

Clinimetrics and Clinical Psychometrics: Macro- and Micro-Analysis

Elena Tomba^a Per Bech^b

^aLaboratory of Psychosomatics and Clinimetrics, Department of Psychology, University of Bologna, Bologna, Italy;

^bPsychiatric Research Unit, Mental Health Centre North Zealand, University of Copenhagen, Hillerød, Denmark

Key Words

Clinimetrics · Clinical psychometrics ·
Macro-/Micro-analysis · Clinical judgment

Abstract

Background: Clinimetrics was introduced three decades ago to specify the domain of clinical markers in clinical medicine (indexes or rating scales). In this perspective, clinical validity is the platform for selecting the various indexes or rating scales (macro-analysis). Psychometric validation of these indexes or rating scales is the measuring aspect (micro-analysis). **Methods:** Clinical judgment analysis by experienced psychiatrists is included in the macro-analysis and the item response theory models are especially preferred in the micro-analysis when using the total score as a sufficient statistic. **Results:** Clinical assessment tools covering severity of illness scales, prognostic measures, issues of co-morbidity, longitudinal assessments, recovery, stressors, lifestyle, psychological well-being, and illness behavior have been identified. **Conclusion:** The constructive dialogue in clinimetrics between clinical judgment and psychometric validation procedures is outlined for generating developments of clinical practice in psychiatry.

Copyright © 2012 S. Karger AG, Basel

Introduction

Clinical judgment in psychiatry includes markers or indexes of high clinical validity, but from a scientific measurement point of view some clinical markers, even when performed by experienced psychiatrists, have been found to be questionable due to low inter-rater reliability. However, we have very few studies focusing directly on clinical judgment analysis [1].

When the Diagnostic and Statistical Manual of Mental Disorders, second revision [2], was released, Spitzer and Wilson [3] remarked that, compared to the DSM-I [4], the ten major diagnostic categories were no longer termed 'reactions', in favor of a return to a more Kraepelinian approach by considering shared phenomenology with the clinical syndromes.

At a symposium in 1977 at which Spitzer released the guiding principles for DSM-III, Feinstein [5] gave a critical review of diagnoses in psychiatry. In this presentation Feinstein [5] introduced his clinimetric ideas, going on to develop these thoughts a decade later in his monograph on clinimetrics [6]. In 1977 Feinstein [5] concluded that both clinical medicine and clinical psychiatry were trying to escape the clinical challenges that confront them. These escapes take a form in which we hope that investigators working in some nonclinical domains, i.e. the biophysicist, the biostatistician, or the pharmacologist, will take care of our classification problem. Feinstein's advice

was therefore to avoid these kinds of traditional clinical escapes and to consider all clinical markers as scientific objects, although their magnitude and complexity are enormous. In his monograph on clinimetrics, Feinstein [6] made a major attempt to cover the diverse clinical markers we have to take into account. His emphasis on the concept of clinimetrics should be seen as an attempt to refer to measurement although many of the clinical markers Feinstein collected might have a questionable measurement aspect. In this review we will try to discuss the measurement aspects in clinimetrics in a dialogue between clinical judgment and psychometrics.

Macro- versus Micro-Analysis of Clinical Markers

In daily clinical practice when making an initial clinical assessment of the patient, the experienced psychiatrist takes the major diagnostic categories (dementia, schizophrenia, mania, depression, anxiety, neuroticism, and extraversion/introversion) into consideration.

With DSM-III [7] the full Kraepelinian approach was accepted, as the major diagnostic categories were defined by their shared symptoms. By focusing on the shared course of symptoms, Kraepelin introduced clinical psychometrics in psychiatry [8].

The initial clinical judgment made by experienced psychiatrists has been referred to by several authors [9–11] as a clinimetric macro-analysis. This macro-analysis rests on the evaluation performed by the experienced psychiatrist. In the next step of the assessment of the patient, we have to do a micro-analysis, i.e. a detailed analysis of the symptoms or other clinical markers in order to perform a measurement using unidimensional scales [11–13].

Clinimetrics

The term clinimetrics was introduced by Feinstein three decades ago [6, 14] to indicate a domain in clinical medicine concerned with indexes or rating scales. According to Feinstein the clinical index of global improvement during treatment is the most unique component in clinical medicine. Thus, when we ask our patients about any change in terms of mental activity since the start of treatment, we are asking whether the patients are much better, slightly better, the same, slightly worse, or much worse. Such questions are considered very important by the patients and consequently give them the feeling that the clinician is thoroughly familiar with the kinds of

problems that confront them. However, according to Wright and Feinstein [15], this single Likert item of global judgment has too many potential problems of measurement to be regarded as a psychometrically valid scale although it is still clinically meaningful.

By placing the evaluation of global improvement within the macro-analysis field in clinical markers, we have no need to make a psychometric analysis, this has to be dealt with using micro-analysis.

Against this background the term clinimetrics seems to rely too much on the psychometric validation procedure for clinical markers. At the same time it seems clear that clinical markers have to be examined using both macro- and micro-analysis.

Implications for Assessment Instruments in Clinical Psychiatry

Macro- and micro-analysis might offer a conceptual and methodological platform for a substantial revision of clinical assessment. In the following sections we will deal with some of the major issues in this respect.

Development of Rating Scales

The Hamilton Depression Scale (HAM-D) was introduced by Hamilton [16, 17] as a 17-item scale for the measurement of depression severity. A macro-analysis of the HAM-D was, however, first performed by Bech et al. [18], using experienced psychiatrists' global assessment of depression severity on a Likert scale from 0 (= no depression) to 10 (= extreme degree of depression) as an index of clinical validity. The experienced psychiatrists had grown up with Kraepelin's diagnostic categories as their 'ground schooling'. It was shown that only 6 of the 17 items were used by these experienced psychiatrists.

In our micro-analysis of this 6-item HAM-D₆ we used the item response theory model developed by Rasch [19] as the psychometric validation procedure [20, 21]. We were then able to show that the rank order of prevalence of these 6 items (depressed mood, work and interests, somatic general, psychic anxiety, feelings of guilt, and psychomotor retardation) was invariant from week to week during treatment with antidepressants, and that this rank order of prevalence was invariant when females were compared with males or when younger patients were compared with elderly patients. This invariant item ordering is often referred to as transferability. The total

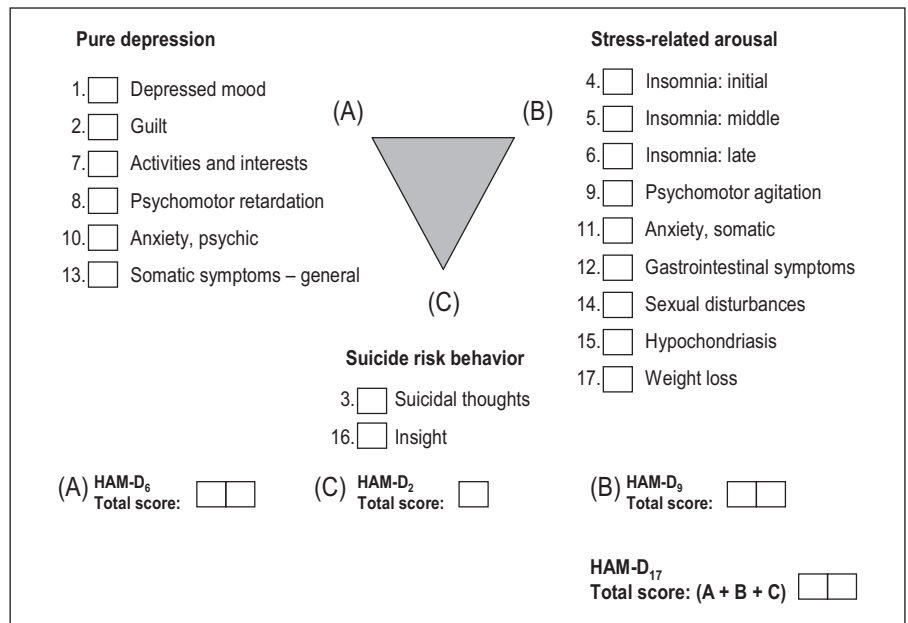


Fig. 1. ABC version of the HAM-D.

score of the HAM-D₆ was therefore a sufficient statistic, in contrast to the HAM-D₁₇ which is a multi-dimensional scale. Thus, we often need to give a profile score rather than a total score of the complete HAM-D₁₇. This is illustrated in the ABC version of the HAM-D (fig. 1). We need to score not only the HAM-D₆ for the pure depression symptoms but also the HAM-D₉ (covering the physiological arousal symptoms, e.g. sleep, agitation, somatic anxiety, sexual problems, appetite, and weight loss). The remaining 2 items, the HAM-D₂, encompass the idiographic symptoms of suicidal thoughts and insight [8].

Besides transferability, there is another important micro-analytic principle in the item response theory model, namely local independence of items. Statistically, local independence is concerned with how different items are stochastically related through independence for any fixed level of depression [22]. However, clinically we are dealing with local independence as a measure (correlation) of the extent to which the score on one item can automatically predict the score on another item [23, 24]. In the Montgomery Åsberg Rating Scale for Depression (MADRS) [24] the score on the item of reported sadness will have a clear local dependence on the item of apparent sadness [25]. It is such a dialogue between the statistical and the clinical aspects of clinimetrics that might improve the scientific level of the clinical markers we are confronted with. Another example of problems with regard to the local dependence of items is to be found in the Beck Depression Inventory (BDI) [26], recently discussed by Greenberg [27].

Careful symptom discrimination via interviewing using an observer-rated instrument may allow us to attribute differential emphasis on specific symptoms. In clinimetrics, major and minor symptoms may be discriminated, unlike in psychometrics, where all items are weighed the same [11]. A recent example was provided by the use of an item of the Clinical Interview for Depression [28], reactivity to social environment, to characterize the clinical features [29] and response to treatment [30] of cyclothymic disorder.

The Bech-Rafaelsen Melancholia Scale (MES) and the Mania Scale (MAS) as well as the schizophrenicity subscale of the Brief Psychiatric Rating Scale (BPRS) and the Hamilton Anxiety Scale (HAS) have all been validated by use of the macro- and micro-analysis [8, 31].

Among the comprehensive distress evaluation scales, a suitable self-rating instrument is the Symptom Questionnaire (SQ). It has been used in psychiatric and psychosomatic settings to assess both symptomatology and emotional well-being and their relationship. Kellner's SQ [32] is a simple and brief self-report questionnaire which contains single-word items instead of questions. The total number of items (92 items) has been increased in order to improve the scale's sensitivity [33]. Responders should answer yes/no or true/false to the item, instead of measuring the frequency and severity of symptoms. The main novelty introduced by Kellner is the presence of items assessing emotional well-being. They were introduced to obtain a more complete evaluation of the patients' clinical

Table 1. Selected instruments for assessing affective symptoms

Index	Description
Hamilton Depression Scale (HAM-D ₆) [18]	An observer-rated instrument for measuring only 6 of the 17 HAM-D (Hamilton, 1960) items corresponding to the experienced psychiatrists' global ratings: depressed mood, guilt, work and interests, psychomotor retardation, psychic anxiety, and general somatics.
Hamilton Anxiety Scale (HAM-A ₆) [36]	The HAM-A ₆ contains the items from the original 14 HAM-A (Hamilton, 1959) that were most discriminating between drug and placebo in patients with a DSM-IV diagnosis of generalized anxiety disorder: psychic tension, phobia, anxiety (psychic), muscular tension, agitation, and concentration difficulties.
Clinical Interview for Depression [28, 37]	A 36-item observer-rated scale for assessing a wide range of affective symptoms (depression, anxiety, and irritability).
Symptom Questionnaire [32]	A 92-item self-rated scale for measuring anxiety, depression, somatic symptoms, and hostility.
Bech-Rafaelsen Melancholia Scale (MES) [38]	An 11-items clinician-rated symptom scale that assesses social life activities and interests, lowered mood, sleep disturbances, anxiety, introversion, concentration difficulties, tiredness, feelings of worthlessness and guilt, decreased verbal activity, suicidal thoughts, and decreased motor activity.
Bech-Rafaelsen Mania Scale (MAS) [39]	A clinician-rated symptom scale designed to assess the presence and severity of clinical features of mania and hypomania and the effectiveness of therapeutic interventions for bipolar disorder. MAS assesses 11 items (elevated mood, pressure of speech, increased social contact, increased motor activity, sleep disturbances, social activities and distractibility, hostility and irritability, increased sexual activity, increased self-esteem, flight of thoughts, and noise level of speech and other vocal activity).
The Hypomania Checklist (HCL-32) [40]	The HCL-32 is a self-administered paper-and-pencil inventory made up of 32 yes/no items used to identify the hypomanic component in patients with depressive episodes in order to help the clinician to diagnose bipolar-II and minor bipolar disorders in psychiatric and general medical practice.

status and to improve the scale's internal consistency. The final SQ version yields 4 principal scales: anxiety, depression, somatization and hostility-irritability, which can be divided into 8 subscales: 4 symptom subscales (anxiety, depression, somatization, hostility) and 4 corresponding well-being subscales (relaxation, contentment, physical well-being and friendliness). The depression scale includes items such as 'hopeless' from the depressive symptoms subscale and items such as 'happy' from the contentment subscale. However, a single high score cannot be equated with a diagnosis of psychopathology, without further clinical assessment. High scores on the somatization scale can indicate the presence of somatization only after exclusion of other medical conditions [34]. In double-blind, cross-over studies, negatively formulated items tended to be more sensitive in discriminating between the effects of a psychotropic drug and placebo [35] (see table 1 [32, 36–40]).

The macro versus micro analysis focuses not only on the evaluation of the severity of symptoms but also on their sequence, rate of progression of illness, and co-morbidity, and on problems of functional life, such as well-

being [41]. Two patients may both meet the diagnostic criteria for major depression, but they may present very different scores on a rating scale. What they do not share may be as important as what they do share: their different scores may reflect different symptom severities, perceptions, and illness attitudes, and these variables may affect the clinical course of the disorder in these two patients.

Prognostic Markers

As discussed elsewhere [8], the neo-Kraepelinic paradigm in DSM-III, DSM-IV or ICD-10 refers to a thorough symptom-based description of mental disorders. However, Kraepelin's description was not completely atheoretical, as the medical model of disorders also 'hovered' in the back of his mind. Table 2 shows the 'medical model' in operation when the experienced psychiatrist performs his or her prognostic or etiological judgment. For example, in (F2a) there is a long-term psycho-social stressor which can explain the symptoms while in (F2b) a single, but exceptional event can explain the symptoms.

Table 2. Etiological considerations in major depression (Clinical Interview for Depression and Related Syndromes, section F) [8]

1	Lack of insight
2a	Continuing psychological stressors
2b	Exposure to an exceptionally catastrophic event
3	Neuroticism
4	Increased reactivity towards environment
5	Diurnal variation (worst in the evening)
6	Diurnal variation (worst in the morning)
7	Quality of depression
8	Persistence of the clinical picture
9	Depressive delusions or hallucinations
10	Previous depressive downs
11	Previous depressive episodes
12	Previous hypomanic ups
13	Previous manic episodes
14	Previous mixed states
15	Hereditary disposition
16	Somatic illness
17	Drug-/substance-induced condition

Emmelkamp [12] provided an example of the macro-analysis when confronted with patients who had developed depressive episodes. A substantial number of depressed patients had been through a chronic (F2a) marital stressor. According to Emmelkamp [12] the data suggest that depression and marital stressors are closely linked. Furthermore, persons who after treatment for depression return to distressed marriages are more likely to experience relapse [42]. When viewing the results of studies on marital stress-induced depression, behavioral marital therapy in maritally distressed couples seems to have an exclusive effect on the marital relationship. This is not found in individual cognitive behavioral therapy, while it is as effective as cognitive therapy in reducing depression [43]. Not surprisingly, behavioral marital therapy was hardly effective in depressed patients who did not experience marital problems [10].

An example of macro-analysis can be found in table 2; in this example, F12 is a macro-analytic attempt to evaluate hidden bipolarity by asking about previous hypomanic ups. We recently made an analysis of the Hypomania Checklist (HCL-32) and showed that around 50% of unipolar depressed patients not responding to antidepressants had a hidden bipolarity [44].

F17 in table 2 is another example. This is an attempt to identify drug-induced disorders by macro-analysis. We recently showed that patients with hepatitis C might develop a major depression when treated with interferon [45].

Comorbidity

Comorbidity is the co-occurrence of multiple disorders of macro-analytic significance. In table 2 somatic illness is indicated (F16) when it is considered to have etiological significance, i.e. the mental condition would not have developed without this somatic illness, e.g. post-stroke depression.

The co-occurrence of several mental disorders has been investigated in longitudinal observation [46]. In major depression, two thirds of the patients meet criteria for another axis I disorder (anxiety disorders in particular) and one third have two or more disorders [47]. The presence of anxiety disorders appears to predict the persistence and recurrence of depressive illness [48, 49]. Some forms of comorbidity may be masked by manifestations of the disorder and only become evident when the most severe symptoms have abated [50]. The current emphasis in psychiatry is on assessment of symptoms resulting in syndromes identified by diagnostic criteria (DSM). However, an emerging awareness that also psychological symptoms which do not reach the threshold of a psychiatric disorder may affect the quality of life and entail pathophysiological and therapeutic implications led to the development of the Diagnostic Criteria for Psychosomatic Research (DCPR) [51–53] together with a specific interview to assess patients [54, 55]. The DCPR were introduced in 1995 and tested in various clinical settings [51–53]. Diagnostic criteria for the following syndromes are included: abnormal illness behavior (disease phobia, thanatophobia, health anxiety, and illness denial), somatization syndromes (persistent somatization, functional somatic symptoms secondary to a psychiatric disorder, conversion symptoms, and anniversary reactions), irritability, type A behavior, demoralization, and alexithymia. The DCPR allow a far more sophisticated qualitative assessment of patients than the one dimensional DSM checklist of psychological symptoms.

The DCPR interview encompassing these various diagnostic rubrics has shown excellent inter-rater reliability, construct validity, and predictive validity for psychosocial functioning and treatment outcomes [54, 56]. It has yielded important information as the clinical characteristics of anxiety disorders [57], major depressive disorders [58], psychiatric disorders in the setting of endocrine disease [59] and cardiovascular illness [60]. Whenever the DCPR have been used in conjunction with the DSM criteria, they have been found to carry additional clinical information [52].

Longitudinal Assessment

Residual symptoms, despite successful response to therapy, appear to be the rule after completion of drug or psychotherapeutic treatment in both mood and anxiety disorders [50, 61]. The presence of residual symptoms has been correlated with poor long-term outcomes [50, 61]. There appears to be a relationship between residual and prodromal symptomatology: in fact most of the residual symptoms occur in the prodromal phase of illness. These findings have led to the hypothesis that upon recovery residual symptoms may progress to become prodromal symptoms of relapse and that treatment directed toward residual symptoms may yield long-term benefits [62]. Treatments which potentially aim to achieve different effects (e.g. pharmacotherapy and psychotherapy) may thus be used in a sequential order. One type of treatment (e.g. psychotherapy) may be employed to improve symptoms which the other type of treatment (e.g. pharmacotherapy) was unable to affect. The aim of the sequential approach is to add therapeutic ingredients as long as they are needed. In this sense, it introduces a conceptual shift in clinical practice. Therapeutic targets are not predetermined but depend on the response of patients to the first course of treatment [63].

Recovery

Clinimetrics also provides relevant clinical implications in the definition of recovery. Commonly, the concept of recovery reflects that of 'improvement' which refers to the clinical distance along which the current state of the patient is compared to the pretreatment position [64]. In this sense, recovery can be expressed either as a categorical variable (present/absent) or as a comparative category (nonrecovered, slightly recovered, moderately recovered, greatly recovered). Both expressions require arbitrary cut-off points related to the amount of improvement.

Feinstein [6] differentiates between the monadic transition index (devoted exclusively to the rating of change during treatment), the dyadic component index (e.g. before and after treatment ratings) and the polyadic component index (when more than two ratings, such as with the sequential model, have been performed). A depressed patient who is asked how he or she feels after 3 weeks of treatment with an antidepressant drug and replies 'just fine' (instead of 'better') uses a self-monadic component. The amount of change induced by treatment, however,

may make him/her overlook the distance from an intended goal, such as the preepisode state. The physician may collude with the patient in this illusion of wellness, since he/she may be more gratified by the amount of improvement induced in the patient than by the current distance from an intended goal [65]. Clinicians may choose recovery as a target that is negotiated between the doctor and the patient. The doctor can insist that the target be reasonable (e.g. not asking to be better than before the illness). Nevertheless, the idea of successful recovery may differ from one patient to the next and should not be constrained too much by the doctor's ideas. We should accept the possibility that a treatment may determine abatement of symptoms in some patients, leave a substantial residual symptomatology in others, yield an unsatisfactory response in others, and provide no benefit or even cause harm in a few. The type of residual symptomatology varies widely from patient to patient and needs to be assessed individually [65]. Frequently, in fact, patients with affective disorders, despite being treated effectively with psychotherapy or pharmacological interventions, can continue to present high subclinical symptomatology and impaired levels of psychological well-being compared to healthy control subjects [66, 67]. The persistence of this impairment after drug treatment of depression is likely to be a poor prognostic index [34, 68].

In a survey on factors identified by depressed outpatients as important in determining remission, those most frequently judged as such were the presence of features of positive mental health, such as optimism and self-confidence, a return to one's usual, normal self, and a return to the usual level of functioning [69]. Ryff's psychological dimensions [70] may be instrumental in assessing both the process and the definition of recovery [71].

Stressors

When Hans Selye developed his theory of the concept of stress, he discriminated between stressors (strain), stress (the bodily reactions to such strain), and distress (the mental reaction to the strain). As discussed elsewhere [8], Selye's original stress model had the daily strain at work and at home as stressors, not the single catastrophic stressor that completely dominates the picture in posttraumatic stress disorder (PTSD).

The ABC version of the HAM-D₁₇ (fig. 1) actually covers the development of PTSD through the nine arousal symptoms in HAM-D₉ (disturbed sleep, somatic anxiety, and agitation) which develop over the first 1–2 months.

Table 3. Selected clinimetric instruments for rating stress and allostatic load

Index	Description
Interview for Recent Life Events [81]	Semistructured research interview covering 64 life events, with determination of categories and independence.
Interview for Life Events [82]	Lengthy interview covering 38 types of life events and their circumstances.
The Psychosocial Index [83]	A 55-item self-report questionnaire for determining the amount of allostatic load together with abnormal illness behavior, psychological well-being, psychological distress, and quality of life.
Job Content Questionnaire [84]	A 49-item self-report questionnaire for the evaluation of workplace environment stress.

After this the pure depression symptoms appear (HAM-D₆) with depressed mood, tiredness, lack of interests, guilt, and psychic anxiety. If untreated, PTSD might be more chronic, demonstrating the HAM-D₂ symptoms (feelings of emptiness, lack of insight, and suicidal thoughts).

In work-related stress conditions the symptoms basically follow the same course, but less clearly [72]. Cumulative exposures to stressful experiences, often subsumed under the term ‘allostatic load’, may increase vulnerability to disease onset [63, 73]. The concept of allostatic load represents the cost of the continual adjustment of the internal milieu required by the organism to adapt to different environmental demands [74]. The majority of studies on allostatic load have focused mainly on identifying its biological components, which are believed to mediate the relationship between stress and illness. Recently, according to the clinimetric approach, Fava et al. [41, 75] have introduced specific criteria for the evaluation of allostatic overload (AO) in the clinical setting. The first criterion (A) deals with the specification of the stressor. Stress is measured by evaluating the occurrence of stressful situations in areas of major importance to the patient, resulting in major changes in the patient’s living conditions, social and family circle, and work, as well as subtle and long-standing life situations, such as those occurring at work. In both cases the situation needs to be experienced by the individual as taxing or exceeding his/her coping skills. The second criterion (B) is concerned with the clinical manifestations ranging from psychiatric to psychosomatic/subclinical symptoms, as defined by the DSM-IV [76] and/or the Diagnostic Criteria for Psychosomatic Research (DCPR) [51], and from impairment in social and occupational functioning to a decrease in well-being [71].

A latent effect of adversities during childhood, when the developing brain is more sensitive to a variety of en-

vironmental signals, has been demonstrated and has created a need to classify the degree of stress children are exposed to for health and disease prevention uses [77].

A number of clinimetric instruments based on structured methods of data collected and used in controlled studies have been developed (table 3) [78–81].

Lifestyle

An increasing body of evidence links the initiation and progression of several medical disorders, such as diabetes, obesity, and cardiovascular illness [82–84]. The benefits of modifying the lifestyle were demonstrated in several randomized controlled trials. Table 4 illustrates some instruments which may assess lifestyle [85–88].

Psychological Well-Being

It is commonly considered that Bentham (1748–1832) made the first attempt to measure psychological well-being via questionnaires [8, 89]. He defined subjective well-being as the difference between the sum of all kinds of pleasure and the sum of all kinds of pain experienced by the individual in a given period of time. Psychometrically, it is very important to defined the time frame, the ‘given period of time’, often referred to clinimetrically as the ‘window’ [8].

Macro-analytically we can classify psychological well-being markers into the ‘hedonic’ scales and the ‘eudemonic’ scales. When going through Murray’s basic human needs and their hierarchic arrangements [90], the hedonic needs might be considered as a global index of measurement [8]. Ryff’s Psychological Well-Being Scale (PWB)

Table 4. Selected instruments for assessing lifestyle (such as alcohol and tobacco use, sleep, and physical activity)

Index	Description
Alcohol Use Disorders Identification Test (AUDIT) [85]	It consists of a 10-item-core self-report or clinician-administered assessment and allows clinicians to collect subclinically significant alcohol indices that have not reached levels of harm associated with alcohol dependence.
Addiction Severity Index (ASI) [86]	This 200-item clinician administered semistructured interview was designed to assess the complex nature of addiction.
Fagerström Test for Nicotine Dependence (FTND) [87]	This screening instrument can be used in an interview or a self-report format covering cigarette consumption, compulsion to use, and dependence. The brevity (6 items) and easy scoring of the FTND make it an efficient way to obtain clinically meaningful information [30].
Pittsburgh Sleep Quality Index (PSQI) [88]	The 19-item PSQI was designed to provide a reliable, valid, and standardized measure of sleep quality. It covers several domains such as subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep medications, and daytime dysfunction. The global score facilitates comparisons between groups and assesses changes over time [30].

Table 5. Selected instruments for assessing psychological well-being

Index	Description
Symptom Questionnaire [32]	A 92-item brief self-rating scale for measuring relaxation, contentment, physical well-being, and friendliness.
Psychological Well-Being Scales (PWB) [70]	An 84-item self-administered questionnaire assessing a multidimensional view of psychological well-being based on 6 dimensions: autonomy, environmental mastery, personal growth, purpose in life, interpersonal relationships, and self-acceptance.
WHO-Five Well-Being Index [91]	A self-rating instruments that includes 5 items that cover positive mood (good spirits, relaxation), vitality (being active and waking up fresh and rested), and general interests (being interested in things).

Table 6. Selected instruments for assessing illness behavior

Index	Description
Illness Attitudes Scales [101, 102]	A 21-item self-rating scale for assessing worry about illness, concern about pain, health habits, hypochondriasis, fear of dying, disease phobia and bodily preoccupations.
Illness Behavior Questionnaire [103]	A self-rating scale which measures hypochondriasis, disease conviction, psychological versus somatic perception of illness, affective inhibition, dysphoria, denial, and irritability.
Illness Perception Questionnaire [104]	A 38-item self-report method for assessing cognitive representations of illness: identity, cause, time-line, consequences, and cure control.

[70] should be consulted when attempting to measure eudemonia, covering such elements as autonomy, personal growth, purpose of life, interpersonal relationships, and self-acceptance. Micro-analytically the PWB is a multidimensional approach implying that a profile score rather than a global score should be used (table 5) [32, 70, 91].

Among the hedonic well-being scales, Hall et al. [91] identified WHO-5 as a clinically very valid scale by judgment analysis. The WHO-5 questionnaire has been widely used and translated into around 30 languages. WHO-5 is a general psychological well-being scale based on five positively phrased items covering the 'window' of the past

2 weeks. Its total score is a sufficient statistic because the scale fulfills the item response theory models [8]. The theoretic score range goes from 0 (worst imaginable quality of life) to 100 (best imaginable quality of life).

Furthermore, the SQ [32] also represents a suitable clinimetric instrument for assessing psychological well-being. It yields 4 principal scales: anxiety, depression, somatization, and hostility-irritability, which can be divided into 8 subscales: 4 symptom subscales (anxiety, depression, somatization, hostility) and 4 corresponding well-being subscales (relaxation, contentment, physical well-being, and friendliness). The SQ well-being scales can, however, also be scored separately (scoring 1 when the answer indicates the presence of well-being), so that the higher the score is, the higher the well-being is. The total scale internal consistency can therefore be computed by comparing the symptom subscale score with the corresponding well-being one. Through this scoring procedure, it is possible to measure also the balance of positive and negative effects, which has been described in major depression [92] and found to be affected by cognitive-behavioral therapy [93]. In another clinical trial [32], the treatment of psychiatric symptoms induced improvement of well-being and, in this case, the SQ subscales describing well-being were more sensitive to drug effects than subscales describing symptoms [34].

Several studies have suggested that psychological well-being plays a buffering role in coping with stress and has a favorable impact on disease [71, 94–96]. Table 5 illustrates some methods for assessing psychological well-being, both eudaimonic and hedonic [32, 70, 91].

Illness Behavior

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 gives rise to psychological responses which influence the patient's experience and behavior, as well as the course and therapeutic outcome of a given illness episode [97]. The study of illness behavior, the ways in which individuals experience, perceive, evaluate and respond to their own health status, has yielded important information in medical patients [98, 99]. There are many manifestations of abnormal illness behavior, which range from hypochondriasis and disease phobia to illness denial and lack of compliance [100].

Table 6 indicates different clinimetric strategies for assessing illness behavior [101–104].

Conclusion

With the introduction of the term 'clinimetrics' Feinstein made an attempt to avoid the temptation to ask other scientists in the nonclinical domains of medicine (e.g. biophysicists or pharmacologists) for help when evaluating patients because the clinicians found the clinical world too complex and too difficult to handle in a scientific way.

We have followed Feinstein's strict clinimetric approach by using macro- and micro-analyses of the various clinical markers. With this pure clinical approach we have attempted to identify clinical assessment tools covering severity scales, prognostic markers, issues of comorbidity, longitudinal assessments, recovery, stressors, lifestyle, psychological well-being, and illness behavior. In this process we have relied on a constructive dialogue between the clinical and the purely psychometric aspects of clinimetrics. Such dialogue may generate important developments: the staging of mental disorders with particular reference to their longitudinal development [105] and their previous treatment history [106]; the construction of subtypes of illness, such as has been suggested in major depression [28, 107, 108]; the broadening of assessment areas, such as demoralization and irritable mood [109–112], and psychological well-being [91].

References

- 1 Bech P, Haaber A, Joyce CR: Experiments on clinical observation and judgement in the assessment of depression: profiled videotapes and Judgement Analysis. *Psychol Med* 1986; 16:873–883.
- 2 American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders (DSM-II), ed 2. Washington, American Psychiatric Association, 1968.
- 3 Spitzer RL, Wilson PT: A guide to the American Psychiatric Association's new diagnostic nomenclature. *Am J Psychiatry* 1968;124: 1619–1629.
- 4 American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders (DSM-I). Washington, American Psychiatric Association, 1952.
- 5 Feinstein AR: A critical overview of diagnosis in psychiatry; in Radkoff M, Stancer HC, Kedward HB, Bruner Mazel V (eds): *Psychiatric Diagnosis*. New York, Bruner/Mazel, 1977, pp 189–206.
- 6 Feinstein AR: *Clinimetrics*. New Haven, Yale University Press, 1987.
- 7 American Psychiatric Association: *The Diagnostic and Statistical Manual of Mental Disorders (DSM-III)*, ed 3. Washington, American Psychiatric Association, 1980.

- 8 Bech P: Clinical psychometrics. Oxford, Wiley Blackwell, 2012.
- 9 Emmelkamp PMG, Bouman TK, Scholing A: Anxiety Disorders. Chichester, Wiley, 1992.
- 10 Emanuels-Zuurveen L, Emmelkamp PMG: Spouse-aided therapy with depressed patients: a comparative evaluation. *Behav Modif* 1997;21:62–77.
- 11 Fava GA, Tomba E, Sonino N: Clinimetrics: the science of clinical measurements. *Int J Clin Pract* 2012;66:11–15.
- 12 Emmelkamp PM: The additional value of clinimetrics needs to be established rather than assumed. *Psychother Psychosom* 2004;73:142–144.
- 13 Bech P: Rating scales in depression: limitations and pitfalls. *Dialogues Clin Neurosci* 2006;8:207–215.
- 14 Feinstein AR: T. Duckett Jones Memorial Lecture – the Jones criteria and the challenges of clinimetrics. *Circulation* 1982;66:1–5.
- 15 Wright JG, Feinstein AR: A comparative contrast of clinimetric and psychometric methods for constructing indexes and rating scales. *J Clin Epidemiol* 1992;45:1201–1218.
- 16 Hamilton M: A rating scale for depression. *J Neurol Neurosurg Psychiatry* 1960;23:56–62.
- 17 Hamilton M: Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* 1967;6:278–296.
- 18 Bech P, Gram LF, Dein E, Jacobsen O, Vitger J, Bolwig TG: Quantitative rating of depressive states. *Acta Psychiatr Scand* 1975;51:161–170.
- 19 Rasch G: Probabilistic Models for Some Intelligence and Attainment Tests, expanded edition. Chicago, The University of Chicago Press, 1980.
- 20 Bech P, Allerup P, Gram LF, Reisby N, Rosenberg R, Jacobsen O, Nagy A: The Hamilton depression scale: evaluation of objectivity using logistic models. *Acta Psychiatr Scand* 1981;3:290–299.
- 21 Bech P, Allerup P, Reisby N, Gram LF: Assessment of symptom change from improvement curves on the Hamilton depression scale in trials with antidepressants. *Psychopharmacology* 1984;84:276–281.
- 22 de Ayala RJ: The theory and practice of item response theory. New York, Guilford Press, 2009.
- 23 Borsboom D: Measuring the mind. Cambridge, Cambridge University Press, 2005.
- 24 Montgomery SA, Asberg M: A new depression scale designed to be sensitive to change. *Br J Psychiatry* 1979;134:382–389.
- 25 Allerup P: Statistical Analysis of MADRS: A Rating Scale. Copenhagen, Danish Institute for Educational Research, 1986.
- 26 Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J: An inventory for measuring depression. *Arch Gen Psychiatry* 1961;4:561–571.
- 27 Greenberg G: Manufacturing Depression: The Secret History of a Modern Disease. London, Bloomsbury, 2010.
- 28 Guidi J, Fava GA, Bech P, Paykel E: The Clinical Interview for Depression. *Psychother Psychosom* 2011;80:10–27.
- 29 Tomba E, Rafanelli C, Grandi S, Guidi J, Fava GA: Clinical configuration of cyclothymic disturbances. *J Affect Disord* 2012;139:244–249.
- 30 Fava GA, Rafanelli C, Tomba E, Guidi J, Grandi S: The sequential combination of cognitive behavioral treatment and well-being therapy in cyclothymic disorder. *Psychother Psychosom* 2011;80:136–143.
- 31 Licht RW, Qvitzau S, Allerup P, Bech P: Validation of the Bech-Rafaelsen Melancholia Scale and the Hamilton Depression Scale in patients with major depression: is the total score a valid measure of illness severity? *Acta Psychiatr Scand* 2005;111:144–149.
- 32 Kellner R: A symptom questionnaire. *J Clin Psychiatr* 1987;48:269–274.
- 33 Kellner R, Rada RT, Andersen T, Pathak D: The effects of chlordiazepoxide on self-rated depression, anxiety, and well-being. *Psychopharmacol* 1979;64:185–191.
- 34 Rafanelli C, Ruini C: Assessment of psychological well-being in psychosomatic medicine. *Adv Psychosom Med* 2012;32:182–202.
- 35 Kellner R, Collins AC, Shulman RS, Pathak D: The short-term anti anxiety effects of propranolol HCL. *J Clin Pharmacol* 1974;14:301–304.
- 36 Bech P: Dose-response relationship of pregabalin in patients with generalized anxiety disorder: a pooled analysis of four placebo-controlled trials. *Pharmacopsychiatry* 2007;40:163–168.
- 37 Paykel ES: The clinical interview for depression. *J Affect Disord* 1985;9:85–96.
- 38 Bech P: The Bech-Rafaelsen Melancholia Scale (MES) in clinical trials of therapies in depressive disorders: a 20-year review of its use as outcome measure. *Acta Psychiatr Scand* 2002;106:252–264.
- 39 Bech P, Bolwig TG, Kramp P, Rafaelsen OJ: The Bech-Rafaelsen Mania Scale and the Hamilton Depression Scale. *Acta Psychiatr Scand* 1979;59:420–430.
- 40 Angst J, Adolfsson R, Benazzi F, Gamma A, Hantouche E, Meyer TD, Skeppar P, Vieta E, Scott J: The HCL-32: towards a self-assessment tool for hypomanic symptoms in outpatients. *J Affect Disord* 2005;88:217–233.
- 41 Fava GA, Rafanelli C, Tomba E: The clinical process in psychiatry: a clinimetric approach. *J Clin Psychiatry* 2012;73:177–184.
- 42 Emmelkamp PMG, Vedel E: Spouse-aided therapy; in Hersen M, Sledge W (eds): *The Encyclopedia of Psychotherapy*. New York, Academic Press, 2002, vol 2, pp 693–698.
- 43 Emmelkamp PMG: Behavior therapy with adults; in Lamberts L (ed): *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*, ed 5. New York, Wiley, 2003, pp 396–449.
- 44 Bech P, Christensen EM, Vinberg M, Bech Andersen G, Kessing LV: From items to syndromes in the Hypomania Checklist (HCL-32): psychometric validation and clinical validity analysis. *J Affect Disord* 2011;132:48–54.
- 45 Leutscher PD, Lagging M, Buhl MR, Pederesen C, Norkrans G, Langeland N, Morch K, Farkkila M, Hjerrild S, Hellstrand K, Bech P, NORDynamic Group: Evaluation of depression as a risk factor for treatment failure in chronic hepatitis C. *Hepatology* 2010;52:430–435.
- 46 Maj M: The aftermath of the concept of 'psychiatric comorbidity'. *Psychother Psychosom* 2005;74:67–68.
- 47 Zimmerman M, Chelminski I, McDermut W: Major depressive disorder and axis I diagnostic comorbidity. *J Clin Psychiatry* 2002;63:187–193.
- 48 Sherbourne CD, Wells KB: Course of depression in patients with comorbid anxiety disorders. *J Affect Disord* 1997;43:245–250.
- 49 Gaynes BN, Magruder KM, Burns BJ, Wagner HR, Yarnall KS, Broadhead WE: Does a coexisting anxiety disorder predict persistence of depressive illness in primary care patients with major depression? *Gen Hosp Psychiatry* 1999;21:158–167.
- 50 Fava GA: Subclinical symptoms in mood disorders: pathophysiological and therapeutic implications. *Psychol Med* 1999;29:47–61.
- 51 Fava GA, Freyberger HJ, Bech P, Christodoulou G, Sensky T, Theorell TT, Wise TN: Diagnostic criteria for use in psychosomatic research. *Psychother Psychosom* 1995;63:1–8.
- 52 Porcelli P, Rafanelli C: Criteria for psychosomatic research (DCPR) in the medical setting. *Curr Psychiatry Rep* 2010;12:246–256.
- 53 Wise TN: Diagnostic criteria for psychosomatic research are necessary for DSM-V. *Psychother Psychosom* 2009;78:330–332.
- 54 Mangelli L, Rafanelli C, Porcelli P, Fava GA: Interview for the diagnostic criteria for psychosomatic research. *Adv Psychosom Med* 2007;28:174–181.
- 55 Porcelli P, Sonino N (eds): *Psychological Factors Affecting Medical Conditions*. Basel, Karger, 2007.
- 56 Porcelli P, Bellomo A, Quartesan R, Altamura M, Iuso S, Liammaneo F, Piselli M, Elisei M: Psychosocial functioning in consultation-liaison psychiatry patients. *Psychother Psychosom* 2009;78:352–358.
- 57 Fava GA, Porcelli P, Rafanelli C, Mangelli L, Grandi S: The spectrum of anxiety disorders in the medically ill. *J Clin Psychiatr* 2010;71:910–914.
- 58 Guidi J, Fava GA, Picardi A, Porcelli P, Bellomo A, Grandi S, Grassi L, Pasquini P, Quartesan R, Rafanelli C, Rigatelli M, Sonino N: Subtyping depression in the medically ill by cluster analysis. *J Affect Disord* 2011;132:383–388.
- 59 Sonino N, Tomba E, Genesia ML, Bertello C, Mulatero P, Veglio F, Fava GA, Fallo F: Psychological assessment of primary aldosteronism. *J Clinical Endocrinol Metab* 2011;96:E878–E883.
- 60 Grandi S, Sirri L, Tossani E, Fava GA: Psychological characterization of demoralization in the setting of heart transplantation. *J Clin Psychiatry* 2011;72:648–654.

- 61 Fava GA, Mangelli L: Subclinical symptoms of panic disorder. *Psychother Psychosom* 1999;68:281–289.
- 62 Fava GA, Kellner R: Prodromal symptoms in affective disorder. *Am J Psychiatr* 1991;148:823–830.
- 63 Fava GA, Tomba E: New modalities of assessment and treatment planning in depression: the sequential approach. *CNS Drugs* 2010;24:453–465.
- 64 Bech P: Measurement of psychological distress and well-being. *Psychother Psychosom* 1990;54:77–89.
- 65 Fava GA: The concept of recovery in affective disorders. *Psychother Psychosom* 1996;65:2–13.
- 66 Rafanelli C, Park SK, Ruini C, Ottolini F, Cazzaro M, Fava GA: Rating well-being and distress. *Stress Med* 2000;16:55–61.
- 67 Fava GA, Rafanelli C, Ottolini F, Ruini C, Cazzaro M, Grandi S: Psychological well-being and residual symptoms in remitted patients with panic disorder and agoraphobia. *J Affect Disord* 2001;65:185–190.
- 68 Fava GA, Grandi S, Zielezny M, Canestrari R, Morphy MA: Cognitive behavioral treatment of residual symptoms in primary major depressive disorder. *Am J Psychiatry* 1994;151:1295–1299.
- 69 Zimmerman M, McGlinchey JB, Posternak MA, Friedman M, Attiullah N, Boerescu D: How should remission from depression be defined? *Am J Psychiatry* 2006;163:148–150.
- 70 Ryff CD: Happiness is everything, or is it? Explorations on the meaning of psychological well-being. *J Pers Soc Psychol* 1989;57:1069–1081.
- 71 Fava GA, Tomba E: Increasing psychological well-being and resilience by psychotherapeutic methods. *J Pers* 2009;77:1903–1934.
- 72 Bech P, Andersen MB, Bech-Andersen G, Tonnesen S, Agnarsson E, Borg V: Work-related stressors, depression and quality of life in Danish managers. *Eur Psychiatry* 2005;20(suppl 3):S318–S325.
- 73 McEwen BS: Protective and damaging effects of stress mediators. *New Engl J Med* 1998;338:171–179.
- 74 McEwen BS, Stellar E: Stress and the individual: mechanisms leading to disease. *Arch Intern Med* 1993;153:2093–2101.
- 75 Fava GA, Guidi J, Semprini F, Tomba E, Sonino N: Clinical assessment of allostatic load and clinimetric criteria. *Psychother Psychosom* 2010;79:280–284.
- 76 American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*, ed 4. Washington, American Psychiatric Association, 1994.
- 77 Shonkoff JP, Boyce WT, McEwen BS: Neuroscience, molecular biology, and the childhood roots of health disparities: building a new framework for health promotion and disease prevention. *JAMA* 2009;301:2252–2259.
- 78 Paykel ES: The Interview for Recent Life Events. *Psychol Med* 1997;27:301–310.
- 79 Brown GW, Harris TO: *Social Origins of Depression*. London, Tavistock, 1978.
- 80 Sonino N, Fava GA: A simple instrument for assessing stress in clinical practice. *Postgrad Med J* 1998;74:408–410.
- 81 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;3:322–355.
- 82 Tomba E: Assessment of lifestyle in relation to health. *Adv Psychosom Med* 2012;32:72–96.
- 83 Mokdad AH, Marks JS, Stroup DF, Gerberding JL: Actual causes of death in the United States, 2000. *JAMA* 2004;291:1238–1245.
- 84 Roger VL: Lifestyle and cardiovascular health: individual and societal choices. *JAMA* 2009;302:437–439.
- 85 Babor TF, de la Fuente JR, Saunders J, Grant M: AUDIT – The Alcohol Use Disorders Identification Test: Guidelines for Use in Primary Health Care – WHO/MNH/DAT 89.4. Geneva, World Health Organization, 1989.
- 86 McLellan AT, Kushner H, Metzger D, Peters R, Smith I, Grissom G, Pettinati H, Argeriou M: The fifth edition of the Addiction Severity Index. *J Subst Abuse Treat* 1992;9:199–213.
- 87 Fagerstrom KO: Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addict Behav* 1978;3:235–241.
- 88 Buysse DJ, Reynolds CF 3rd, 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.
- 89 Bech P: Rating Scales for Psychopathology, Health Status and Quality of Life: A Compendium on Documentation in Accordance with the DSM-III-R and WHO Systems. Berlin, Springer, 1993.
- 90 Murray HA: *Exploration in Personality*. New York, Oxford University Press, 1938.
- 91 Hall T, Krahn GL, Horner-Johnson W, Lamb G, Rehabilitation Research and Training Center Expert Panel on Health Measurement: Examining functional content in widely used health-related quality of life scales. *Rehabil Psychol* 2011;56:94–99.
- 92 Garamoni GL, Reynolds CF 3rd, Thase ME, Frank E, Berman SR, Fasiczka AL: The balance of positive and negative affects in major depression: a further test of the states of mind model. *Psychiatry Res* 1991;39:99–108.
- 93 Nofzinger EA, Schwartz RM, Reynolds CF 3rd, Thase ME, Jennings JR, Frank E, Fasiczka AL, Garamoni GL, Kupfer DJ: Affect intensity and phasic REM sleep in depressed men before and after treatment with cognitive-behavioral therapy. *J Consult Clin Psychol* 1994;62:83–91.
- 94 Pressman SD, Cohen S: Does positive affect influence health? *Psychol Bull* 2005;131:925–971.
- 95 Chida Y, Steptoe A: Positive psychological well-being and mortality: a quantitative review of prospective observational studies. *Psychosom Med* 2008;70:741–756.
- 96 Cosci F, Pennato T, Bernini O, Berrocal C: Psychological well-being, negative affectivity, and functional impairment in fibromyalgia. *Psychother Psychosom* 2001;80:256–258.
- 97 Lipowski ZJ: Physical illness and psychopathology. *Int J Psychiatry Med* 1974;5:483–497.
- 98 Fava GA, Fabbri S, Sirri L, Wise TN: Psychological factors affecting medical condition: a new proposal for DSM-V. *Psychosomatics* 2007;48:103–111.
- 99 Sirri L, Grandi S: Illness behavior. *Adv Psychosom Med* 2012;32:160–181.
- 100 Mechanic D: Social psychological factors affecting the presentation of bodily complaints. *N Engl J Med* 1972;286:1132–1139.
- 101 Sirri L, Grandi S, Fava GA: The Illness Attitude Scales: a clinimetric index for assessing hypochondrial fears and beliefs. *Psychother Psychosom* 2008;77:337–350.
- 102 Kellner R: *Somatization and Hypochondriasis*. New York, Praeger, 1986.
- 103 Pilowsky I: *Abnormal Illness Behavior*. Chichester, Wiley, 1997.
- 104 Weinman J, Petrie KJ, Moss-Morris R, Horne R: The Illness Perception Questionnaire: a new method for assessing the cognitive representations of illness. *Psychol Health* 1996;11:431–445.
- 105 Fava GA, Kellner R: Staging: a neglected dimension in psychiatric classification. *Acta Psychiatr Scand* 1993;87:225–230.
- 106 Tomba E: Nowhere patients. *Psychother Psychosom* 2012;81:69–72.
- 107 Lichtenberg P, Belmarker RH: Subtyping major depressive disorder. *Psychother Psychosom* 2010;79:131–135.
- 108 Bech P: Struggles for subtypes in primary and secondary depression and their mode-specific treatment or healing. *Psychother Psychosom* 2010;79:331–338.
- 109 Fava GA, Sonino N, Wise TN (eds): *The Psychosomatic Assessment*. Basel, Karger, 2012.
- 110 Taylor GJ: Affects, trauma, and mechanisms of symptom formation: a tribute to John C. Nemiah, MD (1918–2009). *Psychother Psychosom* 2010;79:339–349.
- 111 Lemogne C, Nabi H, Zins M, Cordier S, Ducimetiere P, Goldberg M, Consoli SM: Hostility may explain the association between depressive mood and mortality. *Psychother Psychosom* 2010;79:164–171.
- 112 Sensky T: Chronic embitterment and organizational justice. *Psychother Psychosom* 2010;79:65–72.