REVIEW

The possibility of evidence-based psychiatry: depression as a case

**Drozdstoy Stojanov · Jakob Korf · Peter de Jonge ·
Georgi Popov**

Abstract Considering psychiatry as a medical discipline, a diagnosis identifying a disorder should lead to an effective therapy. Such presumed causality is the basis of evidence-based psychiatry. We examined the strengths and weaknesses of research onto the causality of relationship between diagnosis and therapy of major depressive disorder and suggest what could be done to strengthen eventual claims on causality. Four obstacles for a rational evidence-based psychiatry were recognised. First, current classification systems are scientifically nonfalsifiable. Second, cerebral processes are—at least to some extent—nondeterministic, i.e. they are random, stochastic and/or chaotic. Third, the vague or lack of relationship between therapeutic regimens and suspected pathogenesis. Fourth, the inadequacy of tools to diagnose and delineate a functional disorder. We suggest a strategy to identify diagnostic prototypes that are characterised by a limited number of parameters (symptoms, markers and other characteristics). A prototypical diagnosis that may either support or reject particular elements of current diagnostic systems. Nevertheless, one faces the possibility that psychiatry will remain a relatively weak evidence-based medical discipline.

D. Stojanov (✉)
State Hospital for Mental Disorders 'St. Ivan Rilski',
Medical University - Plovdiv,
Plovdiv, Bulgaria
e-mail: stojanovpisevski@gmail.com

J. Korf · P. de Jonge
University Psychiatry Center (UCP),
Groningen, The Netherlands

J. Korf
e-mail: j.korf@psy.umcg.nl

G. Popov
Varna University Hospital 'St.Marina' Clinic,
'General Psychiatry and Addicitons',
Varna University of Medicine,
Varna, Bulgaria

J. Korf
University Medical Centre Groningen (UMCG),
P.O. Box 30.001, 9700 RB, Groningen, The Netherlands

Introduction

In their analysis on the clinical significance of randomised controlled trials, Kraemer and Kupfer (2006) advocated the 'number needed to treat' (NNT) and some other parameters to convey the clinical and practical significance of the outcome of a trial. They indicate that such parameters may deviate largely, so NNTs ranging from 2 to 100 might be relevant depending on the context. It means that one may treat 2 up to 100 subjects to establish the desired effect in a single person. High NNT may point either to weak diagnostic criteria for a particular therapy or to an ineffective intervention. In their paper on the usefulness of the medical model in psychiatry, Shah and Mountain (2007) define 'the "medical model" as a process whereby informed by the best available evidence, doctors advise on, coordinate or deliver interventions for health improvement. It can be summarily stated as "does it work"?' This definition of a diagnosis refers to utility (usefulness in a particular context), rather than to validity (a true aspect of the real world; Kendell and Jablensky 2003). The medical model assumes causality, and that is the basis of rational and evidence-based medicine. The concept of (strict) causality in medicine may originate in classical physics, where force and its consequences are understood as deterministic processes. In medical sciences, the idea of causality is most often expressed in terms of restoration of a steady-state condition. Pathological processes lead to a deviation from

the steady state, and the intervention is applied by the medical doctor to eliminate the pathological state. So the biological basis of recovery from a medical disorder is the capacity of an organism to restore and maintain its internal milieu or homeostasis. A reversible disorder (or disease) points to the capacity of an organism to eliminate or compensate the pathology, whereas a chronic course is indicative of a failing capacity.

The issue we address here is the empirical strength of causal relationships in psychiatry. We consider psychiatry primarily as a medical discipline and a psychiatric disorder as an undesired condition that a medical doctor should help to eliminate. A diagnosis should then recognise a disorder and the therapy proposed should cure the patient, i.e. let disappear the disorder (or at least helps to have a decent life). Psychiatric disorders are primarily concerned with the brain and their manifestation may therefore be more variable than that of somatic disorders, because of the higher complexity of the brain. Such complexity may be greater than in neurological disorders. The clinical consequences, including the time course, of the latter disorders are primarily determined by underlying cellular processes. In contrast, organic pathology is absent or relatively insignificant so far in most psychiatric conditions, in particular, in the so-called functional disorders, including depression. We will consider the issue whether functional psychiatric disorders and, in particular, major depressive disorder can be conceptualised in the conventional framework of strict causality and steady-state dynamics.

The term evidence-based medicine alludes to a firm scientific basis of the medical practice. Often, the evidence is based on clinical trials comparing two or more interventions, including placebo treatment, showing statistically significant different outcomes. The issue is whether a thus defined efficacy allows predicting the outcome of a therapy in the individual patient. An NNT=1 indicates strict causality. The psychiatrist has to propose (or perhaps to decide) which treatment is most adequate for the patient sitting in his office or lying in the hospital bed. This is not meant as a rhetoric question; we acknowledge that many treatments are effective in the individual patient. In psychiatry (perhaps also in other disciplines of medicine, but that is not our concern), however, there are also treatments that are—at least in large cohorts—a little more effective than placebo. One may, for instance, refer to the recent discussion on antidepressant drugs (Kirsch et al. 2008; Turner et al. 2008; Moncrieff and Kirsch 2005). If an active treatment is hardly better than placebo, it is impossible to make a rational decision about the treatment of the individual patient: this is in fact challenging the assumption of causality.

Our central question is: is strict causality possible in psychiatry? We chose major depressive disorder as a representative functional disorder, because of the availability of illustrative data. We discuss four obstacles to achieve strict causality. One argument is the lack of scientific basis of current diagnostic systems and constructs. We challenge the possibility of effective diagnostic and therapeutic interventions with current classification systems. Another obstacle is the nature of cerebral functioning: we argue that brain processes are not as deterministic as often thought. In line with this contention, we describe nondeterministic time courses of recovery from depression that differ fundamentally from that of reversible somatic disorders. Finally, we emphasise that current disease markers or psychological inquiry does not necessarily provide the information required for rational therapeutic strategies. The scope of the present mini-review differs from that of recent related papers, where the philosophical framework and utility or validity of psychiatric diagnosis (Shah and Mountain, 2007; Kendell and Jablensky 2003; Kendell 2001; Kendler 2005; Angst 2007; Jablensky 2005) were discussed. Parts of this paper were recently presented elsewhere (Korf 2008; Stoyanov et al. 2008).

Current diagnostic approaches

One of the most rational frameworks in psychiatry was Sigmund Freud's psychoanalytic approach. He construed a closed diagnostic system leading to a personally dedicated treatment. Despite its apparent rationality, the psychoanalytic approach has limited therapeutic efficacy. Emil Kraepelin is the founder of the present diagnostic approaches and his ideas are the basis of current classification proposals (Shorter 1997). Subsequently a variety of diagnostic systems, including the Diagnostic and Statistical Manual of Mental Disorders (DSM) and the International Classification of Diseases (ICD), were introduced. These systems were developed as a reaction to the psychoanalytic framework and were aimed not only to serve the clinician with a standardised terminology but also to provide a scientific basis of psychiatry (Shorter 1997). It was well realised from the onset onwards that any classification needs a firm scientific basis. The current version of the DSM distinguishes nearly 300 disease entities. Many investigators consider such number as exuberant, but there is as yet no scientific argumentation whether such number is too much, too little or just perfect. The arguments are rhetoric instead: do we need also 300 diagnoses of liver diseases? A way to avoid such fruitless discussions would be to assume that the DSM classification does not define disease entities, but defines (psychological) disorders. But then the question arises: what should be considered as a disorder and what not, and do nonpsychological parameters ever become relevant? The DSM (for instance) allows some

diagnostic variations, as some symptoms are more important for the diagnosis of major depressive disorder (code 296xx) than others: duration of a depressed mood for at least 2 weeks is mandatory, but the presence of only four of the following symptoms is required: disturbed appetite; weight; sleep; psychomotor activity; energy; feelings of worthlessness or guilt; ability to think, concentrate or make decisions; recurrent thought of death or suicidal ideation, plans or attempts. So about 170 different combinations of symptoms fulfill the criterion major depressive disorder.

We now propose a thought—an experiment challenging a classification system and so illustrating its scientific status. Assume that current classification systems are no *ex cathedra* formulated diagnostic standards (which they are to a large extent) but rather as a (possibly pre-scientific) hypothesis that may be falsified in the Popperian sense. Then the question arises when is a parameter sufficiently strong to falsify the DSM? Consider a parameter investigated by the best available methodology found in 80% of the subjects of a cohort and in 20% of a perfectly matched control group [think of, e.g. the dexamethason suppression test (DST) or with a corticotrophin-releasing hormone infusion (Dex-CRH)] of melancholic/ major depression (Carroll et al. 1981; Holsboer 2000). The Carroll (and several later) studies have proposed that the DST might well indicate or reflect pathological mechanisms or processes underlying depression. Such reasoning does—albeit indirectly—most often support the concept of depression as defined by the DSM. The issue here is: could the DST challenge the DSM concept of depression? Not likely. A formal argument is that the DSM does not mention cortisol excretion or dexamethason suppression as a feature of major depression. Besides such formal argument, it is more likely that a parameter that is present in a limited number of patients defined with the DSM criteria will be considered as diagnostically useless. Another example: consider a symptom or parameter that is responsible for only certain aspects of depression (say, irritability; Russo et al. 2004, 2005). Again, would this be sufficient to challenge a classification system? A very similar argumentation holds for the response to a therapeutic intervention. For example, an intervention may prove to be therapeutically effective in a cohort, but only in 60% of the subjects. This is, for instance, the case with antidepressants (Kirsch et al. 2008; Turner et al. 2008; Moncrieff and Kirsch 2005). Apparently also, the therapeutic response is not an argument to redefining a diagnostic classification. Another obstacle for a more fundamental discussion on the utility (or better validity, Kendell and Jablensky 2003) of the DSM is the publication bias. It is generally discouraged to publish psychiatric investigations without adherence to the DSM classification (or another classification system). These considerations together argue that studies on (objective) parameters (e.g. genes, hormones, course of the disorder or response to therapy) are unable to challenge the DSM classification system. Is this really bad for psychiatry? Scientifically speaking, there is no argument: adherence to a nontestable classification system hinders progress. It may be even worse (e.g. Kendell and Jablensky 2003 and references therein): the DSM classification may obstruct scientific progress, because it forces to compare heterogeneous cohorts ('apples and pears') for common characteristics ('genes, colour', etc.). But common sense also tells that not every single parameter is sufficiently strong to challenge a psychiatric classification system. And as mentioned, one cannot attack the DSM with criteria that were never included.

To compare with the medical practice, a first medical diagnosis is a provisional hypothesis about an illness, and further investigations are required to support or to reject the initial diagnosis. The question here is: when are results sufficiently strong to challenge a classification? We thought that an experiment suggests that the rational medical practice does not apply to a psychiatric diagnostic systems. Our reasoning implies that a scientific foundation of psychiatry by using the DSM classification is practically unreachable. To avoid the yes/no argumentation, we suggest a way of challenging current classification systems by searching for diagnostic prototypes (Jablensky 2005; Kendell and Jablensky 2003). Our suggestion is related to the notion of 'zones of rarity' to indicate natural boundaries between diagnostic entities.

Causality and the brain

Biological systems are often considered as linear causal systems. This implies a deterministic relationship between an external or internal event and the emerging state of the organism. In that case a rational diagnosis and predictable outcome of a therapy is—at least in theory—possible if sufficient knowledge is or becomes available to the clinician. This idea does not necessarily apply to psychiatry, as will be illustrated with recent work on depression. The major pathogenic hypotheses of depression assumes the involvement of life events (Kendler et al. 1999; Kendler 2008; Keller et al. 2007), biogenic amines (in particular, serotonin; Russo et al. 2007) and stress hormones (cortisol; Holsboer 2000). Objective response to a stressful life event should be distinguished from subjective response, because life events are stressful only when perceived as such. In particular, the subjective response may increase the risk of psychopathology. The subjective perception depends primarily on the memory of previous experiences, whereas the objective response may be modified by genetic and other biochemical dispositions. It is generally believed that the combination of

objective and subjective features of stress coping contribute to the development of depression (e.g. Kendler 2008). Various aspects of monoamines have been investigated including failing synthesis of brain serotonin or genetically programmed variations of proteins (transporters or catabolic enzymes). Close inspection shows that 10% or less of the variance can be explained by the environmental and biochemical factors (Russo et al. 2007). There are also several promising results achieved in animal models. For instance, Spengler and associates (Murgatroyd et al. 2009) highlighted the crucial role of early life stress events for the emergence of epigenetic marking (hypomethylation) of key regulatory region for the expression of arginine vasopressin in the Para ventricular hypothalamic nucleus. This reflects on the function of hypothalamo-pituitary-adrenal axis in the proper adaptation and represents potential model for the vulnerability of the nervous system to environmental and behavioural stimuli. However, such data might point also to the contribution of several more but as yet unknown factors. The question arises whether a more complete knowledge of these factors is possible and, if so, whether their number is sufficient small to construct a useful diagnosis.

By far, most of the depressive episodes are reversible: meaning that in the general population with or without professional therapeutic interventions, about 80% recovers (Spijker et al. 2002) We measured and modeled the incidence of recovery of a depressive cohort extracted from the general population with or without comorbid psychiatric or somatic pathology (van der Werf et al. 2006; Kaptein et al. 2007). It appeared that the rate of recovery follows an exponential time course, indicating that the probability to recover from depression is independent of the length of the depressive episode. In other words, the chance to recover in the first month after the onset of depression is the same than to recover in, for instance, the third month. The exponential function could explain more than 98% of the variance and was applicable irrespective of comorbidities. The model showed that the 2-week criterion of the DSM is rather arbitrary and not a characteristic (i.e. an 'incubation period') for major depression (van der Werf et al. 2006; Kaptein et al. 2007). Results on hospitalised depressed patients (e.g. Keitner et al. 1992) suggest that the time-to-recovery curves can be modelled with exponential functions as well. A prospective study on recurrent depression showed that the durations of preceding or subsequent depressive episodes and depression-free intervals were all unrelated (Kaptein 2008). The exponential function of recovery from depression can best be explained and modelled by assuming random-mood swings: this means that the brain might transit randomly from one ('depressive') to another ('normal') state. Examples of fast brain-state transitions are sleep stages, fear and happiness. Within seconds, we fall asleep or transit from the slow-wave sleep

into the paradoxical (rapid eye movement) sleep and vice versa. Transitions to or from a depressive state are also fast: a positive response after one night of sleep deprivation may disappear during (or after) a nap of a few minutes (Riemann et al. 1993).

The occurrence of fast and random mood transitions questions how deterministic and thus predictable brain transitions are. Korf and Gramsbergen (2007) argued that the brain operates as an *iso*-energetic system, meaning that there are minimal energy barriers to initiate or execute neural activities (Fig. 1). So neural activities in milliseconds do not depend on energy recruitment but are a manifestation of already present energy resources, designated as potential energy. Potential energy is equivalent with trans-membrane ion gradients (i.e. K^+, Na^+Cl^-) and neurotransmitter pools (i.e. GABA, glutamate) and is (nearly) same all over the brain. The potential energy becomes available with minimal effort: as with a battery by changing a switch. Brain-state transitions may appear randomly not only to the external observer but also to the subject himself.

Conventionally biological systems are considered deterministic, i.e. that a brain state can be predicted from a previous state (Goldberger 1996; Williams 1997; Gottschalk et al. 1995; Mandell and Selz 1995). The *iso*-energetic concept allows indeterminacy, so nonpredictable (i.e. random) or near-random (i.e. stochastic) fluctuations are possible. The *iso*-energetic and some other concepts corroborate the idea that the brain is amenable to chaos theory. In a chaotic (in the mathematical sense) system, small perturbations have large and unpredictable consequences. This idea has previously been emphasised in studies on bipolar depression and related affective states (Gottschalk et al. 1995; Mandell and Selz
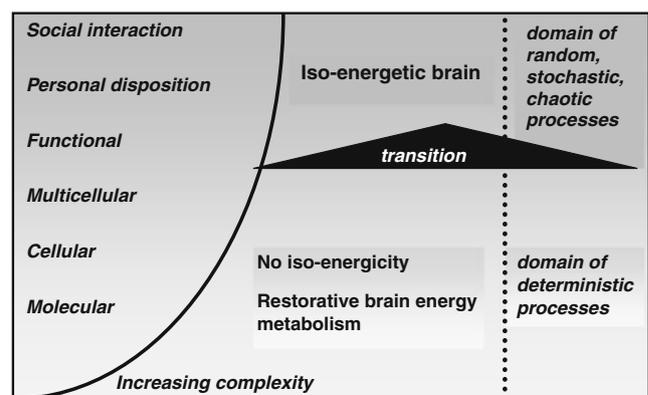

Fig. 1 Schematic representation of the concept of *iso*-energetic brain in relation to complexity of neurobiological processes. To reach *iso*-energicity, energy metabolism (via glucose and oxygen) is required, which is deterministically regulated. *Arrow* indicates direction to reach the *iso*-energetic state. Once the brain near *iso*-energetic nondeterministic processes, such as random, stochastic and chaotic processes, become possible, more complex functions are confined to the *iso*-energetic brain and are thus less deterministic

1995; Pezard et al. 1996), but has not been linked to a brain theory. The brain may exert transitions on external or internal signals that are not well predictable in time and in intensity. The concept of an *iso*-energetic brain implies a rather loose and often untraceable connection between input–output and therefore challenges strict causality underlying classical medical models. An *iso*-energetic brain makes emotional and behavioural reactions to minor external triggers or life events plausible.

Therapeutic interventions

In somatic medicine, at least four types of interventions can be distinguished. The type 1 intervention aims to remove the disease-causing agent either by direct challenging its existence in the patient's body or to assist the disease-eliminating activity of the organism. Antibiotic therapy is the classical example. Compensation is the type 2 intervention: a component with a similar or identical activity is prescribed or measures are taken to enhance the production of a missing compound. Many anti-diabetic treatments or the L-DOPA therapy and dopamine agonists in Parkinson's disease are examples. The type 3 intervention aims to remove a pathological agent with nonspecific toxins, mechanical devices or irradiation. Cancer may serve as an example. And, finally, type 4 treatment is primarily aimed to reduce suffering by, for instance, palliative treatment. Each of these treatments is rational and likely to be evidence-based to a certain extent at least. To what category do psychiatric treatments belong? In depression, the routine treatments include cognitive behavioural therapies and medication (with antidepressant drugs). Other regimens include physical exercise, sleep deprivation, light exposure therapy and—in severe cases—electroconvulsive treatment, alone or combined with the routine treatments. Medication has often been thought as type 1 or type 3 interventions, but there is as yet little evidence that deficiencies of brain monoamines actually *cause* depression. It appeared that low brain serotonin (due to low plasma tryptophan) is associated with failing impulse control and aggression rather than with depression (e.g. Russo et al. 2004, 2005). Cognitive therapies have been developed to correct misconceptions associated with the depression (i.e. to reduce the impact of life events; Beck et al. 1985). Whether their reported positive effects are due to a type 4 rather than a type 1 (or perhaps type 3) effect remains to be proven. Interventions, such as physical exercise, sleep deprivation, light exposure therapy and electroconvulsive treatment, are certainly not type 1, 2 or 3 treatments. The latter are perhaps type 4 treatments (alleviating suffering), but that indexation may feel as artificial. In view of the transition hypothesis (see previous section), the anti-depressive interventions may

facilitate switching from the depressive to the nondepressed state. Sleep deprivation and electroconvulsive treatments are treatments that could show early antidepressant effects. But their effects are unstable; consider the depressiogenic naps (Riemann et al. 1993). Exercise and light therapy may also improve mood or serve as prophylactic. Their effects could be seen as type 1 therapies, as hypo-activity is often a prominent symptom of depression, whereas exposure to light activates or compensates failing mechanisms in seasonal affective disorders, respectively. Alternatively, these treatments may facilitate mood (and so brain) transitions. Therefore, we suggest to adding another type of treatment. Our type 5 treatment is to facilitate brain transitions. This idea might be seen as an extension of type 1 because they assist the organism to battle the ailment, but that is a rather artificial formulation for randomly occurring brain states.

There is an ongoing debate on the therapeutic efficacy of antidepressants (see previous sections). In several meta-analyses, their effects appear to be modest as compared to placebo, explaining about 2% of the variance of depressive symptoms on top of placebo. We have two comments. First, the response of depressed subjects to medication varies widely: both fast and clear-cut responses and no response at all have been reported and second, the placebo response is relatively high and variable. Apparently no meaningful distinction can be made between subjects who need medication and those who recover without medication (or placebo). Is this a matter of lack of knowledge or is this inherent of the current conception of depression? The placebo-drug controversy may serve as another argument supporting the random mood concept, and, if so, the drug treatment is a type 5 intervention.

Most recently, there are experimental data reported (Grayson et al. 2010) which suggest the application of inhibitors of DNA methyl transferase and histone deacetylase (HDAC) as potential therapeutic agents in the causal treatments of psychiatric disorders. This is entailed from the assumption that altered patterns of mRNA and protein expression are downregulated in the pathogenesis of such disorders like unipolar depression or schizophrenia. Therefore, HDAC are supposed to activate the mRNA expression in these conditions.

In short, we argued that—perhaps except cognitive psychotherapy—none of the current antidepressive interventions treats underlying pathopsychological mechanisms directly. So an antidepressive treatment is often not aimed to influence the depressive feelings directly, but to influence the course and severity of a depression. Considering the stochastic mood concept, thus assuming near-random brain transitions, a strict causality between diagnosis and therapeutic response must be considered as unlikely.

Inquiry and markers

Medical disorders, including psychiatric disorders, are diagnosed by the content (e.g. mood), severity, time course and coexistent factors. In the case of depression, for instance, mood characteristics, severity of the depression, duration of the depressive episode and comorbid anxiety, dysthymia or somatic disorders are, among others, important criteria for diagnosis, treatment and prognosis. The diagnosis is mainly based on phenomenology and self-reports ('what the psychiatrist sees and hears'), together referring to the content of a disorder. Such assessments might be biased by the theoretical framework of the therapist and by the feelings of the patient. For instance, the negative attributions of patients about the origin of their depression (feelings of guilt and worthlessness) are well known (Beck et al. 1985). Some physiological parameters may affect subjective mood reports as well. For instance, infusing cortisol gave more negative associations in volunteers and patients, suggesting that hypercortisolaemia might bias the patient's report (Tops et al. 2003, 2004).

The most explored biological markers for depression include hormones and associated receptors (e.g. DST or DEX-CRH), a wide variety of genes, blood or urine levels reflecting metabolic processes, physiological responsiveness and brain imaging (for an extensive review, see Mössner et al. 2007). These markers do not inform on the thought contents but may be related to severity, course and comorbidity. Most often, markers are explored in cross-sectional study designs by comparing a cohort of depressed patients and a matched nondepressed cohort. Occasional examples of longitudinal study designs have appeared. One longitudinal study claims that normalisation of the DST or the Dex-CRH test precedes recovery from depression (Carroll et al. 1981; Holsboer 2000). Despite 30 years of research and strong initial claims, the DST and the Dex-CRH tests have not contributed significantly to a better diagnosis of depression. It is also not clear to what aspect of depression (severity, course, comorbidity) an abnormal cortisol function is associated. Extensive studies have appeared on serotonin-related genes (particularly serotonin transporter regulating genes and their polymorphisms) in relation to depression and life events. Claims that severe and repeated life events provoke depression in the s/s genotype subjects (Caspi et al. 2003) have subsequently not been replicated (Uhe and McGuffin 2008). There is certain progress in the research focused on the exploration of peripheral markers to define the epigenetic risk for depression. Special interest is paid in a pilot study of the promoter methylation of the serotonin transporter gene (Olsson et al. 2010), which, however, needs replication as well. To that end, we and others (Mössner et al. 2007) conclude that none of the current candidate markers supports the diagnosis of depression.

Morphological and functional neuroimaging have the potential to show anatomical brain abnormalities or aberrant physiological responses during a mental task, respectively. Morphological changes are usually considered as irreversible, but in depression, reversible changes of regional brain volumes were reported (Drevets 2003; Kronmüller et al. 2008; Frodl et al. 2008). How fast these volumes change is as yet unknown, but most likely not within days or weeks. Therefore, we consider morphological changes conditional rather than causing depression (Kronmüller et al. 2008), possibly by impairing mood transitions. In this respect, morphological changes are similar to somatic disorders such as cardiovascular heart- and brain-pathology, often associated with depression (de Jonge et al. 2006, 2007). Following exposure to emotional stimuli, abnormal responses of brain amygdala, frontal cortex and some *meso*-limbic areas as detected with functional MRI, oxygen PET and water PET were observed (Drevets 2003). Most often brain images of several subjects were combined to reach statistical significance. It is difficult to decide whether imaging results have to be attributed to abnormal processing of the stimuli or that the depressive state overrules brain processing. Neuroimaging and endocrine studies have been used to define *endo*-phenotypes, i.e. subjects who are more susceptible for stress because of certain genes or gene polymorphisms. Vulnerable *endo*-phenotypes have indeed been reported, suggesting that the subjective perception of stress, and by inference of life events, may render individuals more prone to develop depression (e.g. Jabbi et al. 2007a, b). But again: differences between groups, not between individuals were significant.

Psychiatric disorders are characterised by the content (e.g. mood), severity, time course, and coexistent factors. The thus far tested external markers are insufficient to support or reject DSM diagnosis. Few if any attempts have been made to differentiate between associations with the pathopsychological content or with the course, severity or other characteristics of depression. Another unanswered issue is whether group differences could ever become useful for the individual diagnosis

Discussion

We have emphasised four major challenges for a rational and causal conceptualisation of functional disorders in psychiatry (Table 1). Each of these challenges could well lead to the conclusion that a rational and evidence-based psychiatry is theoretically and practically impossible. The present conclusions are mainly based on data of major depression. Besides making a critical analysis, it is our ambition to suggest answers to the raised challenges. We

Table 1 Some obstacles to reach (strict) causality in psychiatry

1. Classification systems are not falsifiable, because of
 (a) heterogeneity of categories
 (b) inclusion of novel factors and
 (c) publication bias
2. Brain functions are at least, in part, not-deterministic, because of
 (a) iso-energeticity and consequently
 (b) chaotic/random and stochastic processes
3. Therapies are not directed to core symptoms/characteristics, because they
 (a) do not influence subjective feelings
 (b) influence course and not pathogenesis and
 (c) are relatively ineffective
4. Markers and inquiry
 (a) do not inform on the core symptoms
 (b) but possibly on course and severity, and
 (c) inquiry is biased because of subjectivity

start the discussion on strong and weak points of our analysis.

First, we consider some philosophical aspects. More than in somatic medicine the question is whether a psychiatric disorder is a real thing or a construct that does not necessarily refer to a real world. Such criticism is not justified, because any scientific theory, diagnosis or hypothesis is not more than a construct that describes phenomena as long as there is no better idea. This philosophical issue touches the problem of irreducibility of real-world processes and objects. An alternative view is that scientific theories and clinical protocols have no or little common basis. Real progress is possible only when both the scientific and clinical views converge to unambiguous conceptualisations. As long as that is not the case in psychiatry, one has to live with certain and unavoidable inconsistencies between a medical and scientific framework of psychiatry. We considered classification systems primarily as useful constructs and not so much as describing real world objects. As an alternative to evidence, we considered the term 'proof' and 'proven' to emphasise causality. More exactly, these terms indicate causality of two types: (i) strict linear causality in relatively simple systems useful for lower levels of complexity; (ii) nonlinear, dynamic causality, applicable to single cells up to the *endo*-phenotype. There remains the possibility to establish 'bridge' laws (Nagel et al. 1971; Nagel 1979) between multitude complexities at micro (neuronal) and macro (behavioural) levels (also indicated in Fig. 1). In this case, we do not exemplify empirical reduction but we may introduce 'patchy' reduction (Kendler 2005) or inter-theoretic reduction between paradigms. The concept of the *iso*-

energetic brain 24 may be seen as a provisional bridge law (Nagel 1979) and as 'a patchy reductive explanation' (Kendler 2005).

In the following, we suggest a strategy to explore some of the discussed issues. The core of our proposal is to re-analyse epidemiological, psychological and laboratory data but not only according to currently accepted diagnostic systems. For instance, genes or biochemical parameters may well be related to severity, course or other features of a disorder (Kendell and Jablensky 2003). If so, then the same gene variants may be associated not only with depression but also with, e.g. anxiety or schizophrenia or even with nonpsychopathological symptoms as well. Another idea is to relate biochemical parameters to a single or a set of symptoms. This could give information on the significance of some pathophysiological parameters to evoke particular symptoms or group of symptoms present in different disorders and, ideally, how they can be affected by therapy.

Many of current ideas on psychopathology are developed from relatively large cross-sectional or longitudinal epidemiological investigations. Indeed, such studies may disclose general trends, but the results are not necessarily applicable to a single patient. A firm causal basis for psychiatry (and any medical discipline) requires a reasonable guess about the diagnosis in the single individual. In practice, several relatively nondisorder-specific factors converge to a diagnosis that has implications for further therapeutic regimens. Furthermore, we emphasise that a diagnosis has to define a typical condition (disease) of a prototypical individual (patient). One aim of a medical psychiatry could be to define and redefine diagnostic prototypes (Kendell and Jablensky 2003; Jablensky 2005; Nagel 1979; Schaffner 2004; Kendler 2008; Kendler and Campbell 2008). Following this line; we suggest that data analysis in epidemiological and other population-based studies might focus to delineate relatively small subcohorts that are defined by a limited number of parameters. These parameters may include not only psychopathological symptoms but also physiological, biochemical or genetic characteristics. And they may be related to severity, time course and other characteristics of the disorder. Diagnostic prototypes might be identified with already available data or with dedicated prospective studies. A similar approach has been elaborated on the distinction of validity and utility of psychiatric diagnosis (Kendell and Jablensky 2003; Jablensky 2005; Schaffner 2004). One criterion is that in a multidimensional space of symptoms a syndrome is defined by detectable discontinuities. So there should be boundaries of rarity, distinguishing syndrome (disease) from sanity or from other syndromes (diseases or disorders; Jablensky 2005; Schaffner 2004; Kendler 2008). If this cannot be achieved, then there is no justification of diagnosing syndromes. The outcome of such exercise could

be the following: first, one or several of the diagnostic prototypes are identical or very close to categories of existing diagnostic systems (e.g. the DSM), second, these prototypes are completely different and not easy to reconcile with current diagnostic systems, or third such hypothesised diagnostic prototypes do not exist or cannot be recognised as yet. The first outcome provides a strong support for a scientific basis and clinical relevance of—or at least a part of—the tested diagnostic system. The second outcome forces to reconsider the diagnostic system under investigation and might eventually lead to new diagnostic systems. The third outcome might lead to a temporary anarchy of diagnostic systems as no cues for developing a new system are revealed. Future scientific developments may enable more pertinent conclusions in line with the first or second outcome. In the latter case, there is as yet no urgent need to reconsider the diagnostic system under study.

Concluding remarks

The present assay is not aimed to provide a conceptual or philosophical framework of psychiatry, as done recently. These papers (Kendler 2005, 2008; Kendler and Campbell 2008) emphasise the multilevel character of explanatory models in psychiatry and warn against too simple mono-level explanations, e.g. at the molecular or psychological level. We agree with this viewpoint as is also illustrated in Fig. 1 and acknowledge the implications for research paradigms. Our intention is to bring the discussion to a more operational level: what should be done to rationalise psychiatry. The medical model was chosen as starting point in our analysis, because it is in our opinion the best tool to discern shortcomings and weak points. Our concern about utility of validity of current diagnostic constructs is shared by others (e.g. Angst 2007). Our emphasis is on utility, not on validity of diagnostic constructs (Kraemer and Kupfer 2006). From a scientific point of view, the distinction between validity and utility is rather artificial, because validity of a concept can only be proven by practical testing, in the medicine preferentially by showing utility. The medical model is the basis of evidence-based medicine: to establish causality between diagnosis, interventions and therapeutic effects. The stricter the causality is, the more evidence-based psychiatry will become. We have discussed four major obstacles: first the nonfalsifiable nature of current classification systems, second the nondeterministic character of many cerebral processes, third divergence between therapeutic approaches and pathogenesis and fourth inadequacy of diagnostic tools. We have emphasised the use of objective parameters to develop alternative classification frameworks. But we also see shortcomings

of this approach. One has to appreciate that knowledge about the world is far more than what can be communicated between human beings. One unavoidable shortcoming might be due to the impossibility to describe professional knowledge (or perhaps better professional intuition) of the psychiatrist verbally that said without the previously mentioned theoretical biases. So we might eventually have to consider the possibility that professional experience is an essential element not only in everyday health care but also in evidence-based psychiatry.

References

Angst J (2007) Psychiatric diagnosis: the weak component of modern research. World Psychiatry 6:94–95

Beck AT, Rush AJ, Shaw BF, Emery G (1985) Cognitive therapy of depression. Guilford, New York

Carroll BJ, Feinberg M, Greden JF, Tarika J, Albala AA, Haskett RF, James NM, Kronfol Z, Lohr N, Steiner M, de Vigne JP, Young E (1981) A specific laboratory test for the diagnosis of melancholia. Standardization, validation, and clinical utility. Arch Gen Psychiatry 38:15–22

Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. Science 301:386–389

de Jonge P, Roy JF, Saz P, Marcos G, Lobo A, ZARADEMP investigators (2006) Prevalent and incident depression in community-dwelling elderly persons with diabetes mellitus: results from the ZARADEMP project. Diabetologia 49:2627–2633

de Jonge P, Denollet J, van Melle JP, Kuyper A, Honig A, Schene AH, Ormel J (2007) Associations of type-D personality and depression with somatic health in myocardial infarction patients. J Psychosom Res 63:477–482

Drevets WC (2003) Neuroimaging abnormalities in the amygdala in mood disorders. Ann N Y Acad Sci 985:420–444

Frodl TS, Koutsouleris N, Bottlender R, Born C, Jäger M, Scupin I, Reiser M, Möller HJ, Meisenzahl EM (2008) Depression-related variation in brain morphology over 3 years: effects of stress? Arch Gen Psychiatry 65:1156–1165

Grayson DR, Kundakovic M, Sharma RP (2010) Is there a future for histone deacetylase inhibitors in the pharmacotherapy of psychiatric disorders? Mol Pharmacol 77(2):126–135

Goldberger A (1996) Non-linear dynamics for clinicians: chaos theory, fractals and complicity at the bedside. Lancet 347:1312–1314

Gottschalk A, Bauer MS, Whybrow PC (1995) Evidence of chaotic mood variation in bipolar disorder. Arch Gen Psychiatry 52:947–959

Holsboer F (2000) The corticosteroid receptor hypothesis of depression. Neuropsychopharmacology 23:477–501

Jabbi M, Korf J, Kema IP, Hartman C, van der Pompe G, Minderaa RB, Ormel J, den Boer JA (2007a) Convergent genetic modulation of the endocrine stress response involves polymorphic

variations of 5-HTT, COMT and MAOA. Mol Psychiatry 12:483–490

Jabbi M, Kema IP, van der Pompe G, te Meerman GJ, Ormel J, den Boer JA (2007b) Catechol-*o*-methyltransferase polymorphism and susceptibility to major depressive disorder modulates psychological stress response. Psychiatr Genet 17:183–193

Jablensky A (2005) Categories, dimensions and prototypes: critical issues for psychiatric classification. Psychopathology 38:201–205

Kaptein KI (2008) Recovery from depression: new insights in time-course analyses. Thesis. University of Groningen. pp. 114–128. ISBN 978-90-367-3549-0

Kaptein KI, De Jonge P, Korf J, Spijker J, De Graaf R, Van Der Werf SY (2007) Random-mood interpretation of determinants for major depression. Psychol Med 37:1261–1271

Keitner GI, Ryan CE, Miller IW, Norman WH (1992) Recovery and major depression: factors associated with twelve-month outcome. Am J Psychiatry 149:93–99

Kendell RE (2001) The distinction between mental and physical illness. Br J Psychiatry 178:490–493

Kendell R, Jablensky A (2003) Distinguishing between the validity and utility of psychiatric diagnoses. Am J Psychiatry 160:4–12

Kendler KS (2005) Toward a philosophical structure for psychiatry. Am J Psychiatry 162:433–440

Kendler KS (2008) Explanatory models for psychiatric illness. Am J Psychiatry 165:695–702

Kendler KS, Campbell J (2008) Editorial review. Interventionist causal models in psychiatry: repositioning the mind–body problem. Psychol Med 10:1–7

Kendler KS, Karkowski LM, Prescott CA (1999) Causal-relationship between stressful life events and the onset of major depression. Am J Psychiatry 156:837–841

Keller MC, Neale MC, Kendler KS (2007) Association of different adverse life events with distinct patterns of depressive symptoms. Am J Psychiatry 164:1521–1529

Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT (2008) Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Med 5:e45

Korf J (2008) On causality in psychiatry. Psychiatry and freedom: conference for philosophy and mental health, Dallas, TX, USA. Abstract, p 24. http://www.utsouthwestern.edu/psychiatryandfreedom

Korf J, Gramsbergen JB (2007) Timing of potential and metabolic brain energy. J Neurochem 103:1697–1708

Kraemer HC, Kupfer DJ (2006) Size of treatment effects and their importance to clinical research and practice. Biol Psychiatry 59:990–996

Kronmüller KT, Pantel J, Köhler S, Victor D, Giesel F, Magnotta VA, Mundt C, Essig M, Schröder J (2008) Hippocampal volume and 2-year outcome in depression. Br J Psychiatry 192:472–473

Mandell AJ, Selz KA (1995) Nonlinear dynamical patterns as personality theory for neurobiology and psychiatry. Psychiatry 58:371–390

Moncrieff J, Kirsch I (2005) Efficacy of antidepressants in adults. Br Med J 16:155–157

Mössner R, Mikova O, Koutsilieri E, Saoud M, Ehlis AC, Müller N, Fallgatter AJ, Riederer P (2007) Consensus paper of the WFSBP task force on biological markers: biological markers in depression. World J Biol Psychiatry 8:141–174

Murgatroyd C, Patchev AV, Wu Y, Micale V, Bockmühl Y, Fischer D, Holsboer Fn, Wotjak CT, Almeida OFX, Spengler D (2009) Dynamic DNA methylation programs persistent adverse effects of early-life stress. Nature Neurosci. Published online 8 November 2009

Nagel E (1979) Teleology revisited and other essays in the philosophy and history of science. The John Dewey essays in philosophy 3. Columbia University Press, New York

Nagel E, Bromberger S, Grünbaum A (1971) Observation and theory in science. The Alvin and Fanny Blaustein Thalheimer lectures 1969. Johns Hopkins University Press, Baltimore, MD

Olsson CA, Foley DL, Parkinson-Bates M, Byrnes G, McKenzie M, Patton GC, Morley R, Anney RJL, Craig JM, Saffery R (2010) Prospects for epigenetic research within cohort studies of psychological disorder: a pilot investigation of a peripheral cell marker of epigenetic risk for depression. Biol Psychol 83(2):59–65

Pezard L, Nandrino JL, Renault B, el Massioui F, Allilaire JF, Müller J, Varela F, Martinerie J (1996) Depression as a dynamical disease. Biol Psychiatry 39:991–999

Riemann D, Wiegand M, Lauer CJ, Berger M (1993) Naps after total sleep deprivation in depressed patients: are they depressiogenic? Psychiatry Res 49:109–120

Russo S, Boon JC, Kema IP, Willemse PH, den Boer JA, Korf J, de Vries EG (2004) Patients with carcinoid syndrome exhibit symptoms of aggressive impulse dysregulation. Psychosom Med 66:422–425

Russo S, Kema IP, Haagsma EB, Boon JC, Willemse PH, den Boer JA, de Vries EG, Korf J (2005) Irritability rather than depression during interferon treatment is linked to increased tryptophan catabolism. Psychosom Med 67:773–777

Russo S, Kema IP, Bosker F, Haavik J, Korf J (2007) Tryptophan as an evolutionarily conserved signal to brain serotonin: Molecular evidence and psychiatric implications. World J Biol Psychiatry 13:1–11

Schaffner KF (2004) Commentary on Stotz and Griffiths, Burian, and Waters: genes, concepts, DST implications, and the possibility of prototypes. Hist Philos Life Sci 26:81–90

Shah P, Mountain B (2007) The medical model is dead—long live the medical model. Br J Psychiatry 191:375–377

Shorter E (1997) A history of psychiatry. Wiley, New York, pp 1–436

Spijker J, de Graaf R, Bijl RV, Beekman AT, Ormel J, Nolen WA (2002) Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Br J Psychiatry 181:208–213

Stoyanov D, Popov G, Korf J (2008) Synergistic psychiatry and molecular neuroscience in the re-evaluation of the categories in psychiatry. Psychiatry and freedom: conference for philosophy and mental health. Dallas, TX, USA. Abstract, pp 35–36. http://www.utsouthwestern.edu/psychiatryandfreedom

Tops M, van der Pompe G, Baas D, Mulder LJ, Den Boer JA, Meijman TF, Korf J (2003) Acute cortisol effects on immediate free recall and recognition of nouns depend on stimulus valence. Psychophysiology 40:167–173

Tops M, van der Pompe G, Wijers AA, Den Boer JA, Meijman TF, Korf J (2004) Free recall of pleasant words from recency positions is especially sensitive to acute administration of cortisol. Psychoneuroendocrinology 29:327–338

Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R (2008) Selective publication of antidepressant trials and its influence on apparent efficacy. New Engl J Med 358:252–260

Uhe R, McGuffin P (2008) The moderation by the serotonin transporter gene of environmental adversity in the aetiology of mental illness: review and methodological analysis. Mol Psychiatry 13:131–146

van der Werf SY, Kaptein KI, de Jonge P, Spijker J, de Graaf R, Korf J (2006) Major depressive episodes and random mood. Arch Gen Psychiatry 63:509–518

Williams GP (1997) Chaos theory tamed. Joseph Henry, Washington, DC