

Reliability of clinical ICD-10 diagnoses among electroconvulsive therapy patients with chronic affective disorders

Klaus Damgaard Jakobsen^{*,,***}**
Thomas Hansen^{*,**}**
Henrik Dam^{**}**
Ejnar Bundgaard Larsen^{***}**
Ulrik Gether^{*}**
Thomas Werge^{*}

^{*} Research Institute of Biological Psychiatry,
Mental Health Centre St. Hans, Roskilde

^{**} University Department of Psychiatry,
Mental Health Centre Hvidovre, Broendby

^{***} Center for Pharmacogenomics, University
of Copenhagen, Copenhagen N

^{****} University Department of Psychiatry,
Mental Health Centre Rishospitalet,
Copenhagen O

^{*****} University Department of Psychiatry,
Mental Health Centre Glostrup, Glostrup

DENMARK

ABSTRACT – Background and Objectives: Diagnostic reliability is of major concern both to clinicians and researchers. The aim has been to investigate the trustworthiness of clinical ICD-10 affective disorder diagnoses for research purpose.

Methods: 150 ECT patients with chronic affective disorders were investigated. A standardized schema for basic anamnesis and the Operational Criteria Checklist for Psychotic and Affective Illness (OPCRIT) were used. The sensitivity, specificity, positive and negative predictive values of clinical affective disorder ICD-10 diagnoses and the formal agreement between clinical ICD-10, OPCRIT ICD-10 and DSM-IV diagnoses were determined using unweighted κ -statistics.

Results: The sensitivity, specificity, positive and negative predictive values of the clinical bipolar diagnoses was 0.55, 0.75, 0.42 and 0.84, respectively. The sensitivity, specificity, positive and negative predictive values of the clinical unipolar diagnoses was 0.79, 0.55, 0.77 and 0.58, respectively. The agreement between clinical ICD-10 and OPCRIT ICD-10 bipolar vs. non-bipolar diagnoses was low, $\kappa = 0.28$. The agreement between clinical ICD-10 and OPCRIT ICD-10 unipolar vs. non-unipolar diagnoses was low, $\kappa = 0.35$. The agreement between OPCRIT ICD-10 and DSM-IV diagnoses on bipolar vs. non-

bipolar disorders was high, $\kappa = 0.91$, and the agreement on unipolar vs. non-unipolar disorders was fairly high, $\kappa = 0.78$.

Conclusions: This study demonstrates that the reliability of clinical ICD-10 diagnoses of affective disorders from chronic subjects with a history of ECT is problematic despite sample homogeneity on basic clinical, demographic and epidemiological parameters.

Received 7 February 2008

Revised 22 June 2008

Accepted 11 July 2008

Introduction

A fundamental objective for research into affective disorders is to identify core features linked with the illnesses. Like schizophrenia spectrum disorders, affective disorders are clinical syndromes rather than distinct disease entities. The present diagnostic criteria, both ICD-10¹ and DSM-IV², are categorical systems that assign diagnoses according to a set of hierarchical operational criteria that have their origin in psychopathologic tradition rather than biological aetiology.

This strategy has ensured high reliability of the present diagnostic classifications in terms of satisfactory content validity. But it remains unclear how well these conventionally defined syndromes are related to the underlying disorders, they are supposed to capture. The construct validity of the contemporary operational diagnostic classifications remains indeterminate³.

To overcome part of the cleft between phenomenological psychopathology and biology, research into affective disorders is increasingly focussing on markers that can be associated to the particular clinical syndromes under study. Thus, research programs focus on presumed endophenotypes⁴ associated with the illnesses in search for indicators

that can confirm diagnoses, secure more accurate treatment and predict prognoses.

The studies on putative endophenotypes of affective disorders cover a broad spectrum of research disciplines e.g. neuropsychology, functional imaging, neuroendocrinological testing and genetics. The common pathway of all these investigations on particular subsets of patient populations always seems to be recruitment of subjects to a specific study on the basis of the clinical syndrome or diagnoses. It is therefore essential that the clinical diagnoses are reliable, for optimal medical care or best possible research. The reliability of clinical ICD-10 affective disorders diagnoses therefore remains a central issue both clinically and in research programs.

The affective disorder spectrum has been proposed by Akiskal^{5,6}. The construct validity of the disorders within the affective spectrum⁷ and their relation to other spectra, e.g. the schizophrenia spectrum, remains debatable as the present operational criteria⁸ are conventionally defined syndromes made to enhance reliability of diagnoses.

Both incidence and prevalence of affective disorders is growing^{9,10} for unknown reasons, enhancing the need for advanced diagnostic methods¹¹ to ensure correct affective disorder diagnoses.

The aims of the study have been to analyse the reliability of clinical ICD-10 affective disorder diagnoses among chronic severely ill patients with a lifetime history of electroconvulsive therapy (ECT). In order to see if chronicity in terms of duration of illness enhance criterion validity of clinical ICD-10 affective disorder diagnoses. The operational criteria OPCRIT checklist for psychotic and affective illness^{12,13} was used as a gold standard to calculate predictive values, sensitivity and specificity of the clinical diagnoses. The current principal and clinical ICD-10 diagnoses were compared with the ICD-10 and DSM-IV diagnoses of the OPCRIT instrument.

Methods

Ethics

The study has been carried out in accordance to the Helsinki Declaration. The Danish Data Protection Agency and the Danish Scientific-Ethical Committees (file # 01-024/01) have approved the study. All patients had given written informed consent prior to inclusion into the project: Dansk Psykiatrisk Biobank. At the time of recruitment and rating no subject was subdue to civil or forensic psychiatric restraint.

Sample

Danish Psychiatric Biobank (DPB) was founded in 2001 as a joint venture between the six university departments of Psychiatry of the previous Copenhagen Hospital Corporation, renamed, Region Hovedstadens Psykiatri, Denmark.

In- and outpatients from the catchments areas of the greater metropolitan area of

Copenhagen were asked on a routinely basis to participate in the Biobank, i.e. to allow access to their medical records, to participate in psychopathological interviews or neuropsychological tests and to donate a blood sample.

A total of 155 patients with clinical ICD-10 affective disorders, i.e. bipolar disorders (BPD) and unipolar (first and recurrent) major depressive disorders (MDD), whom all had a lifetime history of ECT were randomly sampled at four major Mental Health Centres in the Region during a six-month period, autumn 2006. ECT treatment was used as a proxy for affective disorder illness, as ECT is a very unusual treatment option¹⁴ for schizophrenia spectrum disorders, in Denmark. The most recent and principal ICD-10 clinical diagnoses were recorded.

Sample homogeneity in terms of basic clinical, demographic and epidemiological data is shown in Table I.

Diagnostic interviews and rating scales

All 155 subjects were assessed using a standardized schema, Danish Psychiatric Biobank Schema (DPB-schema), to extract basic clinical, demographic and epidemiological information from interviews with patients and their medical records. The DPB-schema contains three scales that requires rating by the investigator: Clinical Global Impression Scale for Severity of Illness (CGI)¹⁵, Global Assessment of Functioning scale (GAF)¹⁶ and a Suicide Attempt Scale (SAS) constructed according to WHO standards¹⁷. The inter-rater reliability has been investigated with respect to CGI and SAS scales based on data from DPB-subjects that have been assessed independently (and at different times). The inter-rater

Table I
Basic Demographic and epidemiological data*

	All	Men	Women	Unipolar	Bipolar
N	155	47	108	107	48
Age (years)	53,4 +/- 14,5	52,8 +/- 12,6	53,7 +/- 15,3	53 (40-62,3)	56,5 (48-62)
Age at onset (years)	27,5 (18-46)	32,5 (20-46)	25,5 (17-45,5)	30 (18-49)	24 (17,8-37,5)
Age at first treatment (years)	32 (24-47)	35 (25-47)	32 (23-47)	35 (25-49)	30,5 (23-44)
Age at first hospitalization (years)	38 (28-51,5)	43 (32-54)	36 (26-50)	40 (30-54) ^A	32 (23-46) ^A
Duration of illness (years)	18 (9-32)	16 (7,3-29,8)	19 (9-36)	16 (7-30,5) ^B	25,5 (14-38) ^B
Number of depressions, (#)	4 (3-9)	3,5 (2-5) ^C	5 (3-10) ^C	4 (3-8)	5 (2,3-10)
Number of treated depressions, (#)	3 (2-6)	3 (1,5-4) ^D	4 (2-8) ^D	3 (2-6)	4 (2-7,5)
Number of manic or mixed episodes, (#)	0 (0-2)	0 (0-2)	0 (0-1)	0 (0-0) ^E	5 (2-8) ^E
Number of treated manic or mixed episodes, (#)	0 (0-0)	0 (0-0,3)	0 (0-0)	0 (0-0) ^F	3 (1-5) ^F
Suicide Attempt Scale score**	1 (1-3)	1 (1-3)	1 (1-3)	1 (1-3)	1 (1-2)
Global assessment of functioning score (GAF)	50 (40-51)	50 (45-60) ^G	45 (39,5-50) ^G	50 (40-57,5)	45 (40-50)
Hamilton score	22 (14,5-28)	20,2 +/- 10,1	21,6 +/- 8,6	21 +/- 9,1	21,8 +/- 8,9

* Data are given as means (+/- SD), when normally distributed, otherwise as medians (25-75% quartiles).

** Lifetime suicidality according to WHO standards, 0 = none up till 5 = more than one determinant or violent suicide attempt.

A. Age at first hospitalisation is significantly different between unipolar and bipolar subjects, $P = 0.005$.

B. Duration of illness is significantly different between unipolar and bipolar subjects, $P = 0.010$.

C. # Numbers of depressions is significantly different between men and women, $P = 0.030$.

D. # Numbers of treated depressions are significantly different between men and women, $P = 0.036$.

E. # Numbers of manic or mixed episodes are significantly different between unipolar and bipolar subjects, $P = 0.001$.

F. # Numbers of treated manic or mixed episodes are significantly different between unipolar and bipolar Subjects, $P = 0.001$.

G. GAF score differed significantly between men and women, $P = 0.010$.

agreement on 305 DPB-subjects of the CGI and of the SAS on 308 DPB-patients is very good ($\kappa = 0.87$ & 0.81 , respectively).

The inter-rater agreement of GAF scores has not been investigated in the framework of DPB as the author or any other DPB-rater, only rarely have performed GAF ratings. The GAF estimate was collected from the medical records of the DPB-participants

and thus derives from ratings performed by the clinical psychiatrists. Due to multiple sites of investigation with multiple clinical psychiatrists, data on the inter-rater reliability on the GAF scores have not been possible due to logistic reasons.

Patients were diagnostically evaluated using the 90-item, operational criteria OPCRIT checklist for psychotic and affective

illness^{18,19}, version OPCRIT 4 Windows. The OPCRIT checklist was preferred to other diagnostic instruments due to its simple, operational, computerised, and polydiagnostic approach.

All OPCRIT checklist assessments and DPB-schema ratings were done on a life-time basis/approach using interviews with patients and all available medical records.

Diagnoses generated from patients files are known to be comparable to, live best estimate procedures, if medical records are comprehensive²⁰, which is the case in this study due to long the duration of illness of the patients. The OPCRIT instrument^{21,22} is acknowledged for the ability to generate valid, life-time best estimate diagnoses, when using interviews with patients and their medical records. All diagnostic assessment of subjects, their medical records and rating procedures were done by the first author (KDJ), being a SCAN-certified research consultant psychiatrist, as previously described^{23,24}. An OPCRIT co-rating group running for 2 years, lead by KDJ, was established. Resident, research and several SCAN-certified consultant psychiatrists from the university centres of psychiatry of Regions Hovedstadens Psykiatri participated in the OPCRIT co-rating group with the purpose to avoid rater drift.

The OPCRIT 4 windows version generates 15 different ICD-10 diagnoses and 15 different DSM-IV diagnoses. The DSM-IV diagnoses were translated into ICD-10 diagnoses using the DSM-IV, International Version with ICD-10 Codes². Due to small sample size and the fact, that the DSM-IV diagnoses of bipolar II is not part of the ICD-10 criteria, which is the currently used system in Denmark, DSM-IV bipolar I & II diagnoses were merged as BPD. All unipolar diagnoses were merged

as equivalent to the DSM-IV diagnoses of MDD on the basis of the Hamilton Depression Scale (Ham-D)²⁵ scores, see section on results.

The Ham-D scores were collected from previous Ham-D ratings documented in the medical records, and for subjects without previous Ham-D assessment, this was done at the live interviews with the patients. The highest Ham-D score ever, was chosen as representative for worst depressive episode, for each patient.

The Global Assessment of Functioning (GAF)²⁶ and the Clinical Global Impression Scale for Severity of Illness (CGI)¹⁵ were chosen to estimate the overall level of functioning / severity of illness of the subjects based on their medical records, personal interviews or both. The CGI scores were dichotomised into two groups: 0-4 and 5-7 (data not shown).

A Suicide Attempt Scale was constructed according to WHO standards¹⁷ in order to investigate differences in suicidal behaviour in the sample. Data was collected from subject's medical records, the following scores were used: 0 = none, 1 = suicidal thoughts recorded, 2 = One mild / non-determinant attempt recorded, 3 = More than one mild / non-determinant attempt recorded, 4 = One violent / determinant attempt recorded, 5 = More than one violent / determinant attempt recorded.

Age at onset, age at first treatment and age at first hospitalisation were defined, as the earliest age, at which subjects / medical records could remember / describe their first signs of a depressive or manic episode, their first treatment and their first hospitalisation for an affective disorder episode. The degree of chronicity was defined as duration of illness and estimated as time from age at onset till 2007.

Statistical analyses

All data have been analysed with SigmaStat® version 3.1. The Kolmogorov-Smirnov normality test was used to differentiate parametric from nonparametric anamnestic, demographic and epidemiological data. Results on these parameters were characterized by means (+/- SD) (for normally distributed data) or medians (25/75 quartiles) (for non-normally distributed data). Student's t-test was used to compare normally distributed data, while the Mann-Whitney rank sum test was used to compare non-parametric data. Chi-square test was used to compare data with dichotomised or categorical distribution. Fisher Exact test was used on dichotomised data of small amount. Sensitivity, specificity, positive and negative predictive values were calculated to investigate the reliability of ICD-10 diagnoses from the clinical setting versus OPCRIT derived diagnoses, using the OPCRIT diagnoses as reference, see footnote below*. Unweighted κ -statistics were used to estimate formal agreement between clinical ICD-10, OPCRIT ICD-10 and DSM-IV derived diagnoses according to Altman²⁷, see footnote below**.

Results

Basic clinical, demographic and epidemiological data

The clinical, demographic and epidemiological data of the 155 subjects obtained

with the DPB-schema are shown in Table I. Approximately two-third of the sample is women reflecting the fact that MDD is predominantly seen in women. About one-third of the subjects have BPD and two-third has MDD revealing an overrepresentation of BPD patients among chronic subjects that are treated with ECT in this sample.

As shown in Table I, no significant differences were found between men and women or between BPD and MDD subjects on age, age at onset and age at first treatment, indicating a homogeneous distribution in the sample. Significant differences between BPD and MDD were found on age at first hospitalisation ($p = 0.005$) and duration of illness ($p = 0.010$), but not between men and women. Thereby, showing an earlier need of hospital care, and longer duration of illness among BPD patients.

There were significant differences between men and women in numbers of depressive episodes ($p = 0.030$) and numbers of treated depressive episodes ($p = 0.036$), but not between BPD and MDD subjects, demonstrating that women have more or request more treatment for depressions than men in this sample.

The severities of the depressive episodes were homogeneously distributed in the sample as no differences were found between BPD and MDD patients or genders on Ham-D or Suicide Attempt Scale scores.

The lifetime ratings on functional level and severity of illness revealed that women had significantly lower GAF and CGI scores ($p = 0.010$ and $p = 0.040$). Women are more dys-

* Sensitivity = number of true positives / number of true positives + number of false negatives; specificity = number of true negatives / number of true negatives + number of false positives; positive predictive value = numbers of true positives / numbers of true positives + numbers of false positives; negative predictive value = numbers of true negatives / numbers of true negatives + numbers of numbers of false negatives.

** Strength of agreement: $K < 0.20$ = poor; $0.21 < K < 0.40$ = fair; $0.41 < K < 0.60$ = moderate; $0.61 < K < 0.80$ = good; $0.81 < K < 1.00$ = very good.

functional compared to men, who on the contrary are more ill. The GAF and CGI scores did not differ between diagnostic groups.

The study did not find any significant gender or disorder related differences in regard to: family history of MDD, BPD or other psychiatric illnesses, treatment responses on medication or ECT, psychosis, tobacco, alcohol or drug addiction. Nine of 155 subjects were treated with thyroid substitution, 115 subjects had normal thyroid stimulation hormone (TSH) values, seven subjects had elevated TSH levels and one subject had too low.

Sensitivity, specificity, and prediction values of ICD-10 affective disorders

The overall predictive value of the clinical ICD-10 diagnoses (right vs. wrong) was moderate (66%). The total correct predictions of BPD and MDD were moderate ($\approx 70\%$), see Table II.

However, sensitivity and positive predictive values of BPD were unexpectedly low (55% & 42%). Contrary to this, both the

specificity and the negative predictive values of BPD were higher (75% & 84%). The sensitivity and the positive predictive values of MDD were comparable higher (79% & 77%) and the specificity and the negative predictive values (55% & 58%) contrarily lower than those of BPD.

Seven (25%) and five subjects (23%) of those with clinical ICD-10 BPD and MDD, who shifted diagnoses, were converted to ICD-10 diagnoses outside the affective disorder spectrum by OPCRIT. The seven BPD shifted to two subjects with schizophrenia and five subjects with unspecified non-organic psychoses. The five MDD converted to three subjects with schizophrenia and two subjects with unspecified non-organic psychoses. Further, 21 subjects with clinical ICD-10 BPD were converted by OPCRIT to ICD-10 MDD cases and 17 subjects with clinical ICD-10 MDD shifted to OPCRIT ICD-10 BPD.

The formal agreement between the clinical ICD-10, OPCRIT ICD-10 and DSM-IV diagnoses has been calculated using unweighted κ -statistics. The concordance between clinical ICD-10 and OPCRIT ICD-10 diagnoses on BPD vs. non-BPD and MDD vs. MDD

Table II
Sensitivity, specificity and prediction values of ICD-10 unipolar and bipolar disorders

	All*	Unspecified non-organic psychoses	Schizophrenia	Unipolar**	Bipolar***
Clinical ICD-10	155			105	50
OPCRIT ICD-10	152	7	5	102	38
OPCRIT DSM-IV	151	16	4	96	35
Sensitivity				0,79	0,55
Specificity				0,55	0,75
Positive predictive value	0,66			0,77	0,42
Negative predictive value				0,58	0,84
Total correct predictions	0,66			0,71	0,70

* Identical OPCRIT ICD-10 and DSM-IV diagnoses vs. not identical diagnoses.

** Identical OPCRIT ICD-10 and DSM-IV unipolar diagnoses F 32-33) vs. not identical diagnoses.

*** Identical OPCRIT ICD-10 and DSM-IV bipolar diagnoses F 30-31) vs. not identical diagnoses.

disorders were unexpectedly low ($\kappa = 0.28$ and $\kappa = 0.35$) despite the long duration of illness of in the sample. The agreement between OPCRIT ICD-10 and DSM-IV diagnoses were high ($\kappa = 0.91$ and $\kappa = 0.78$) between the specific disorders, i.e. BPD vs. non-BPD and MDD vs. non-MDD, but only moderate ($\kappa = 0.62$) in the total sample, i.e. affective vs. non-affective disorders.

Discussion

The present study on the diagnostic reliability of clinical affective disorder diagnoses among chronic patients confirm the same problems of diagnostic shifting, as already shown by Kessing^{28,29}, between the clinical and OPCRIT ICD-10 diagnoses of BPD and MDD demonstrated by the overall predictive- and κ -values. Out of fifty subjects that shifted diagnoses, one fourth of the patients with clinical ICD-10 BPD and MDD converted diagnoses by OPCRIT to ICD-10 diagnoses outside the affective disorder spectrum, which emphasise recent considerations³⁰⁻³² on the issue of the borders between psychotic mood disorders and the schizophrenia spectrum.

The low sensitivity of the clinical BPD diagnoses indicates that a considerable number of patients suffer from unrecognised BPD as shown by Akiskals group^{33,34}. While, the low positive predictive value of the clinical BPD diagnoses suggests that a substantial fraction of these patients are misdiagnosed. These observations may have severe clinical implications in terms of treatment and prognoses.

BPD treatment regimes incl. mood stabilizers may be of less harm to MDD patients but antidepressant treatment regimes with-

out mood stabilizers may be hazardous^{35,36} to genuine BPD patients. Albeit these considerations, it is somewhat reassuring that the negative predictive value indicate that the risk of overlooking true BPD diagnoses among affective disorder patients are lower.

Only a minority of the clinically affective disorder patients have true schizophrenia spectrum disorders according to OPCRIT, which imply that the majority of diagnostic changes are confined within the affective spectrum. As a consequence, the four key values of diagnostic reliability (positive and negative predictive values, sensitivity and specificity) of MDD are mirror images of those of BPD. These results is in accordance with those found in large studies³⁷⁻³⁹.

These findings were unexpected considering the subjects lifetime history of ECT and median duration of illness of almost twenty years, but confirm the problems on the reliability of affective disorder diagnoses as already shown in large Danish epidemiologic studies^{40,41}. We did anticipate the clinical diagnoses of these chronic patients to be more reliable due to their long careers of known psychiatric illness.

These uncertainties on the reliability of clinical ICD-10 diagnoses do have research implications. The sensitivity and predictive values of the clinical diagnoses is overall moderate, indicating that clinical ICD-10 affective disorder diagnoses are unreliable, even when originating from very chronic patients. Clinical ICD-10 affective disorders diagnoses therefore seem unwarranted for research purpose. Standardised diagnostic instruments, e.g. the OPCRIT checklist, seem to be highly recommendable in order to obtain valid affective disorder diagnoses regardless the setting.

The reasons for the lacking reliability of clinical ICD-10 affective disorder diagnoses

may either be of nosocomial or psychopathological origin. These considerations are illustrated by the fact that a subset of OPCRIT ICD-10 BPD patients did not report and had no records of manic or mixed episodes when using the DPB-schema for basic clinical information. While, using the OPCRIT procedure revealed such episodes, causing a diagnostic shift from clinical MDD to OPCRIT ICD-10 BPD status. The opposite was the case on a similar subset of clinical BPD patients that shifted to OPCRIT ICD-10 MDD status.

In clinical practice psychiatrists are forced to focus on the clinical representation of the illness in the need of treatment for the particular patient during the distinct affective disorder episode. If depression is the main representation of a specific episode, previous hypomanic and manic episodes may pass unobserved by both the doctor and the patient. Particularly, if the hypomanic phases are an integrated part of the patient's life as periods of extra vitality.

The clinical over-diagnoses of BPD in this sample are most probable due to nosocomial factors relating to the long duration of illness, which extent to the former use of ICD-8 diagnoses and the transition to ICD-10, 1994 (ICD-9 was never introduced in Denmark). Remnants of previous ICD-8 classification may be preserved within the ICD-10 diagnoses among the oldest patients.

The high concordance between OPCRIT ICD-10 and DSM-IV BPD and MDD diagnoses reflects the similarities between these two diagnostic systems on the specific disorders. However, the lower concordance of the general affective vs. non-affective diagnoses presumably reflects the differences between ICD-10 and DSM-IV with respect to first rank symptoms within the affective spectrum.

The low reliability of affective disorder diagnoses, demonstrated by the present study on chronic subjects, highlights the loss of construct validity when using the present conventional operational classifications. The ICD-10 diagnoses of Single Depression, mild episode (F32.0) and Bipolar Disorder, other type (F31.8) illustrates this. Most dysphoric patients, regardless other psychiatric conditions except mania, will fulfil the criteria for ICD-10 Mild Depression.

Similarly, making a distinction between Bipolar I & II Disorders is impossible in ICD-10. Instead one has to use, e.g. F31.8 Bipolar Disorder, other type, in order to classify Bipolar II in ICD-10.

Regarding the future ICD-11 and DSM-V one must hope that the concept of depression will be more exact⁴² and that of Bipolar Disorders further differentiated and specified^{43,44}. Classifications of distinctive affective disorders on the basis of aetiology or pathophysiology are in their waiting as the present endophenotypes are vague. One good example though, is seasonal affective disorder, which respond to natural or artificial light^{45,46}.

The non-psychopathological assessment of the patients also raises several important issues that need commenting. Firstly, age, age at onset, age at first treatment, numbers of depressive episodes, suicidal- and Ham-D scores do not differ between gender or disease, thus confirming that these severely and chronically ill, affective disorder patients are highly homogeneous and randomly sampled regardless disease category and sex.

Still, demographic and epidemiological differences do emerge; most notably, age at first hospitalisation is significantly lower and duration of illness is significantly longer among BPD subjects, which may reflect the wider (albeit non-significant)

age-span on age at onset among MDD than BPD patients. These results are in accordance with previous findings⁴⁷ that BPD patients fall ill teen years earlier and have twice as many hospitalisations than MDD subjects.

Further, differences between men and women were seen. Numbers of depressions and treated depressions are significantly higher, whereas GAF and CGI scores are significantly lower in women. Our data seem to indicate that women in this highly chronic sample of depressive patients are more dysfunctional (lower GAF) and either seek or are hospitalised more frequently than men, who appear more severely ill (higher CGI). This was not anticipated and may reflect distinct disease entities of depression across gender or be due to cultural bias⁴⁸ related to the sex specific behaviour of patients and gender influence on the psychiatric care.

The weakness with regard to the sample is the recruitment of subjects among chronic hospitalised ECT patients with BPD and MDD. Such a cohort of severely ill may not generalise to the background population. Still, the sample is representative of the present hospital psychiatry in the greater Copenhagen metropolitan area.

This study focuses on a highly selected subpopulation of affective disorder patients to examine whether diagnostic reliability and anamnestic homogeneity would increase compared to more representative samples. The overall conclusions, nonetheless, are that clinical ICD-10 diagnoses of affective disorders of chronic severely ill subjects are not more reliable than those reported from broader and more representative samples.

Acknowledgements

The study was financed by grants to Thomas Werge from the Copenhagen Hospital Corporation Research Fond, the Danish National Psychiatric Research Foundation, the Danish Agency for Science, Technology and Innovation (Centre for Pharmacogenomics) and the Danish Medical Research Council.

References

1. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic Criteria for Research. 1993. Geneva, WHO.
2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, International Version. Washington, DC: APA; 1995.
3. Farmer A, McGuffin P, Williams J. Defining and Classifying Disorder. Measuring Psychopathology. Oxford, UK: Oxford University Press; 2002. p. 42-59.
4. Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* 2003; 160(4): 636-645.
5. Akiskal HS, Pinto O. The evolving bipolar spectrum. Prototypes I, II, III, and IV. *Psychiatr Clin North Am* 1999; 22(3): 517-534, vii.
6. Akiskal HS, Benazzi F. The DSM-IV and ICD-10 categories of recurrent [major] depressive and bipolar II disorders: evidence that they lie on a dimensional spectrum. *J Affect Disord* 2006; 92(1): 45-54.
7. Akiskal HS, Akiskal KK, Lancrenon S et al. Validating the bipolar spectrum in the French National EPIDEP Study: overview of the phenomenology and relative prevalence of its clinical prototypes. *J Affect Disord* 2006; 96(3): 197-205.
8. Vieta E, Phillips ML. Deconstructing bipolar disorder: a critical review of its diagnostic validity and a proposal for DSM-V and ICD-11. *Schizophr Bull* 2007; 33(4): 886-892.
9. Kessler RC, Akiskal HS, Angst J et al. Validity of the assessment of bipolar spectrum disorders in the WHO CIDI 3.0. *J Affect Disord* 2006; 96(3): 259-269.

10. Moreno C, Laje G, Blanco C, Jiang H, Schmidt AB, Olfson M. National trends in the outpatient diagnosis and treatment of bipolar disorder in youth. *Arch Gen Psychiatry* 2007; 649: 1032-1039.
11. Craddock N, Jones I, Kirov G, Jones L. The Bipolar Affective Disorder Dimension Scale BADD(S) –a dimensional scale for rating lifetime psychopathology in bipolar spectrum disorders. *BMC Psychiatry* 2004; 4: 19.
12. McGuffin P, Farmer A, Harvey I. A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system. *Arch Gen Psychiatry* 1991; 488: 764-770.
13. Williams J, Farmer AE, Ackenheil M, Kaufmann CA, McGuffin P. A multicentre inter-rater reliability study using the OPCRIT computerized diagnostic system. *Psychol Med* 1996; 264: 775-783.
14. DPS. ECT behandling i Danmark. ECT udvalgets betænkning. Dansk Psykiatrisk Selskab. 2002.
15. Guy W. Early Clinical Drug Evaluation ECDEU) Assessment Manual. Rockville, US, Nat Inst Ment Health. 1976.
16. Endicott J, Spitzer RL, Fleiss JL, Cohen J. The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry* 1976; 336: 766-771.
17. World Health Organisation. Homepage on Suicide. WHO, editor. 2007.
18. McGuffin P, Farmer A, Harvey I. A polydiagnostic application of operational criteria in studies of psychotic illness. Development and reliability of the OPCRIT system. *Arch Gen Psychiatry* 1991; 488: 764-770.
19. Williams J, Farmer AE, Ackenheil M, Kaufmann CA, McGuffin P. A multicentre inter-rater reliability study using the OPCRIT computerized diagnostic system. *Psychol Med* 1996; 264: 775-783.
20. Leckman JF, Sholomskas D, Thompson WD, Belanger A, Weissman MM. Best estimate of lifetime psychiatric diagnosis: a methodological study. *Arch Gen Psychiatry* 1982; 398: 879-883.
21. Azevedo MH, Soares MJ, Coelho I et al. Using consensus OPCRIT diagnoses. An efficient procedure for best-estimate lifetime diagnoses. *Br J Psychiatry* 1999; 175: 154-157.
22. Craddock M, Asherson P, Owen MJ, Williams J, McGuffin P, Farmer AE. Concurrent validity of the OPCRIT diagnostic system. Comparison of OPCRIT diagnoses with consensus best-estimate lifetime diagnoses. *Br J Psychiatry* 1996; 1691: 58-63.
23. Jakobsen KD, Frederiksen JN, Hansen T, Jansson LB, Parnas J, Werge T. Reliability of clinical ICD-10 schizophrenia diagnoses. *Nord J Psychiatry* 2005; 593: 209-212.
24. Jakobsen KD, Frederiksen JN, Parnas J, Werge T. Diagnostic agreement of schizophrenia spectrum disorders among chronic patients with functional psychoses. *Psychopathology* 2006; 396: 269-276.
25. Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* 1967; 64: 278-296.
26. Endicott J, Spitzer RL, Fleiss JL, Cohen J. The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. *Arch Gen Psychiatry* 1976; 336: 766-771.
27. Altman GD. Some Common Problems in Medical Research. *Practical Statistics for Medical Research*. 396-435. London, UK, Chapman & Hall/CRC. 1999.
28. Kessing LV. Diagnostic stability in depressive disorder as according to ICD-10 in clinical practice. *Psychopathology* 2005; 381: 32-37.
29. Kessing LV. Diagnostic stability in bipolar disorder in clinical practise as according to ICD-10. *J Affect Disord* 2005; 853: 293-299.
30. Lake CR, Hurwitz N. Schizoaffective disorders are psychotic mood disorders; there are no schizoaffective disorders. *Psychiatry Res* 2006; 143 (2-3): 255-287.
31. Lake CR, Hurwitz N. Schizoaffective disorder merges schizophrenia and bipolar disorders as one disease-there is no schizoaffective disorder. *Curr Opin Psychiatry* 2007; 204: 365-379.
32. Lake CR. Hypothesis: Grandiosity and Guilt Cause Paranoia; Paranoid Schizophrenia is a Psychotic Mood Disorder; a Review. *Schizophr Bull* 2007 December 1.
33. Akiskal HS, Benazzi F. Atypical depression: a variant of bipolar II or a bridge between unipolar and bipolar II? *J Affect Disord* 2005; 842-3: 209-217.
34. Akiskal HS, Benazzi F. The DSM-IV and ICD-10 categories of recurrent [major] depressive and bipolar II disorders: evidence that they lie on a dimensional spectrum. *J Affect Disord* 2006; 921: 45-54.
35. Post RM, Altshuler LL, Frye MA et al. Rate of switch in bipolar patients prospectively treated with second-generation antidepressants as augmentation to mood stabilizers. *Bipolar Disord* 2001; 35: 259-265.
36. Post RM, Leverich GS, Nolen WA et al. A re-evaluation of the role of antidepressants in the treatment of bipolar depression: data from the Stanley Foundation Bipolar Network. *Bipolar Disord* 2003; 56: 396-406.

37. Akiskal HS, Akiskal KK, Lancrenon S et al. Validating the bipolar spectrum in the French National EPIDEP Study: overview of the phenomenology and relative prevalence of its clinical prototypes. *J Affect Disord* 2006; 963: 197-205.
38. Akiskal HS, Akiskal KK, Lancrenon S, Hantouche E. Validating the soft bipolar spectrum in the French National EPIDEP Study: the prominence of BP-II 1/2. *J Affect Disord* 2006; 963: 207-213.
39. Akiskal HS, Benazzi F. The DSM-IV and ICD-10 categories of recurrent [major] depressive and bipolar II disorders: evidence that they lie on a dimensional spectrum. *J Affect Disord* 2006; 921: 45-54.
40. Kessing LV. Diagnostic stability in bipolar disorder in clinical practise as according to ICD-10. *J Affect Disord* 2005; 853: 293-299.
41. Kessing LV. Diagnostic stability in depressive disorder as according to ICD-10 in clinical practice. *Psychopathology* 2005; 381: 32-37.
42. Kessing LV. Epidemiology of subtypes of depression. *Acta Psychiatr Scand Suppl* 2007; 433: 85-89.
43. Vieta E. Defining the bipolar spectrum and treating bipolar II disorder. *J Clin Psychiatry* 2008; 694: e12.
44. Vieta E, Suppes T. Bipolar II disorder: arguments for and against a distinct diagnostic entity. *Bipolar Disord* 2008; 101 Pt 2: 163-178.
45. Westrin A, Lam RW. Seasonal affective disorder: a clinical update. *Ann Clin Psychiatry* 2007; 194: 239-246.
46. Westrin A, Lam RW. Long-term and preventative treatment for seasonal affective disorder. *CNS Drugs* 2007; 2111: 901-909.
47. Spiessl H, Hubner-Liebermann B, Cording C. [Differences between unipolar and bipolar affective disorders. Review and results from a clinical population]. *Fortschr Neurol Psychiatr* 2002; 708: 403-409.
48. Bjerkeset O, Romundstad P, Gunnell D. Gender differences in the association of mixed anxiety and depression with suicide. *Br J Psychiatry* 2008; 192: 474-475.

Address for correspondence:
Klaus Damgaard Jakobsen, MD
Section 807
Mental Health Center Hvidovre
Broenbyoestervej 160
DK-2605 Broendby
Denmark
Phone. +45 3632 3836
Fax. +45 3632 3889
E-mail. kdjakobsen@gmail.com