

The concept of psychiatric illness – a core problem in psychiatric epidemiology

Inger Sandanger^{1,3}, Jan F. Nygård¹ and Tom Sørensen²

¹Department of Behavioural Sciences in Medicine, University of Oslo

²Institute Group of Psychiatry, University of Oslo

³National Insurance Administration

Correspondence: Inger Sandanger, E-mail: inger.sandanger@basalmed.uio.no

ABSTRACT

The basic problem in psychiatric epidemiology is the ontology, to establish what the nature of psychiatric illness, disease or disorder is. Psychiatric epidemiology operates with two main concepts of psychiatric illness: the overload concept and the criteriological concept. From these two different ideas of what a psychiatric illness is, follow varying consequences for studies, clinical work and administrative purposes. Differences between screening instruments and diagnostic tools in their concepts and operationalisations result in different research results, such as varying gender differences. The screener and the diagnostic instrument should both be tested against other illness indicators. A continuous screener can be interesting as a measure of level of symptoms, but not a good measure of the probability for diagnosis. We need in fact several definitions of psychiatric disorders; for clinical intervention use, for different types of research and for administrative/bureaucratic purposes. The importance of classification depends on the purpose.

NORSK SAMMENDRAG

Det grunnleggende problem i psykiatrisk epidemiologi er ontologien, den psykiatriske lidelses natur. Psykiatrien opererer med to hovedbegreper for psykiatrisk lidelse: overbelastningsbegrepet og kriteriebegrepet. Fra disse to forskjellige idéene om psykisk lidelse springer det ut ulike konsekvenser for forskningsstudier, klinisk arbeid og administrative formål. Forskjeller grunnet i begrepene og operasjonaliseringene for screener og diagnostisk instrument resulterer i ulike forskningsresultater, blant annet varierende kjønnsforskjeller. Screener og diagnostiske instrumenter bør begge prøves ut mot andre sykdomsindikatorer. En kontinuerlig screener kan være interessant som mål på nivå av plager, men ikke så god som sannsynlighet for diagnose. Vi trenger egentlig flere definisjoner for psykiske lidelser; for klinisk behandling, forskjellige forskningsformål og for administrative formål. Betydningen av klassifikasjon er avhengig av formålet.

INTRODUCTION

"The wit of man has rarely been more exercised than in the attempt to classify the morbid mental phenomena covered by the term insanity. The result has been disappointing". D.H. Tuke 1893.

The basic problem in psychiatric epidemiology is the ontology, to establish what the nature of psychiatric illness, disease or disorder is.

Is psychiatric illness in fact a detectable, objective entity, qualitatively different from general human experiences, an event that happens to a person, like a bone fracture, or is it a detectable, fluctuating state, which we call illness above a certain level or quantity, like temperature and fever?

Who defines psychiatric illness, the person who experiences it, the medical profession, or the society? Is psychiatric illness subjective or objective? Must there be pain or feeling of illness or discomfort? Should it be observable, and by whom? Does there have to be dysfunction or impairment in any way,

behavioural or physiological? Ought there be problems with sociocultural adaptation? Value choices of these questions will structure the attention, the foreground, the background and the clusters of illness we see, and subsequently the whole nosology. Nosologies and hence the understanding of diagnoses will be formed by the theories of the time, logic and philosophy of what psychiatric illness is. As such, diagnoses are based on theories and are not atheoretical. Consequently there will be scientific, value related and practical repercussions of every nosology and diagnosis. It influences the clinical thinking, the scientific study design and the thinking among third parties, such as insurance companies, private or public. Who or what does a nosology or specific diagnosis serve the most, the need of a society who wants disturbed people out of sight, the medical industry, who will sell more the wider the definition of illness is, the need of systematic registers for the purpose of order and statistics, the researcher with variable needs, or the clinician trying to help her/his patient?

Today, psychiatric epidemiology operates with two main concepts of psychiatric illness. The first is the overload concept, indicating that we all have nervous symptoms, which will be called illness when the symptoms exceed a certain level, a nominalistic approach. The question is "How much of it has s/he got?" (the temperature and fever analogy). The second alternative concept is the criteriological, essentialistic view that people either have or do not have disorders. The disorder exists as a particular construct of symptoms in a person, as discussed above. It corresponds to the question "Has s/he got it?" a question more easily asked about a fracture. Operationalisation is a result of the chosen concept of illness. From the two ontologies of psychiatric disorder follow two operationalisations, one continuous score and one yes/no box definition.

AIM

The intention of this article is to discuss how dependent the empirical results in epidemiological studies are upon the concepts and operationalisations that have been used. From two different ideas of what a psychiatric illness is, we will briefly discuss some of the varying consequences for studies, clinical work and administrative purposes (health insurance). The differences between screening instruments and diagnostic tools in their concepts and operationalisation will be discussed, as well as how analysis of gender differences in psychiatric illness is influenced by the two different concepts and their operationalisations.

SOME HISTORY OF THE CONCEPTS OF PSYCHIATRIC ILLNESS

It is conspicuous how the methods follow from the prevailing conception of psychiatric illness. In the studies before 1940, which Dohrenwend calls the first generation epidemiology, most studies of occurrence were done by agency records and from information from key informants (doctors and other sources) [1-6]. This reflected a firm belief in genetical and constitutional origins of psychiatric disease, a medical condition [7].

In the post war period, the social origins of psychiatric disease came into focus, partly as a result of the development of the social sciences. Methods for random sampling and surveys became available. The interest turned to psychosocial etiology and intervention. Nervous illness was seen as a variety of nuisances or disturbances. Specific diagnoses were not considered as important as finding "cases" or persons with "impairment", for psychosocial interventions and prevention. Low diagnostic reliability among psychiatric specialists accelerated the development of caseness instruments. The evaluation of soldiers during the war had enforced the development of brief instruments for evaluating mental health [8-10]. When additional proof was wanted, psychiatrists interviewed individuals di-

rectly to diagnose disease, or they reviewed protocols from the interview. The caseness was the important outcome and reflected a high degree of anarchy in diagnostic thinking within the psychiatric profession. Much effort was spent on evaluating the respondents with regard to symptom pattern, impairment, caseness probability and need for psychiatric attention [11], and Leighton et al.'s work started the bridging to diagnostic instruments.

The introduction of the Diagnostic and Statistical Manual of Mental Disorder revision III (DSM-III) in 1980 started a mission to unite psychiatry into one practice of diagnosing. The DSM-III aimed to be purely descriptive in its classification of psychiatric disorders. The attention to the origin of illness turned away from a psychosocial and psychodynamic understanding towards a neurobiological model. This coincided with the advances in psychopharmacological and neurobiological research and brought about a renewed interest in genetic studies and personality. Psychiatric disorders were increasingly regarded as reflections of neurobiological imbalance or disturbance. The third generation of epidemiological studies has therefore focused on specific diagnostics. Among instruments developed are the Research Diagnostic Criteria [12], the Diagnostic Interview Schedule (DIS) [13], the Present State Examination (PSE) [14], the Structural Clinical Interview for DSM-III (SCID) [15], the Composite International Diagnostic Interview (CIDI) [16], and the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) [17].

The introduction of descriptive diagnoses with the DSM-III, and its followers, and more recently the International Classification of Disease version 10, (ICD10), has made diagnostic practice more universal. However, philosophers and others interested in the concept of mental illness have criticised the diagnoses on the grounds of tacit presuppositions in both scientific theory and concepts [18].

SCREENING AND DIAGNOSIS

The diagnosis

In the following we will discuss the concept of psychiatric diagnosis with regard to the present descriptive psychiatric diagnostic system, which in principle does not consider the etiology or the prognosis of the disease. Most doctors or psychiatrists feel that they recognise a depression when they see a person who is depressed. This recognition is based on a scientific pluralism used in a holistic approach. The doctor looks for a prototype of a disorder (descriptive diagnostics), may take some blood tests (biological theory), relates perhaps to transference and the interaction between the patient and her/himself (psychoanalytical theory), explores the functioning within the family system (psychodynamic and systems theory), and questions the work environment and the role of the person in the

local community (social psychiatry). Most of these domains have been affected by the time the patient reaches the psychiatric service. What the doctor really does is to put all available information together, i.e. the art of diagnosing [19]. The doctor sets a diagnosis based on the criteria of the ICD10 system, but the validity of these criteria is in daily practice probed by the phenomenological information s/he gathers about the patient in addition to the diagnostic criteria. As such, the diagnostic criteria acquire high validity in clinical settings. However, the diagnostic criteria themselves constitute only some of the many dimensions that reflect the psychiatric disorder. We indulge in limited operational definitions based on consensus and compromises in the opinions and experiences of professionals.

Screening

Screening is conceptually a test of asymptomatic persons in order to diagnose illness at an early stage. Positive screenees are submitted to diagnostic procedures to see if they are actually ill, and thereafter to treatment to prevent the development of more severe illness at an early stage, when it is easier to treat or cure.

Usually a screener is regarded as a fast way to sort out cases from non-cases [20]. Famous studies like The Midtown Manhattan Study (USA) and the Stirling County study (Canada) have used different screeners [8,11,21]. The estimates by screener and diagnostics have often been comparable in size [22], but have not picked out all the same persons. A Swedish study tested the Hopkins Symptoms Check List-25 items (HSCL-25) using the Present State Examination (PSE-9) as a caseness criterion and found a positive predictive value of 58% in a sample of general practice patients [23]. Dohrenwend, however, has criticised efforts to demonstrate criterion-oriented, concurrent validity, such as Nettelbladt did, on the basis that no generally agreed upon criteria of psychological health or disorder exists, and that attempts to demonstrate concurrent validity therefore are inappropriate [24,25]. A screener may also be used so that a few criteria have to be fulfilled in order to get a full examination. This is done by the Composite International Diagnostic Interview, CIDI [26], but most often the screener is a rating scale of some sort, for number and/or intensity of symptoms, which yields a continuous score, and where a certain value has been chosen as cut-off for dichotomisation into cases (those who will have the full diagnostic procedure) and non-cases (those who will not).

Relation between screener and diagnosis

When using screener and diagnostic instrument rooted in two different conceptions or ontologies of psychiatric disorders (the continuous score “how much of it does s/he have”, and the criteria “does s/he have it?”), it would be surprising if the two case finders should demonstrate a high degree of agreement. The screener

and the diagnostic instrument should both be tested against other illness indicators, such as reduced functioning in work, hobbies or daily doings, use of prescription medication, use of any health service for the condition, both medical and paramedical, and sick-leave, in the interest of pursuing the search for a meaningful case concept in epidemiology [27,28].

The curve between the overload concept screener and criteriologial diagnosis by a given system can have the form of a linear curve and the screener can be used as a *severity measure*: the higher the score, the more severe the problem, useful to compare means (fig. 1). It can also be used as a *probability measure*; if the symptom score is above a certain value, the person is probably a case. When reaching a certain threshold, the probability of being a case rises dramatically. Figure 1 shows the theoretical shape of a severity curve and a probability curve for a continuous screener, where 15% of the population are presumed to be diagnostic cases with a screener with cut-off of 1.75. On the ordinate is the prevalence (percentage) of diagnostic non-cases for each screener cut-off value. On the abscissa are the possible cut-off values by the screener. If the cut-off is set to 1 there will be zero non-cases, all will be cases. If the cut-off is set to 4, 100% will be non-cases. At cut-off 1.75, 85% of the population are non-cases and 15% are cases, by both curves. At cut-off 1.50, the prevalence of non-cases is about 60% on the severity curve, that is 40% are cases. On the probability curve only a few percent are cases, which shows that such a low cut-off is meaningless if the screener behaves as probability curve because it will include almost everybody. However, it is meaningful to measure the percentage of the population that will have a symptom score (severity) of 1.50.

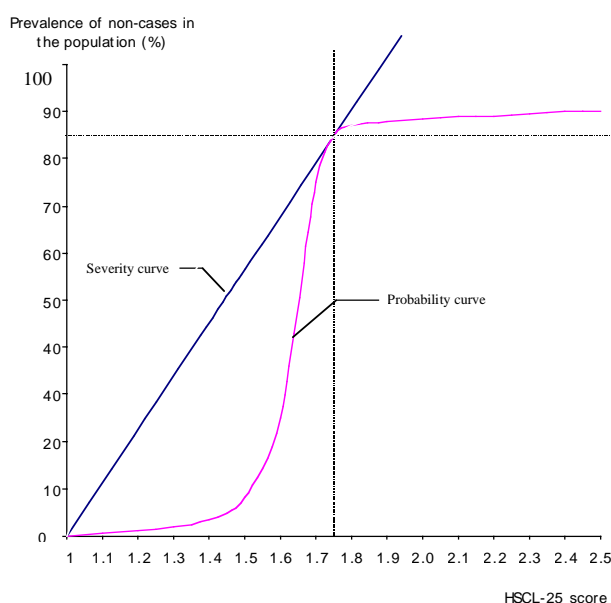


Figure 1. Severity curve and probability curve, where 15% of the population a priori are presumed to be cases with a cut-off of 1.75.

Using the probability model is an approximation to an essentialistic comprehension of psychiatric disorder and can be viewed as a compromise between the nominalistic and essentialistic concept. The important thing is to know the form of the relationship between the screener (if it is a continuous score, a rating scale) and the percentage that have a diagnosis. If it has a probability curve shape, it is a good screener, if it is more like a severity linear curve, it is best used as a measure in its own right.

Receiver operating curves (ROC), sensitivity, specificity and positive predictive value give information about agreement of screener and diagnostics, but plotting the curve gives a good visualisation of how the screener functions. If using a screener to find cases, it should be tested out that the relation makes a probability curve.

GENDER

A study by Broverman showed that clinicians agreed in their descriptions of a "healthy man" and a "healthy adult independent of sex", giving them equal features, but that these were different from their description of a "healthy woman". Women were described in terms which would be considered negative for men or "healthy adults independent of sex", such as submissiveness, dependency etc. [29,30]. There were obviously different concepts for a healthy woman and a healthy man. If characteristics frequently found in women, and which are perceived as negative or indications of psychopathology in men, are used as a concept of psychopathology, one is bound to find more pathology among women. How much of the gender difference can be ascribed to definitions which take female symptoms more into consideration than male symptoms? The definition of depression in the Stirling County Study of 1952 and 1970 had dysphoric mood as an essential feature, as well as disturbances of sleep, appetite and energy, in conjunction with at least a mild degree of impairment in everyday functioning, persisting over a minimum duration. Using the data from 2015 randomly selected persons in a Norwegian population study [31], the authors used the Stirling County definition and took out the Hopkins Symptoms Checklist-25 items [32,33] "feeling blue", "difficulties falling asleep/staying asleep", "poor appetite", "feeling low in energy, slowed down" and "feeling everything is an effort", and created a new "case" for depression comprising those who had these symptoms. The prevalence of depression with this type of caseness (score ≥ 1.75) was 6.3% for men and 8.7% for women, which did not differ significantly. On the other hand, the prevalence of HSC-25 ≥ 1.75 cases and the prevalence of IDC10 depression cases were both significantly and twice as high for women than men [31]. In other words, choice of items is important for gender prevalence difference [34] and rises the question about different weightings of items in screeners and diagnostic criteria for women

and men. At the bottom of this question are the concepts of disorder, which lead to the item choices.

VALIDITY AND PURPOSE OF CLASSIFICATION

From two different ideas of what a psychiatric illness is, we will briefly discuss some of the varying consequences for studies, clinical work and administrative purposes (health insurance). These problems are not limited to the field of psychiatry, but concerns most areas in medicine. Examples of this are the problems with establishing the diagnosis of gastric ulcer [35] or deciding when a cell with changed cell morphology is a cancer cell [36]. The present discussion is limited to psychiatric diagnoses.

If we use a set of diagnostic criteria on persons from the general population, based on a concept of psychiatric disorder derived from the clinic, and without getting any of the additional, phenomenological information of the clinical setting, a high validity cannot be expected compared with the clinical setting. Some individuals will not fulfil the criteria, but still be quite ill or disturbed. Others, who fulfil the criteria for the diagnosis, will not show any reduction in functioning and will not seek more help than healthy persons. Ill persons in the population may show different adaptations, which keep them going without therapy, and which referred persons does not have.

In many psychiatric studies, results from screening instruments are published as occurrences of illness, prevalence. When doing this, the authors may refer to a high agreement with diagnoses, tested in a clinical population. However, the group of persons who form a clinical population at for instance an out-patient clinic, has such a number and variety of symptoms that most short instruments will show high agreement, because patients are selected by the symptoms we want to test. The meaning of findings by a screener must be thoroughly discussed in the light of the authors' understanding of psychiatric illness. If the purpose of an epidemiological study were to find how many persons needed treatment, should individuals without impairment be included? The answer 'no' seems very likely [37,38]. If we were doing etiological research on the other hand, it seems obvious that all cases should be included by definition. We need in fact several definitions of psychiatric disorders; for clinical intervention use, for different types of research and for administrative/bureaucratic purposes (health insurance registers etc.) [39-41]. Each different purpose requires its own evaluation systematically designed and adapted to the situation, or to the context. The clinician need a diagnosis made of all available information about the patient and his illness, as mentioned above, in order to choose the best treatment for the single, given patient. The National Insurance Administration bases the amount of pension or sickness benefit on the diagnosis and the reduction in functioning that it leads to. §8-4 in the Norwegian National Insurance Scheme (Folke-

trygdloven) states that "Sickness benefit is given to the person who is disabled to work because of a reduced functioning that clearly is caused by illness or accident". Disability pension is given when "The medical illness must have led to a permanent reduction in function of such a character and degree that it is the main cause for the reduction of income/workability" (§12-6): "... by judgement of how much the income is reduced (how high is the degree of disability), possibilities for income in all kinds of work that the person can now do, must be compared to the possibilities for income that the person had before the illness/accident." The diagnosis must accordingly be the cause of reduced income in all available work for the patient.

In the health insurance statistics one counts sickness benefits and disability pensions by diagnoses. If occurrence of an illness is counted by diagnoses in the sickness benefit registers, you get only those who both are ill and cannot work. The health insurance would have benefited from a diagnostic system that included functioning and prognosis of the illness, not only the descriptive, criteriological diagnoses. For the purpose of disability pension, one would have liked to have a prognosis attached to the diagnosis. In the same manner compensation for occupational hazards is given by diagnosis where such a hazardous etiology is established for that diagnosis in general. For research purposes one would like a diagnosis that would catch homogenous groups for natural course or comparison of interventions and study the causes for different outcomes, like function reduction.

Dependent on the purpose of the diagnosis: to institute a treatment, to give a health insurance benefit or to study the occurrence, causes and prognoses of illnesses, the diagnostic knowledge must be different. Context for the use of diagnostics is taken into account.

The importance of context was most clearly advocated by Kurt Lewin who insisted that behavioural phenomena must be understood in terms of an appre-

ciation of the immediate and dynamic causal matrix in which behavioural events are embedded [42], condensed in his formula $f(b)=p \times e$ (behaviour is a function of the person and the environment). Even in a pure biological framework for understanding of psychiatric disorders, there are problems with explaining the pathway from genes to the complex phenomena of behaviour if context is not taken into consideration [43,44] or to understand what causes changes in the brain function by plasticity and alternative signal routes [45,46].

A classification for all purposes would indeed be convenient, but it is based on a static model of psychiatric disorder, independent of context, which does not exist. To quote Derogatis et al. in their introduction to the Hopkins Symptom Check List (HSCL) [47]: "It has been shown repeatedly that questions concerning the validity of a scale only have meaning with regard to specific interpretative decisions, i.e. *whether the scale is or is not valid for a specific purpose*, and that the nature of the decision shapes the particular type of validity that has relevance for the question at hand".

EPILOGUE

The importance of classification depends on the purpose. Brown stated in 1976 "It does not matter to me whether depression is viewed as a disease or as a way of adapting. It is sufficient that almost over the whole world it is experienced as troublesome" [48]. Psychiatric epidemiology has so far often been too deterministic in its thinking: a case or not a case. The important debate is about the limitations and problems in defining *clinical significance* in epidemiological studies [49-52]. The development of psychiatric epidemiology is definitely not only about measuring the occurrence and relationships of psychiatric disorders, but very much about what to measure. If we do not know what the illness is, how can we study its causes or its prognosis with or without intervention/treatment?

REFERENCES

1. Bök JA. A genetic and neuropsychiatric investigation of a North-Swedish population with special regards to schizophrenia and mental deficiency. *Acta Genet* 1953; **4**: 1-100.
2. Fremming KH. The expectation of mental infirmity in sample of the Danish population. Occasional Papers on Eugenics, Vol. 7. London: Cassel, 1951.
3. Ødegaard Ø. The distribution of mental disorders in Norway. *Acta Psychiatr Neurol* 1945; **20**: 247-284.
4. Ødegaard Ø. A statistical investigation of the incidence of mental disorder in Norway. *Psychiatr Q* 1946; **20**: 381-399.
5. Ødegaard Ø. The incidences of mental diseases as measured by census investigations versus admission statistics. *Psychiatr Q* 1952; **26**: 212-218.
6. Strømgren E. Statistical and genetical population studies within psychiatry. Methods and principal results. Proceedings of the 1st International Congress of Psychiatry, Paris, 1950: 155-189.
7. Kraepelin E. *Psychiatrie: ein Lehrbuch für Studierende und Ärzte*, 5. Auflage. Leipzig: Barth, 1896.
8. Langner TS. A twentytwo item screening score of psychiatric symptoms indicating impairment. *J Health Hum Behav* 1962; **3**: 269-276.

9. Macmillan AM. The health opinion survey: technique for estimating prevalence of psychoneurotic and related types of disorders in communities. *Psychol Rep* 1957; **3**: 325-329.
10. Star SA. The screening of psychoneurotics in the army: Technical development of tests. In: Stouffer SA et al., eds. *Measurement and prediction*. Princeton: Princeton University Press, 1950: 486-547.
11. Leighton DC, et al. The character of danger. *Psychiatric symptoms in selected communities. The Stirling County study of psychiatric disorder and sociocultural environment, Vol. III*. New York: Basic Books, 1963.
12. Feighner JP, et al. Diagnostic criteria for use in psychiatric research. *Arch Gen Psychiatry* 1972; **26**: 57-63.
13. Robins LN, et al. The National Institute of Mental Health Diagnostic Interview Schedule: Its history, characteristics, and validity. *Gen Psychiatry* 1981; **38**: 381-389.
14. Wing JK, Cooper JE, Sartorius N. *The description and classification of psychiatric symptoms: An instruction manual for the PSE and CATEGO systems*. London: Cambridge University Press, 1974.
15. Spitzer RL, et al. *Structured clinical interview for DSM-III-R (SCID)*. New York: New York State Psychiatric Institute, Biometrics Research Department, 1987.
16. Robins LN, et al. The Composite International Diagnostic Interview. *Arch Gen Psychiatry* 1988; **45**: 1069-1077.
17. Wing JK, et al. Schedules for clinical assessment in neuropsychiatry. In: Mezzich JE, Jorge MJ, Salloum IM, eds. *Psychiatric epidemiology. Assessment concepts and methods*. Baltimore: Johns Hopkins University Press, 1994: 268-280.
18. Sadler JZ, Wiggins OP, Schwartz MA, eds. *Philosophical perspectives on psychiatric diagnostic classification*. (The Johns Hopkins Series in Psychiatry and Neuroscience, McHugh PR, Johnson RT, eds.). Baltimore: Johns Hopkins University Press, 1994.
19. Miettinen OS. Evidence in medicine: invited commentary. *Can Med Assoc J* 1998; **158**: 215-221.
20. Duncan-Jones P, Henderson S. The use of a two phase design in a prevalence survey. *Soc Psychiatry* 1978; **13** (4): 231-237.
21. Macmillan AM. A survey technique for estimating the prevalence of psychoneurotic and related types of disorders in communities. *Epidemiol Ment Dis* 1959: 203-228.
22. Sørensen T. Commuting, community integration and psychiatric disorder. In: Sørensen T, Abrahamsen P, Torgersen S, eds. *Psychiatric disorders in the social domain. Contributions to a scientific basis for understanding and action*. Oslo: Norwegian University Press, 1991: 41-49.
23. Nettelbladt P, et al. Test characteristics of the Hopkins Symptom Check List-25 (HSCL-25) in Sweden, using the Present State Examination (PSE-9) as a caseness criterion. *Soc Psychiatry Psychiatr Epidemiol* 1993; **28**: 130-133.
24. Dohrenwend BP, et al. Some problems of validity with the Psychiatric Status Schedule as an instrument for case identification and classification in the general population. *Arch Gen Psychiatry* 1980; **37**: 720-721.
25. Dohrenwend BP. The problem of validity in field studies in psychological disorder. In: Mezzich JE, Jorge MJ, Salloum IM, eds. *Psychiatric epidemiology. Assessment concepts and methods*. Baltimore: Johns Hopkins University Press, 1994: 201-223.
26. Wittchen U-H. Epidemiological research in mental disorders: lessons for the next decade of research – the NAPE Lecture 1999. *Acta Psychiatr Scand* 2000; **101**: 2-10.
27. Bolton D. Philosophy of mind and psychiatry. *Curr Opin Psychiatry* 1998; **11**: 563-566.
28. Fulford KWM. History and Philosophy. *Curr Opin Psychiatry* 1998; **11**: 541-543.
29. Broverman IK, et al. Sex-role stereotypes and clinical judgements of mental health. *J Consult Clin Psychol* 1970; **34** (1): 1-7.
30. Broverman IK, et al. Sex-role stereotypes: A current appraisal. *J Soc Issues* 1972; **28** (2): 59-78.
31. Sandanger I, et al. Prevalence, incidence and age of onset of psychiatric disorders in Norway. *Soc Psychiatry Psychiatr Epidemiol* 1999; **34**: 570-579.
32. Derogatis LR, et al. The Hopkins symptom checklist (HSCL): A self-report symptom inventory. *Behav Sci* 1974; **19**: 1-15.
33. Winokur A, et al. Symptoms of emotional distress in a family planning service: stability over a four-week period. *Br J Psychiatry* 1984; **144**: 395-399.
34. Murphy JM. What happens to depressed men? *Harvard Rev Psychiatry* 1995; **3**: 47-49.
35. Bernersen B, et al. Erosive prepyloric changes in dyspeptics and non-dyspeptics in a defined population. The Sorreisa Gastrointestinal Disorder Study. *Scand J Gastroenterol* 1992; **27** (3): 233-237.
36. Fahey MT, Irwig L, Macaskill P. Meta-analysis of Pap test accuracy. *Am J Epidemiol* 1995; **141** (7): 680-689.
37. Richters JE, Hinshaw S. Psychiatry's turbid solution. *Clin Psychol* 1993; **3**: 276-280.
38. Wakefield JC. Disorder as harmful dysfunction: A conceptual critique of DSM-III-R's definition of mental disorder. *Psychol Rev* 1992; **99**: 232-247.
39. Grant DC, Harari E. Diagnosis and serious mental illness. *Aust N Z J Psychiatry* 1996; **30**: 445-449.
40. Karls JM, Wandrei KE. PIE: A new language for social work. *Soc Work* 1992; **37** (1): 80-85.

41. Kringlen E. Is the concept of schizophrenia useful from an aetiological point of view? A selective review of findings and paradoxes. *Acta Psychiatr Scand* 1994; **90** (Suppl. 384): 17-25.
42. Lewin K. The conflict between Aristotelian and Galileian modes of thought in contemporary psychology. In: *A dynamic theory of personality: Selected papers*. New York: McGraw-Hill, 1935: 1-42.
43. Carson RC. Aristotle, Galileo, and the DSM taxonomy: The case of schizophrenia. *J Consult Clin Psychol* 1996; **64** (6): 1133-1139.
44. Dohrenwend BP, Dohrenwend BS. Social and cultural influences on psychopathology. *Annu Rev Psychol* 1974; **25**: 417-452.
45. Greenfield S. *The private life of the brain*. London: Allen Lane, 2000.
46. Wakefield JC. The concept of mental disorder: On the boundary between biological facts and social value. *Am Psychol* 1992; **47** (3): 373-388.
47. Derogatis LR, Lipman RS, Rickels K. The Hopkins Symptom Checklist (HSCL): A measure of primary symptom dimensions. In: Pichot P, ed. *Psychological measurements in psychopharmacology – modern problems in pharmacopsychiatry*. Basel: Karger, 1974: 79-110.
48. Brown GW. *Social causes of disease*. London: Tavistock Publications, 1976: 291-333.
49. Frances A. Problems in defining clinical significance in epidemiological studies. *Arch Gen Psychiatry* 1998; **55**: 119.
50. Regier DA, et al. Limitations of diagnostic criteria and assessment instruments for mental disorders. Implications for research and policy. *Arch Gen Psychiatry* 1998; **55** (2): 109-15.
51. Spitzer RL. Diagnosis and need for treatment are not the same. *Arch Gen Psychiatry* 1998; **55**: 120.
52. Miettinen OS. The modern scientific physician: 7. Theory of medicine. *Can Med Assoc J* 2001; **165** (10): 1327-1328.