judgment revisited: the distraction of quantitative models. *Ann Intern Med* 1994; **120**: 799–805.
- 71 Sonino N, Peruzzi P. A psychoneuroendocrinology service. *Psychother Psychosom* 2009; **78**: 346–51.
- 72 Wade DT, Smeets RJE, Werbunt JA. Research in rehabilitation medicine. Methodological challenges. *J Clin Epidemiol* 2010; **63**: 69–74.
- 73 Bottomley A, Jones D, Claassens L. Patient-reported outcomes: assessment and current perspectives of the guidelines of the Food and Drug Administration and the reflection paper of the European Medicine Agency. *Eur J Cancer* 2009; **45**: 347–53.
- 74 Clancy C, Collins FS. Patient-Center Outcomes Research Institute: the Intersection of Science and Health Care. *Sci Transl Med* 2010; **2**: 37cm18.
- 75 Concato J, Feinstein AR. Asking patients what they like: overlooked attributes of patients satisfaction with primary care. *Am J Med* 1997; **102**: 399–406.

Paper received August 2011, accepted October 2011