

Measuring clinical findings: the value of clinimetrics

Giovanni A. Fava^{1,*}, Nicoletta Sonino^{1,2}, Jenny Guidi^{3,4}

¹Department of Psychiatry, State University of New York at Buffalo, Erie County Medical Center, 462 Grider Street, Buffalo, NY 14215, United States

²Department of Statistical Sciences, University of Padova, via Battisti 241, 35121 Padova, Italy

³Department of Psychology “Renzo Canestrari”, University of Bologna, viale Berti Pichat 5, 40127 Bologna, Italy

⁴Massachusetts General Hospital, Clinical Trials Network and Institute, Boston, 1 Bowdoin Square, MA 02114, United States

*Corresponding author. Department of Psychiatry, Erie County Medical Center, 462 Grider St., Buffalo, NY 14215, United States.

E-mail: giovanniandrea.fava@unibo.it

Abstract

The term “clinimetrics” indicates a domain concerned with indexes, rating scales and other expressions that are used to describe or measure symptoms, physical signs and other clinical variables. Clinimetrics, the science of clinical measurements, has a set of rules that govern the structure of indexes and their consistency, validity and discrimination properties, as well as the choice of component variables. A unique feature of clinimetric indexes are broad global ratings of clinical manifestations that would otherwise be scattered, such as in the case of comorbidity (e.g. the Charlson Comorbidity Index). There have been major developments in the assessment of subjective health status and of psychosocial aspects. The use of indexes may improve the quality of data collection both in medical practice and clinical trials. Clinimetrics enlarges the scientific basis of clinical practice and provides an intellectual home for clinical judgment.

Bullet Points

- The term clinimetrics indicates a domain concerned with indexes, rating scales, and other measurements of clinical phenomena.
- Clinimetrics has a set of rules that help selecting the most suitable clinical measurements.
- Clinimetrics may improve the scientific quality of data and illness configuration.

Keywords: clinimetrics; measurement; comorbidity; patient-reported outcome measures; health attitudes

Introduction

In their everyday practice, clinicians integrate customary taxonomy with a wide range of evaluations that include variables such as patterns of symptoms, timing of phenomena, severity of illness, effects of comorbid conditions, rate of progression of illness, functional capacity, psychosocial impact of medical disorders and their treatments. This type of information may be reliably assessed by clinimetric methods. 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]. Examples of clinimetric indexes mentioned by Feinstein [1, 2] were 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]. In 1987 Feinstein published a monograph concerned with the rules that govern the structure of indexes, the choice of component variables, the evaluation of consistency and validity [5].

The concept of clinimetrics was far ahead of its time and not surprisingly had only a marginal impact for the first 20 years [6].

In the past two decades there has been an increasing appreciation of its role, due to a number of converging developments. There has been growing awareness on how a strictly biological model of disease that ignores a person’s life experience is insufficient to account for all the factors that make individuals susceptible to disease and variably responsive to treatment [7–9]. The changed spectrum of health conditions (shifted toward aging and chronicity) and the inter-individual variabilities in health priorities suggest that the aim of treatment should refer to personal goals [7]. Identification of all modifiable biological and non-biological factors involved is then required [6, 7]. The clinical problems of patients with multiple chronic conditions (often defined as complex patients) and their implications in terms of health care costs have attracted increasing consideration [8]. Patient empowerment and activation represent key aspects to improving health in case of significant comorbidity. However, such interventions should be grounded in the biopsychosocial context of an individual [9], prioritizing patient experiences and quality of life [10].

The aim of this critical review was to illustrate the state of the art of clinimetrics, the characteristics and types of clinimetric indexes, and their implications for clinical practice and research.

Received 30 January 2025; revised 17 April 2025; accepted 9 May 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of the Fellowship of Postgraduate Medicine.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Some innovative features of clinimetrics compared to its original formulations [1, 2, 5] have been outlined in a review of the field that was published in 2012 [11]. Subsequent clinimetric studies have yielded additional developments that deserve to be addressed. There has been a huge redundancy in instruments measuring clinical variables in the past two decades [6, 11, 12]. It is thus important to appreciate the characteristics of the available tools for specific purposes.

For this purpose, the Authors performed a Medline search of the English language articles concerned with the keyword “clinimetric*”, from 1982 to March 2025. Since such procedure was likely to identify only papers that made specific reference to clinimetrics and miss other investigations that were concerned with clinical measurements, it was supplemented by a manual search of the literature. The most relevant contributions in terms of clinical implications were selected.

What are the clinimetric measurements and what do they measure?

The structure, selection of variables, and organization of clinimetric measurements may vary a lot. The traditional use that indexes 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]. Clinimetric tools are not always called indexes: they may be designated as scales, scores, factors, stages, classes, ratings, systems or criteria. They may be applied to any clinical data that derive from observation, medical history, physical examination, laboratory, imaging, pathology, and functional analysis.

Clinimetrics often requires the joint collaboration of the patient, that particularly applies to the psychosocial domain, where both observer and self-rating scales have been developed for measuring the same clinical phenomena [6, 11]. Interviewing or observer-rated methods make full use of the clinical experience and of the discrimination skills of the interviewer. Clinical interpretation of symptoms according to diagnostic criteria has indeed become the mainstay of psychiatric assessment [13]. On the other hand, self-rating methods allow a more direct assessment of the patient’s subjective perceptions, such as quality of life. Patient Reported Outcome Measures (PROMs) are self-reported scales and indexes specifically designed to improve the detection of the patients’ subjective experience [14]. Originally designed to measure treatment outcomes, PROMs can be used in clinical practice as screening tools, to evaluate the severity, burden and impact of symptoms on quality of life and psychological well-being, and to assess functioning in daily life, productivity, and emotional stability. Clinimetrics addresses also the various components of the patient-physician relationship [15] that are essential for shared decision-making and self-management [16].

A unique feature of clinimetric tools is the possibility to provide a broad global rating of clinical phenomena, such as comorbidity [8]. Many clinimetric indexes express the global rating of a condition with a score [5, 11, 12]. The Apgar score pioneered the field [4]; other common examples are the Gleason score for prostate cancer [17] and the score for predicting risk of stroke in atrial fibrillation [18]. In fields such as neurology and geriatrics the main challenge is classifying disability rather than disease [6–8]. Thus, a number

Table 1. Clinimetric characteristics.

- | |
|---|
| 1) Standardization: reliability (inter-rater agreement, consistency) and validity |
| 2) Purpose and framework (clinical functions, justification and applicability) |
| 3) Comprehensibility |
| 4) Replicability (e.g. clarity of instructions) |
| 5) Ability to discriminate accurately |
| 6) Face validity (biological and clinical coherence of components) |
| 7) Content validity (comprehensiveness, weighting of components, quality of basic data) |
| 8) Ease of practical application |

of clinimetric instruments have been developed for measuring daily functional capacities [19, 20], pain [21], gait and balance [22, 23], and dyskinesia [24, 25].

How to evaluate measurements in clinimetrics

Clinimetrics can be evaluated by several criteria (Table 1). As to standardization, reliability and validity are considered the main requirements for an index. 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 the *kappa* coefficient, that represents agreement corrected for chance [5]. The problem of observer variability occurs in virtually all medical fields. For instance, in classic studies considerable variation was found in interpretation of mammograms [26], assessment of physical signs in respiratory disease [27], and histopathologic diagnosis of melanoma [28]. 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 is supposed to measure.

An essential requisite of a clinimetric measurement is its sensitivity, defined as its ability to discriminate between patients and controls and/or subgroups of patients and to reflect changes in treatment trials [29]. This property is often indicated as responsiveness [30]. However, responsiveness, if not associated with the ability of an index to differentiate among distinct populations or clinical subgroups (i.e. discrimination properties), may be affected by non-specific effects [31].

There are major differences between clinimetrics and the traditional psychometric approach [29, 31]. A basic psychometric assumption is the need for homogeneity of components of rating scales, as measured by statistical tests such as Cronbach’s *alpha* and factorial analysis [29, 31]. The redundant components that provide a scale with a high score for homogeneity are likely to obscure its capacity to record changes in clinical status. In the clinimetric approach, neither homogeneity of components nor unidimensionality are required, conforming to the heterogeneous features of clinical variables. Indeed, what matters is the capacity of an index to discriminate between different (sub)groups of individuals and to reflect changes in experimental settings such as drug trials [29]. Further, in classical psychometrics, all items of a rating scale have the same weight and measure parallel forms of the same symptom. Yet, in clinical practice (as reflected by the clinimetric approach) not all symptoms have the same weight (major and minor symptoms can be differentiated, such in Jones’ criteria for rheumatic fever) [1]. Finally, a high correlation is often

regarded as evidence that the two scales measure the same factor. However, a high correlation does not indicate similar discrimination properties: a common content of two scales may insure a high positive correlation between them, but the items they do not share may be important in determining their sensitivity [29]. As a result, even though the basic principles of clinimetrics and psychometrics may seem to be similar, the two domains have substantial differences in strategies and goals, that need to be differentiated in the description of an index.

What type of clinimetric indexes are available and how can they be used?

Clinimetric indexes may be differentiated 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 signs and symptoms and presence of associated disturbances (comorbidity). General indexes refer to general health and functional states that are not distinctive for a particular disease or condition [5]. Indexes may either be restricted to a specific clinical issue (e.g. pain) or attempt to synthesize different and heterogeneous manifestations of illness in an overall judgment of the person's condition. The amount of clinical information that general indexes may yield is considerable. Such global ratings (unitary concepts) allow to bring together clinical manifestations that would otherwise be scattered and to find a context that is more meaningful than what is provided by medical diagnoses [6]. For instance, the Charlson Comorbidity Index [32] is the most widely recognized clinimetric method for identifying the patient's unique clinical situation, with particular reference to prognostic variables. It not only indicates how comorbidity affects the outcome of the primary diagnosis [33, 34], but it is also a global evaluation of the overall severity of illness.

Examples of clinimetric unitary concepts are:

a) *frailty*, that connotes increased vulnerability to physiological and psychological stressors in older people [8, 35, 36]. The Clinical Frailty Scale provides an overall clinical picture that does not simply result from the sum of morbidities [35];

b) *allostatic load*, that refers to the cumulative burden of chronic stress and life events [37–39]. When environmental challenges exceed the individual ability to cope, allostatic overload ensues [38]. Clinimetric methods, based on clinical interview [38] or self-rated questionnaire [40] have been introduced;

c) *euthymia*, that indicates presence of positive affect, balanced levels in psychological well-being dimensions, and resilience [41]. It contributes to a favorable disease course and outcome. Both observer- and self-rated clinimetric methods for its assessment are available [41].

Clinimetrics is in line with the notion that health applies to the whole person and not to specific organs, has subjective dimensions (how an individual experiences it and how such experience is communicated to the physician), and has a positive and dynamic connotation in terms of functional ability [6, 11, 42]. Indeed, the subjective perception of health status is as valid as that of the clinician in evaluating outcomes [14, 43–45]. There may be important discrepancies between the clinical state of an individual according to objective measurements and the unique personal perception of health: individuals with severe medical diseases by objective parameters may maintain a high level of functioning, while disorders that are judged as minor

ailments by physicians may have devastating effects [46, 47]. Such discrepancies have a primary role in assessing recovery [48] and in rehabilitation medicine [49, 50].

Further, a number of clinimetric indexes has been developed to assess individual attitudes and behavior related to health and disease, that shape lifestyle, presentation of symptoms, access to patient care, interactions between patients and physicians, adherence to medical advice, response to treatment [51]. Health attitudes and behavior may range from anxiety and worry about illness to various forms of denial, such as delay of seeking care and lack of adherence to treatment. An impressive body of evidence has linked the progression of severe medical disorders to specific lifestyle behaviors [16]. Clinimetric indexes for assessing variables such as alcohol consumption, nicotine dependence, and sleep quality are available [52]. In particular, health behavior is likely to play a major role in the process of convalescence, in self-management of chronic conditions, in determining a state of recovery, and whenever a rehabilitation process is involved [51].

Clinimetric strategies are also useful to evaluate the balance of potential benefits and adverse effects of treatments applied to the individual patient [53]. There have been major clinimetric developments in the assessment and definition of adverse events, such as those that have occurred with psychotropic drugs [54]. The concept of iatrogenic comorbidity refers to the unfavorable modifications in the characteristics, course and response to treatment of an illness that may be related to previously administered therapies [54]. Vulnerabilities may manifest themselves during treatment administration and/or after its discontinuation. Research on attitudes toward medications has shed new light on the phenomena related to treatment adherence [55]. Clinimetric tools may thus help placing treatment options within the framework of potential benefits, likelihood of responsiveness, and vulnerability to the adverse effects of treatment [56]. Indexes for measuring medication appropriateness have been validated [57].

In the medical setting, psychiatric evaluation is primarily geared to identification of mental disorders according to diagnostic criteria, with particular reference to depressive and anxiety disorders [13]. However, it does not allow to characterize the individual psychosocial issues that affect illness course and management [46, 58]. Clinimetric indexes, based on clinical interview or questionnaires, may thus be helpful [46].

Implications for clinical trials

The clinimetric approach has a number of implications for clinical trials [59]. When no significant differences emerge, a factor that is frequently overlooked is that the selected measures may be valid and reliable, but they lack adequate discriminating properties. Such considerations particularly apply to studies concerned with small samples, or when differences are expected to be small [29].

As important is the clinimetric concept of incremental validity (each distinct aspect of measurement should deliver a unique increase in information in order to qualify for inclusion) [11, 60]. In clinical research, many scales are often used under the misguided assumption that nothing will be missed. On the contrary, violation of the principle of incremental validity is likely to lead to inconsistent results [11].

Another major issue in the interpretation of clinical trials is that treatment outcome is the cumulative result of different factors [61]. In fact, a selected treatment interacts with several classes of variables (globally labeled as non-specific): living conditions (e.g. housing, nutrition, work, environment, social

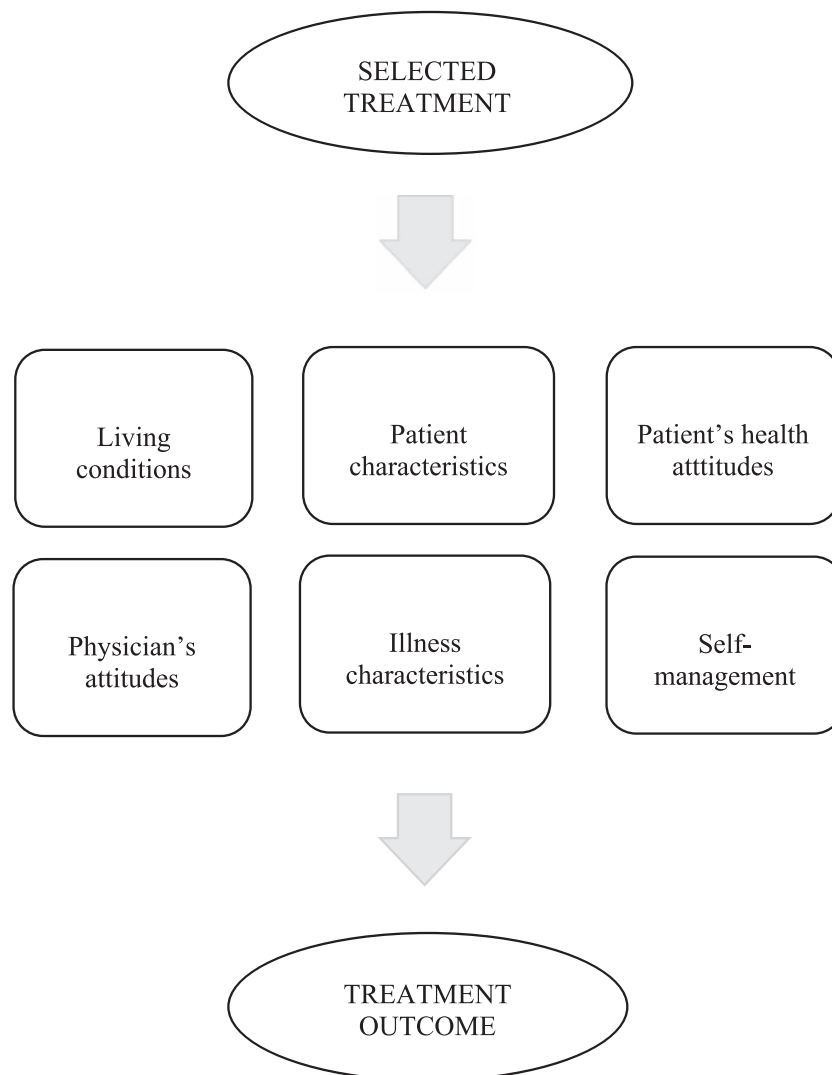


Figure 1 Interacting clinimetric variables that affect treatment outcome

support), patient characteristics (e.g. age, sex, genetics, general health conditions, personality, well-being), patient's health attitudes and behavior, treatment setting (e.g. physician's attitude and attention), illness features and previous treatment, level of self-management [61]. All these variables, that can be measured (and thus controlled for) with clinimetric indexes (Fig. 1), may bring therapeutic or counter-therapeutic effects. In some patients their interactive combination may lead to clinical improvement, whereas in other cases yields no effect or worsens the condition [61]. How a person experiences the treatment process, his/her interaction with the physician and significant others, his/her role in collaborating with the therapeutic plan (self-management), are all fundamental components of therapy viewed as a total human response.

Recently, Horwitz et al. [59] have suggested the use of composite clinical scores in clinical trials. Such scores may be calculated as a weighted sum of a patient's clinical features, comorbidity, and other characteristics that may affect individual susceptibility to adverse outcomes. Physicians commonly evaluate the risk for adverse outcomes and select treatment accordingly [56, 59]. The calculated score may indicate the clinical subgroups that might benefit most from a given treatment [59].

Conclusions

Clinimetrics, the science of clinical measurements, enlarges the scientific basis of clinical practice and expands the narrow range of information that is currently used in research. Selecting the most suitable clinimetric assessment strategies may demarcate major prognostic and therapeutic differences among patients who otherwise seem to be deceptively similar because they have the same diagnosis and laboratory results [2, 6, 11]. The concept of clinimetrics is closely related to the need for physicians to distinguish diagnosis from health [62], disease, illness, sickness and predicaments [63], absence of disease and well-being and wellness [41, 64, 65].

Clinical excellence involves both the art and the science of medicine: a comprehensive knowledge of presentations and courses of diseases, skills in data gathering, competency in communication, and considerable judgment in the use of information and problem solving [66]. Skills in data gathering are concerned with observation, clinical interviewing and physical examination. Clinimetrics calls for a renewed appraisal of clinical interviewing as the basic method to understand how a person feels and what are the presenting complaints, obtain medical history, evaluate personal attitudes and behavior related to health

and disease [67, 68]. Clinimetric indexes offer unprecedented opportunities for improving the quality of data collection, whether obtained by clinical interviews or questionnaires, and such quality is likely to affect assessment and treatment. The clinimetric perspective provides an intellectual home for the reproduction and standardization of clinical judgment [69]. It calls for a joint evaluation of biological, psychological and social factors that allow the clinician to make full use of all available information.

Author contributions

All Authors conceived and wrote the entire manuscript.

Conflict of interest statement: The Authors have no conflicts of interest to declare.

Funding

The Authors have no funding to declare.

Key references

5. Feinstein AR. *Clinimetrics*. New Haven, Yale University Press, 1987.
8. Charlson ME, Wells MT. Comorbidity: from a confounder in longitudinal clinical research to the main issue in population management. *Psychother Psychosom* 2022;**91**:145–151.
11. Fava GA, Tomba E, Sonino N. Clinimetrics: the science of clinical measurements. *Int J Clin Pract* 2012;**66**:11–15.
59. Horwitz RI, Baker JB, Ghatak A, et al. Randomization bias, multi-morbidity, and the Composite Clinical Score. *Psychother Psychosom* 2024;**93**:219–223.
68. Fava GA, Sonino N, Aron DC, et al. Clinical interviewing: an essential but neglected method of medicine. *Psychother Psychosom* 2024;**93**:94–99.

Self-assessment questions

- 1) Broad global ratings of clinical phenomena are a distinctive feature of clinimetrics.
- 2) Clinimetric indexes need to have homogeneous components.
- 3) The Charlson Comorbidity Index is an example of clinimetric measurement.
- 4) Both benefits and adverse effects of treatment can be assessed with clinimetric tools.
- 5) In clinimetrics it is important to have multiple types of evaluation of the same variable.

Answers:

- 1) True
- 2) False
- 3) True
- 4) True
- 5) False

Research questions

1. What is clinimetrics?
2. What are the rules of clinimetrics?
3. How can clinimetrics be used in research and practice?

References

1. Feinstein AR. T. Duckett Jones memorial lecture. The Jones criteria and the challenges of clinimetrics. *Circulation* 1982;**66**:1–5. <https://doi.org/10.1161/01.CIR.66.1.1>
2. Feinstein AR. An additional basic science for clinical medicine: IV. The development of clinimetrics. *Ann Intern Med* 1983;**99**:843–8. <https://doi.org/10.7326/0003-4819-99-6-843>
3. The Criteria Committee of the New York Heart Association. *Disease of the Heart and Blood Vessels*. 6th edn. Boston:Little Brown, 1964.
4. Apgar V. A proposal for a new method of evaluation of the newborn infant. *Curr Res Anesth Analg* 1953;**32**:260–7. <https://doi.org/10.1213/00000539-195301000-00041>
5. Feinstein AR. *Clinimetrics*. New Haven:Yale University Press, 1987.
6. Fava GA. Forty years of clinimetrics. *Psychother Psychosom* 2022;**91**:1–7. <https://doi.org/10.1159/000520251>
7. Tinetti ME, Fried T. The end of the disease era. *Am J Med* 2004;**116**:179–85. <https://doi.org/10.1016/j.amjmed.2003.09.031>
8. Charlson ME, Wells MT. Comorbidity: from a confounder in longitudinal clinical research to the main issue in population management. *Psychother Psychosom* 2022;**91**:145–51. <https://doi.org/10.1159/000521952>
9. Horwitz RI, Singer BH, Hayes-Conroy A et al. Biosocial pathogenesis. *Psychother Psychosom* 2022;**91**:73–7. <https://doi.org/10.1159/000521567>
10. Kalyani RR, Allende-Vigo MZ, Antinori-Lent KJ et al. Prioritizing patient experiences in the management of diabetes and its complications: an endocrine society position statement. *J Clin Endocrinol Metab* 2024;**109**:1155–78. <https://doi.org/10.1210/clinem/dgad745>
11. Fava GA, Tomba E, Sonino N. Clinimetrics: the science of clinical measurements. *Int J Clin Pract* 2012;**66**:11–5. <https://doi.org/10.1111/j.1742-1241.2011.02825.x>
12. Leff LE, Koperwas ML. Calculated medicine: seven decades of accelerating growth. *Am J Med* 2024;**137**:582–8. <https://doi.org/10.1016/j.amjmed.2024.03.025>
13. APA. *Diagnostic and Statistical Manual of Mental Disorders*. Fifth edn. DSM-5. Washington, DC: American Psychiatric Association, 2013
14. Carrozzino D, Patierno C, Guidi J et al. Clinimetric criteria for patient-reported outcome measures. *Psychother Psychosom* 2021;**90**:222–32. <https://doi.org/10.1159/000516599>
15. Sledge WH, Feinstein AR. A clinimetric approach to the components of the patient-physician relationship. *JAMA* 1997;**278**:2043–8. <https://doi.org/10.1001/jama.1997.03550230019007>
16. Kris-Etherton PM, Petersen KS, Després JP et al. Strategies for promotion of a healthy lifestyle in clinical settings: pillars of ideal cardiovascular health: a science advisory from the American Heart Association. *Circulation* 2021;**144**:e495–514. <https://doi.org/10.1161/CIR.0000000000001018>
17. Humphrey PA. Gleason grading and prognostic factors in carcinoma of the prostate. *Mod Pathol* 2004;**17**:292–306. <https://doi.org/10.1038/modpathol.3800054>
18. Lip GY, Halperin JL. Improving stroke risk stratification in atrial fibrillation. *Am J Med* 2010;**123**:484–8. <https://doi.org/10.1016/j.amjmed.2009.12.013>
19. Bouça-Machado R, Fernandes A, Ranzato C et al. Measurement tools to assess activities of daily living in patients with Parkinson's disease: a systematic review. *Front Neurosci* 2022;**16**:945398. <https://doi.org/10.3389/fnins.2022.945398>
20. Harrison JK, McArthur KS, Quinn TJ. Assessment scales in stroke: Clinimetric and clinical considerations. *Clin Interv Aging* 2013;**8**:201–11. <https://doi.org/10.2147/CIA.S32405>

21. Nijs J, Kosek E, Chiarotto A et al. Nociceptive, neuropathic, or nociplastic low back pain? The low back pain phenotyping (BAC-PAP) consortium's international and multidisciplinary consensus recommendations. *Lancet Rheumatol* 2024;**6**:e178–88. [https://doi.org/10.1016/S2665-9913\(23\)00324-7](https://doi.org/10.1016/S2665-9913(23)00324-7)
22. Barker AL, Nitz JC, Low Choy NL et al. Measuring fall risk and predicting who will fall: Clinimetric properties of four fall risk assessment tools for residential aged care. *J Gerontol A Biol Sci Med Sci* 2009;**64**:916–24. <https://doi.org/10.1093/gerona/glp041>
23. Bloem BR, Marinus J, Almeida Q et al. Measurement instruments to assess posture, gait, and balance in Parkinson's disease: critique and recommendations. *Mov Disord* 2016;**31**:1342–55. <https://doi.org/10.1002/mds.26572>
24. Goetz CG, Nutt JG, Stebbins GT. The unified dyskinesia rating scale: presentation and clinimetric profile. *Mov Disord* 2008;**23**:2398–403. <https://doi.org/10.1002/mds.22341>
25. Chouinard G, Cosci F, Chouinard VA et al. The extrapyramidal symptom rating scale and its abbreviated version: a critical review of clinimetric properties. *Psychother Psychosom* 2023;**92**:359–66. <https://doi.org/10.1159/000535113>
26. Elmore JG, Wells CK, Lee CH et al. Variability in radiologists' interpretations of mammograms. *N Engl J Med* 1994;**331**:1493–9. <https://doi.org/10.1056/NEJM199412013312206>
27. Smyllie HC, Blendis LM, Armitage P. Observer disagreement in physical signs of the respiratory system. *Lancet* 1965;**2**:412–3. [https://doi.org/10.1016/s0140-6736\(65\)90759-2](https://doi.org/10.1016/s0140-6736(65)90759-2)
28. Shoo BA, Sagebiel RW, Kashani-Sabet M. Discordance in the histopathologic diagnosis of melanoma at a melanoma referral center. *J Am Acad Dermatol* 2010;**62**:751–6. <https://doi.org/10.1016/j.jaad.2009.09.043>
29. 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. <https://doi.org/10.1016/j.jclinepi.2004.12.006>
30. Husted JA, Cook RJ, Farewell VT et al. Methods for assessing responsiveness: a critical review and recommendations. *J Clin Epidemiol* 2000;**53**:459–68. [https://doi.org/10.1016/S0895-4356\(99\)00206-1](https://doi.org/10.1016/S0895-4356(99)00206-1)
31. Wright JG, Feinstein AR. A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales. *J Clin Epidemiol* 1992;**45**:1201–18. [https://doi.org/10.1016/0895-4356\(92\)90161-F](https://doi.org/10.1016/0895-4356(92)90161-F)
32. Charlson ME, Carrozzino D, Guidi J et al. Charlson comorbidity index: a critical review of clinimetric properties. *Psychother Psychosom* 2022;**91**:8–35. <https://doi.org/10.1159/000521288>
33. Feinstein AR. The pre-therapeutic classification of co-morbidity in chronic disease. *J Chronic Dis* 1970;**23**:455–68. [https://doi.org/10.1016/0021-9681\(70\)90054-8](https://doi.org/10.1016/0021-9681(70)90054-8)
34. de Groot V, Beckerman H, Lankhorst GJ et al. How to measure comorbidity. A critical review of available methods. *J Clin Epidemiol* 2003;**56**:221–9. [https://doi.org/10.1016/s0895-4356\(02\)00585-1](https://doi.org/10.1016/s0895-4356(02)00585-1)
35. Church S, Rogers E, Rockwood K et al. A scoping review of the clinical frailty scale. *BMC Geriatr* 2020;**20**:393. <https://doi.org/10.1186/s12877-020-01801-7>
36. Kim DH, Rockwood K. Frailty in older adults. *N Engl J Med* 2024;**391**:538–48. <https://doi.org/10.1056/NEJMra2301292>
37. McEwen BS. Protective and damaging effects of stress mediators. *N Engl J Med* 1998;**338**:171–9. <https://doi.org/10.1056/NEJM199801153380307>
38. Fava GA, McEwen BS, Guidi J et al. Clinical characterization of allostatic overload. *Psychoneuroendocrinology* 2019;**108**:94–101. <https://doi.org/10.1016/j.psyneuen.2019.05.028>
39. Guidi J, Lucente M, Sonino N et al. Allostatic load and its impact on health: a systematic review. *Psychother Psychosom* 2021;**90**:11–27.
40. Sonino N, Fava GA. A simple instrument for assessing stress in clinical practice. *Postgrad Med J* 1998;**74**:408–10. <https://doi.org/10.1136/pgmj.74.873.408>
41. Guidi J, Fava GA. The clinical science of euthymia: a conceptual map. *Psychother Psychosom* 2022;**91**:156–67. <https://doi.org/10.1159/000524279>
42. Dowie A. The enigma of health: cultural, health political, and philosophical aspects. *J R Coll Physicians Edinb* 2021;**51**:37–42. <https://doi.org/10.4997/jrcpe.2021.109>
43. Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *JAMA* 1994;**272**:619–26. <https://doi.org/10.1001/jama.1994.03520080061045>
44. Testa MA, Simonson DC. Assessment of quality-of-life outcomes. *N Engl J Med* 1996;**334**:835–40. <https://doi.org/10.1056/NEJM199603283341306>
45. Feinstein AR, Josephy BR, Wells CK. Scientific and clinical problems in indexes of functional disability. *Ann Intern Med* 1986;**105**:413–20. <https://doi.org/10.7326/0003-4819-105-3-413>
46. Fava GA, Patierno C, Sonino N et al. Key psychosocial issues in medical care. *Acta Psychiatr Scand* 2024;**149**:368–77. <https://doi.org/10.1111/acps.13667>
47. Kroenke K. A practical and evidence-based approach to common symptoms: a narrative review. *Ann Intern Med* 2014;**161**:579–86. <https://doi.org/10.7326/M14-0461>
48. Kluivers KB, Hendriks JC, Mol BW et al. Clinimetric properties of 3 instruments measuring postoperative recovery in a gynecologic surgical population. *Surgery* 2008;**144**:12–21. <https://doi.org/10.1016/j.surg.2008.03.027>
49. Wade DT, Smeets RJ, Verbunt JA. Research in rehabilitation medicine: methodological challenges. *J Clin Epidemiol* 2010;**63**:699–704. <https://doi.org/10.1016/j.jclinepi.2009.07.010>
50. Thomas RJ. Cardiac rehabilitation—challenges, advances, and the road ahead. *N Engl J Med* 2024;**390**:830–41. <https://doi.org/10.1056/NEJMra2302291>
51. Fava GA, Cosci F, Sonino N et al. Understanding health attitudes and behavior. *Am J Med* 2023;**136**:252–9. <https://doi.org/10.1016/j.amjmed.2022.10.019>
52. Tomba E. Assessment of lifestyle in relation to health. *Adv Psychosom Med* 2012;**32**:72–96. <https://doi.org/10.1159/000330007>
53. Vandenbroucke JP, Psaty BM. Benefits and risks of drug treatments: how to combine the best evidence on benefits with the best data about adverse effects. *JAMA* 2008;**300**:2417–9. <https://doi.org/10.1001/jama.2008.723>
54. Fava GA, Rafanelli C. Iatrogenic factors in psychopathology. *Psychother Psychosom* 2019;**88**:129–40. <https://doi.org/10.1159/000500151>
55. De Las Cuevas C, de Leon J. Reviving research on medication attitudes for improving pharmacotherapy: focusing on adherence. *Psychother Psychosom* 2017;**86**:73–9. <https://doi.org/10.1159/000450830>
56. Richardson WS, Doster LM. Comorbidity and multimorbidity need to be placed in the context of a framework of risk, responsiveness, and vulnerability. *J Clin Epidemiol* 2014;**67**:244–6. <https://doi.org/10.1016/j.jclinepi.2013.10.020>
57. Hanlon JT, Schmader KE. The medication appropriateness index at 20: where it started, where it has been, and where it may be going. *Drugs Aging* 2013;**30**:893–900. <https://doi.org/10.1007/s40266-013-0118-4>

58. Young-Hyman D, de Groot M, Hill-Briggs F et al. Psychosocial care for people with diabetes: a position statement of the American Diabetes Association. *Diabetes Care* 2016;**39**:2126–40. <https://doi.org/10.2337/dc16-2053>
59. Horwitz RI, Baker JB, Ghatak A et al. Randomization bias, multimorbidity, and the composite clinical score. *Psychother Psychosom* 2024;**93**:219–23. <https://doi.org/10.1159/000539522>
60. Sechrest L. Incremental validity: a recommendation. *Educ Psychol Meas* 1963;**23**:153–8. <https://doi.org/10.1177/001316446302300113>
61. Fava GA, Guidi J, Rafanelli C et al. The clinical inadequacy of the placebo model and the development of an alternative conceptual framework. *Psychother Psychosom* 2017;**86**:332–40. <https://doi.org/10.1159/000480038>
62. Wylie CM. The definition and measurement of health and disease. *Public Health Rep* 1970;**85**:100–4.
63. Taylor DC. The components of sickness: diseases, illnesses, and predicaments. *Lancet* 1979;**314**:1008–10. [https://doi.org/10.1016/S0140-6736\(79\)92573-X](https://doi.org/10.1016/S0140-6736(79)92573-X)
64. Huppert FA. Psychological well-being: evidence regarding its causes and consequences. *Applied Psychology: Health and Well-Being* 2009;**1**:137–64.
65. Eriksson M, Ekstrom-Bergstrom A, Arvidsson S et al. Meaning of wellness in caring science based on Rodger's evolutionary concept analysis. *Scand J Caring Sci* 2024;**38**:185–99. <https://doi.org/10.1111/scs.13196>
66. Detsky AS. Learning the art and science of diagnosis. *JAMA* 2022;**327**:1759–60. <https://doi.org/10.1001/jama.2022.4650>
67. Fava GA, Sonino N, Aron DC et al. Clinical interviewing: an essential but neglected method of medicine. *Psychother Psychosom* 2024;**93**:94–9. <https://doi.org/10.1159/000536490>
68. Sonino N, Fava GA, Aron DC et al. The role of interviewing in endocrine practice. *J Endocrinol Invest* 2025, March 29, <https://doi.org/10.1007/S40618-025-02565-w>
69. Feinstein AR. "Clinical judgment" revisited: the distraction of quantitative models. *Ann Intern Med* 1994;**120**:799–805. <https://doi.org/10.7326/0003-4819-120-9-199405010-00012>