Isr J Psychiatry Relat Sci Vol 43 No. 1 (2006) 34–39

Changes in Stability of First-Admission Psychiatric Diagnoses Over 14 Years, Based on Cross-Sectional Data at Three Time Points

Alexander M. Ponizovsky, MD, PhD,[1] Alexander Grinshpoon, MD, MHA,[1] Inna Pugachev, MA,[1] Daniela Nahon, BA,[1] Michael Ritsner, MD, PhD,[2] and Moshe Z. Abramowitz, MD[1]

[1] *Mental Health Services, Department of Research and Evaluation, Ministry of Health, Jerusalem, Israel*

[2] *Sha'ar Menashe Mental Health Center, Hadera, and the Bruce Rappaport Faculty of Medicine, Technion, Haifa, Israel*

Abstract: Background: Stability of diagnoses over time is an important criterion of reliability of any diagnostic system. *Aims:* To compare the stability of diagnoses among hospitalized psychiatric patients over 14 years, based on cross-sectional data at three time points. *Method*: Diagnoses extracted from the National Psychiatric Case Register concerning psychiatric patients first admitted during 1989 (n=2,996) were compared to those who were admitted during 1996 (n=3,021) and 2003 (n=4,041). Stability of diagnostic categories was measured by positive predictive value of admission diagnosis at discharge from hospital. *Results*: There was no significant difference in diagnostic stability for most diagnostic categories between 1989 and 1996 patient cohorts. However, over the seven following years (cohort 2003) the diagnostic stability had been substantially increased: by 19% for affective disorders, by 18% for childhood disorders, by 17% for organic conditions, by 14% for neurotic disorders, by 12% for both schizophrenia and drug and alcohol dependence. *Conclusions:* In long-term perspective, reliability of most diagnostic categories of mental disorders has clearly improved.

Diagnosis (from the Greek words *dia* = by and *gnosis* = knowledge) is the process of identifying a disease by its signs, symptoms and results of various diagnostic procedures as well as the conclusion reached through that process (1). In the absence of laboratory tests and techniques based on a solid understanding of pathogenesis, criteria available to psychiatry for validating diagnostic categories have been restricted to clinical features, outcome and family history (2). A clinician making an initial diagnosis of psychiatric disorder in a first-episode patient at admission in psychiatric hospital has insufficient information on family history and he/she, clearly, cannot foresee outcome even in short-term. Under these circumstances, a clinician's initial diagnosis is based exceptionally on signs and symptoms of a specific disorder (a "diagnosis of recognition"). Most signs and symptoms, however, have poor reliability, with most clinicians disagreeing about their significance for making a specific diagnosis (3). As the signs and symptoms have such low predictive value, unreliable "diagnosis of recognition" requires further verifying after receiving more pertinent information for a longer period of observation; such verifying diagnosis is usually available at time of discharge from hospital.

The discrepancy between the two types of diagnosis refers to an important issue — the stability of psychiatric diagnosis over time. The primary aim of our study was to report diagnostic stability, defined as the proportion of patients who received a diagnosis at discharge from psychiatric hospital in the same main category as admission diagnosis. Our secondary aim was to examine to what extent the transition from ICD-9 to ICD-10 could affect diagnostic stability of any diagnostic category.

The introduction of ICD-10 (4) worldwide in the beginning of the 1990s is considered as a progress allowing psychiatry to approach the rest of medicine (5). It was hoped that the improved conceptualization of the ICD-10 and the provision of extensive diagnostic guidelines would contribute to better diagnostic reliability, but it did not necessarily alter diagnostic habits of experienced psychiatrists. These

Address for Correspondence: A. Ponizovsky, MD, PhD, Mental Health Services, Ministry of Health, 2 Ben Tabai St., Jerusalem 93591, Israel. E-mail: alexander.ponizovsky@moh.health.gov.il

criteria, however, are being increasingly regarded by residents and younger colleagues as reflecting the essence or nature of the psychiatric disorder in question. The question of whether the implementation of ICD-10 could affect stability of hospital psychiatrists' diagnoses remains unclear. In the present study we sought to 1) compare the patterns of diagnoses between two patient cohorts first-admitted in all Israel psychiatric settings during 1989 and 1996, when ICD-9 and ICD-10, respectively, were used, and 2) determine the relative stability of ICD-10 diagnostic categories, by comparing the admission and discharge diagnosis in 1996 and 2003 inpatient cohorts.

Method

Data collection

This cross-sectional study was designed as a diagnostic cohort study of data coming from the Israeli National Psychiatric Hospitalization Registry of the Ministry of Health that contains complete information on all psychiatric admissions in Israel since 1950. Approximately 16,000 hospitalizations are recorded annually; almost 4,000 of them are first-time admissions (6). In this study, immigrants with fewer than five years in Israel were excluded from the analysis (N=712 in 1989 and N=983 in 1996) due to uncertainty of information on their hospital status in the country of origin. After this exclusion, demographic characteristics and primary diagnoses of the remaining 2,996 persons who had their first psychiatric admission in 1989 (cohort 1) were directly compared with corresponding data on the 3,021 persons who were hospitalized for the first time in 1996 (cohort 2) and the 4,041 persons whose first admission was in 2003 (cohort 3). ICD-9 diagnoses for cohort 1 and ICD-10 diagnoses for cohorts 2 and 3 were utilized. Diagnostic stability was measured by positive predictive values (PPVs) of admission diagnoses in predicting the diagnosis upon discharge. For this, the percentage of admission diagnoses remaining unchanged at the time of discharge from the hospital for each diagnostic category was calculated: the higher this proportion, the higher PPV of primary diagnosis, and thus the greater stability.

Data analysis

The data were analyzed using NCSS version 2000 for Windows (Number Cruncher Statistical Systems; Kaysville, Utah, 1998). Chi-square statistics were used for testing the significance of differences in proportions. Kappa reliability test (Cohen's κ) (7) was used in order to measure the degree of concordance between admission and discharge diagnoses within each cohort. According to Landis and Koch (8), who proposed a scale to describe the degree of concordance, we considered κ magnitudes ranging from 0.61 to 0.80 as "substantial" or satisfactory. Mann-Whitney tests for non-matched (independent) samples were used to compare PPVs of diagnostic categories across the cohorts. For all analyses, the level of statistical significance was defined as α less than 0.05.

Table 1. *Gender and age breakdown in three cohorts of first admitted psychiatric patients*

Age group, yr.	Cohort 1989 (N=2,998)				Cohort 1996 (N=3,021)				Cohort 2003 (N=4,041)			
	Male		Female		Male		Female		Male		Female	
	N	%	N	%	N	%	N	%	N	%	N	%
11-19	333	19.7	216	16.5	229	13.3	218	16.8	430	18.6	315	18.2
20-29	536	31.7	297	22.7	606	35.4	292	22.5	717	17.7	394	22.8
30-39	313	18.5	211	16.1	295	17.1	219	16.9	437	18.9	253	14.6
40-49	155	9.2	115	8.8	189	11.0	167	12.9	278	12.0	238	13.8
50-59	107	6.3	130	9.9	112	6.5	117	9.0	191	8.2	199	11.5
60 +	245	14.5	340	26.0	222	12.9	285	22.0	258	11.2	331	19.1
Total	1689	100.0	1309	100.0	1723	100.0	1298	100.0	2311	100.0	1730	100.0

Cohort 1989 vs. cohort 1996: χ^2= 14.8, d.f.=5, p<.01; Cohort 1996 vs. cohort 2003: χ^2= 3.06, d.f.=5, p>.05

Table 2. *Positive predictive value (PPV) of ICD diagnostic categories at discharge from hospital in three cohorts*

Diagnosis	Cohort 1989 (n=2,996) Number of diags. at admis.	PPV at discharge (%)	k_1	Cohort 1996 (n=3,021) Number of diags. at admis.	PPV at discharge (%)	k_2	z_1	Cohort 2003 (n=4,041) Number of diags. at admis.	PPV at discharge (%)	k_3	z_2
Organic conditions	253	77.5	0.70	201	72.6	0.75	1.12	243	89.0	0.68	4.49[2]
Schizophrenia and other nonorganic psychoses	998	68.0	0.60	982	81.6	0.74	6.92[2]	1343	94.2	0.68	9.80[2]
Affective disorders	455	67.9	0.63	688	65.4	0.60	0.85	1013	83.8	0.62	8.77[2]
Neurotic disorders	297	75.1	0.61	304	69.1	0.62	1.64	467	82.9	0.60	4.71[2]
Personality disorders	247	78.1	0.56	137	62.8	0.59	3.23[2]	161	56.3	0.47	1.26
Childhood disorders	62	54.8	0.48	56	62.5	0.62	0.84	131	80.4	0.62	2.69[2]
Drug and alcohol dependence	99	81.8	0.74	139	69.8	0.69	2.11[1]	177	81.9	0.59	2.66[2]
Mental retardation	42	83.3	0.60	52	80.8	0.71	0.32	42	82.3	0.62	0.21
Codes for administrative use	543	21.9	0.28	462	29.0	0.31	0.47	463	45.8	0.33	5.27[2]

Cohen's κ — concordance between admission and discharge diagnoses within Cohort 1989 (κ_1), Cohort 1996 (κ_2), and Cohort 2003 (κ_3)

Mann-Whitney non-matched two-samples tests: z_1 — Cohort 1989 vs. Cohort 1996; z_2 — Cohort 1996 vs. Cohort 2003;
[1] — $p< 0.05$; [2] — $p<0.001$

Results

Table 1 compares age and gender breakdowns in the three cohorts. The cohorts were comparable, except for female patients aged 40-49 and 60+ years, who were overrepresented in 1989 and 1996, respectively ($\chi^2= 14.8$, d.f.=5, $p<.01$).

Table 2 shows the stability of each diagnostic category at discharge from hospital as measured by PPV of the admission diagnosis and Cohen's κ agreement coefficient within each cohort. As can be seen, the degree of concordance was similarly satisfactory for most diagnostic categories, excepting the category of personality disorders (κ=0.56, 0.59 and 0.47, respectively) and codes for administrative use (κ=0.28, 0.31 and 0.33, respectively) for all cohorts as well as for childhood disorders in cohort 1989 (κ=0.48), and drug/alcohol dependence in cohort 2003 (κ=0.59). The highest concordance rates across the cohorts were noted for the categories of organic conditions (κ=0.70, 0.75, and 0.68, respectively), and drug/alcohol dependence in cohort 1989, and schizophrenia in cohort 1996 and 2003 (κ=0.74 and 0.68, respectively).

To establish the temporal stability of diagnoses, between-cohort comparisons of the corresponding PPV were made (Table 2). There was no significant difference in PPV for most diagnostic categories between 1989 and 1996 patient cohorts. The exceptions were the three categories: schizophrenia, stability of which increased over this time by 14% ($z=6.92$, $p<0.001$), and personality disorders and drug/alcohol dependence, stabilities of which decreased, respectively, by 15% ($z=3.23$, $p<0.001$) and 12% ($z=2.11$, $p<0.05$). However, over the following seven years (from cohort 1996 to 2003) the diagnostic stability of the six categories substantially increased: by 19% for affective disorders ($z=8.77$), by 18% for childhood disorders ($z=2.69$), by 17% for organic conditions ($z=4.49$), by 14% for neurotic disorders ($z=4.71$), by 12% for both schizophrenia ($z=9.80$) and drug and alcohol dependence ($z=2.66$,

all p<0.001). The PPV of diagnosis of mental retardation remained unchanged and that of personality disorders decreased, yet not significantly.

Discussion

In this study, we explored the stability of diagnoses among three epidemiologically defined cohorts of patients who were first admitted in Israel psychiatric settings during 1989, 1996 and 2003. We examined diagnostic stability in two ways: 1) as the degree to which the admission diagnosis remains unchanged at the time of discharge, and 2) as the degree to which ICD-9 diagnostic rates during 1989 correspond with ICD-10 diagnostic categories during 1996 and 2003. We found that diagnostic stability of most diagnostic categories remained unchanged under the ICD-10, but it clearly improved for the category of schizophrenia including other non-organic psychoses.

Limitations

To better understand and further the discussion of the study results, several limitations should be addressed here. When one examines changes over time, it must be kept in mind that there are intrinsic difficulties in causal attribution of the changes to a single factor, since many potentially confounding factors also change over this temporal interval. In our case, it is difficult to say that the change in ICD classification is the only factor involved or that changes in the stability of diagnosis are due to this factor alone. From 1989 to 1996, many potentially relevant changes in psychiatric practice and diagnosis occurred in the country. For example, changes in the Mental Health Act (9) concerning involuntary hospitalization order, change in policy of the hospital (do not make a definite diagnosis at admission), and the introduction of novel medications (antipsychotics and antidepressants), have probably accounted for some of the changes in the diagnostic consistency over time. Moreover, a shorter time spent in hospital (average length of stay was 286, 151 and 120 days in 1989, 1996 and 2003, respectively) as has been the trend for the last 14 years may cause the diagnosis upon admission to resemble the discharge diagnosis more closely. The mass influx of immigrant physicians into the mental health system (10) has also resulted in accentuating the need to unite all diagnosticians in Israel under a shared conceptual framework. In this respect, the publication of the Hebrew translation of the ICD-10 in 1996 and its being adopted by the Ministry of Health as the official nomenclature has increased the possibility of creating such a much needed framework. Any of these factors acting alone or in combination may serve as additional explanations for the change of diagnosis from admission to discharge between the time points studied.

Using a retrospective database from the National Psychiatric Case Register, in which the quality of the data depends on accurate reporting by hospitals, and that do not include variables reflecting the environmental or situational context in which the diagnostic process is likely to occur, we were not able to address all these real-life questions. However, unlike most studies, which are based on highly selected populations and data sources, this report, relying on the large-scale cohorts of all first admissions from national population, provides unique information about stability of psychiatric diagnosis in Israel for two seven-year periods, independent of other factors by which this stability could been explained. With these reservations, our findings can be attributed to change in classification system.

Diagnosis stability

Overall, our data show the relative stability of most diagnoses in both diagnostic systems, with the highest level of stability for mental retardation and the lowest one for childhood disorders. Consistent with most, but not all, previous polydiagnostic studies (11-16), we found that of all diagnostic categories, schizophrenia was the most stable diagnosis in the ICD-10 cohort. The high stability of ICD-10 schizophrenia in this study (PPV=81.6% in 1996 and 94.2% in 2003) confirmed findings from long-term follow-up studies reporting the stability estimates from 89% to 93% (14-18). In contrast to schizophrenia, the level of stability of affective disorders in 1989 and 1996 (PPVs=67.9% and 65.4%, respectively) is much lower than that in 2003 (83.8%) and previous estimates for bipolar/manic psychoses (86%) (13). It should be kept in mind, however, that differences across studies can result from methodology differences, e.g., the use of more stringent research diag-

nostic criteria versus more arbitrary registry diagnoses, the quality of which depends on accurate reporting by hospitals. A recent paper in Israel reported that in 87% to 89% of cases with psychotic disorder or with schizophrenia the case registry agreed with Research Diagnostic Criteria diagnoses, demonstrating a high level of sensitivity of the registry diagnosis (19).

We analyzed the PPV of a range of psychoses including schizoaffective and delusional disorders that are known for their lower diagnostic stability (20, 21). It is possible that in separate analysis of this large group other results would be obtained. We suggest, however, that the high stability of schizophrenias as a group which appears might be due to the balance between the two opposite tendencies in diagnosis of these two unstable components. Because the diagnosis of schizoaffective disorder tends to migrate to a group of affective (mood) disorders (11), while that of delusional psychosis to direction of schizophrenia (17), stability of the whole group remains unchanged from admission to discharge. Previous findings, that diagnosis of psychotic disorder is distinguished by greater stability than those of non-psychotic conditions (13), were confirmed by our data only for ICD-10. However, even the stability rates of ICD-10 personality disorders (62.8%) were found to be twice as high as compared to what has been previously reported (36%) (22). The diagnostic stability ratings of ICD-9 personality disorders (78.1%) and neurotic disorders (75.1%) were higher than those of schizophrenia and affective disorders. Low stability of the ICD-10 diagnostic category of personality disorders among first-admitted patients could be explained by an increasing consensus among mental health professionals that most patients with such diagnoses should be treated in outpatient settings and referred to the hospital either when they constitute a suicide risk or a danger to others (23). This tendency results in hospitalization of a more disturbed population in which a correct diagnosis at admission may be more difficult to obtain. A similar explanation is applicable to the lack of stability for the diagnostic category of drug/alcohol dependence. Current research supports the notion that the stability of a diagnosis of alcohol dependence depends on the severity of the illness (24).

Perhaps another explanation is the growing reluctance of mental health professionals to tag first-admitted individuals immediately with a psychiatric diagnosis, with all the far-reaching repercussions attached to such diagnoses. In lieu of this, many diagnosticians prefer to defer diagnosis upon admission, or to use a more generic Z code (i.e., Z03.2-Observation for suspected mental and behavioral disorders) for the purpose of filling the admission forms.

Diagnostic discrepancy

Changes in ICD-10 were guided by empirical data focused on improving diagnostic validity and reliability. The main advantages of ICD-10 compared to previous ICD-9 classification comprise: 1) the use of criteria-oriented instead of typological diagnostic approach; 2) the more operationalized descriptions of the diagnoses; 3) similarity in structure and terms to the DSM system; and 4) provision of diagnostic guidelines (25). The major source of disagreement between the two classification systems is the relocation of several diagnostic categories to different sections. Data from many field trials demonstrates that ICD-10 is comparable to ICD-9 in terms of reliability (11, 12). Our study confirms this conclusion for most diagnostic categories, although significant discrepancies between certain categories of the two diagnostic systems were observed. For example, from ICD-9 to ICD-10 the stability of schizophrenia diagnosis rose by 14%, with reciprocal stability decline of personality disorders and drug and alcohol dependence by 15% and 12%, respectively. The differences are highly statistically significant, and it is unlikely that they could result from variations in the samples. Therefore, we suggest the differences as intrinsic for diagnostic criteria of the systems, yet we cannot rule out that diagnostic habits of psychiatrists may have changed over the years.

In conclusion, diagnostic stability of most diagnostic categories remained unchanged under the ICD-10 compared to ICD-9, but it clearly improved for the category of schizophrenia including other non-organic psychoses. The problems of validity and reliability of psychiatric diagnosis are often conflated in the literature (26, 27). A demonstration of a high stability of diagnoses, i.e., reliability of a given diagnostic systems, does not mean that the system is valid. Therefore, further efforts should be directed to investigate the validity of psychiatric diagnoses.

Acknowledgements

AP was supported in part by the Ministry of Absorption.

References

1. Wikipedia, 2005. Available: http://en.wikipedia.org.
2. Robins E, Guze SB. Establishment of diagnostic validity in psychiatric illness: Its application to schizophrenia. Am J Psychiatry 1970;126:983-987.
3. Verdoux H, van Os J. Psychotic symptoms in non-clinical populations and the continuum of psychosis. Schizophr Res 2002;54:59-65.
4. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders. Diagnostic Criteria for Research. Geneva: WHO, 1993.
5. Parnas J, Zahavi D. The role of phenomenology in psychiatric classification and diagnosis. In: Maj M, Gaebel W, Lopez-Ibor JJ, Sartorius N, editors. Psychiatric diagnosis and classification. Chichester, U.K.: Wiley, 2002: pp. 137-162.
6. Lichtenberg P, Kaplan Z, Grinshpoon A, Feldman D, Nahon D. The goals and limitations of Israel's psychiatric case register. Psychiatr Serv 1999;50:1043-1048.
7. Cohen JA. A coefficient of agreement for nominal scales. Educ Psych Meas 1960;20:37-46.
8. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometr 1977; 33:363-374.
9. Mental Health Act, State of Israel, 1991.
10. Ponizovsky A, Ginath Y, Factourovich A, Levin K, Maoz B, Ritsner M. The impact of professional adjustment on the psychological distress of immigrant physicians. Stress Med 1996;12, 247-251.
11. Marneros A, Deister A, Rohde A. Stability of diagnoses in affective, schizoaffective and schizophrenic disorders. Cross-sectional versus longitudinal diagnosis. Eur Arch Psychiatry Clin Neurosci 1991;241:187-192.
12. Fennig S, Kovasznay B, Rich C, Ram R, Pato C, Miller A, Rubinstein J, Carlson G, Schwartz JE, Phelan J, et al. Six-month stability of psychiatric diagnoses in first-admission patients with psychosis. Am J Psychiatry 1994; 151:1200-1208.
13. Chen YR, Swann AC, Burt DB. Stability of diagnosis in schizophrenia. Am J Psychiatry 1996;153:682-686.
14. Mason P, Harrison G, Croudace T, Glazebrook C, Medley I. The predictive validity of a diagnosis of schizophrenia. A report from the International Study of Schizophrenia (ISoS) coordinated by the World Health Organization and the Department of Psychiatry, University of Nottingham. Br J Psychiatry 1997;170:321-327.
15. Amin S, Singh SP, Brewin J, Jones PB, Medley I, Harrison G. Diagnostic stability of first-episode psychosis. Comparison of ICD-10 and DSM-III-R systems. Br J Psychiatry 1999;175:537-543.
16. Huguelet P, Schneider El Gueddari N, Glauser D. Stability of DSM-III-R diagnoses: Study of a case register. Psychopathol 2001;34:118-122.
17. Veen ND, Selten JP, Schols D, Laan W, Hoek HW, van der Tweel I, Kahn RS. Diagnostic stability in a Dutch psychosis incidence cohort. Br J Psychiatry 2004;185: 460-464.
18. Bromet EJ, Naz B, Fochtmann LJ, Carlson GA, Tanenberg-Karant M. Long-term diagnostic stability and outcome in recent first-episode cohort studies of schizophrenia. Schizophr Bull 2005;31:639-649.
19. Weiser M, Kanyas K, Malaspina D, Harvey PD, Glick I, Goetz D, Karni O, Yakir A, Turetsky N, Fennig S, Nahon D, Lerer B, Davidson M. Sensitivity of ICD-10 diagnosis of psychotic disorders in the Israeli National Hospitalization Registry compared with RDC diagnoses based on SADS-L. Compr Psychiatry 2005;46:38-42.
20. Schwartz JE, Fennig S, Tanenberg-Karant M, Carlson G, Craig T, Galambos N, Lavelle J, Bromet EJ. Congruence of diagnoses 2 years after a first-admission diagnosis of psychosis. Arch Gen Psychiatry 2000;57:593-600.
21. Forrester A, Owens DG, Johnstone EC. Diagnostic stability in subjects with multiple admissions for psychotic illness. Psychol Med 2001;31:151-158.
22. Stanton MW, Joyce PR. Stability of psychiatric diagnoses in New Zealand psychiatric hospitals. Aust N Z J Psychiatry 1993;27:2-8.
23. Sanguineti VR, Samuel SE, Schwartz SL, Robeson MR. Retrospective study of 2,200 involuntary psychiatric admissions and readmissions. Am J Psychiatry 1996; 153:392-396.
24. Culverhouse R, Bucholz KK, Crowe RR, Hesselbrock V, Nurnberger JI Jr, Porjesz B, Schuckit MA, Reich T, Bierut LJ. Long-term stability of alcohol and other substance dependence diagnoses and habitual smoking: An evaluation after 5 years. Arch Gen Psychiatry 2005; 62:753-760.
25. Dittmann V, Freyberger HJ, Albus M, Blanz B, von Cranach M, Gastpar M, Gutzmann H, Maier W, Mombour W, Stieglitz RD. ICD-10 field trial in German-speaking countries — summary, judgement and perspectives. Pharmacopsychiatry 1990;23:202-204.
26. Jansson L, Handest P, Nielsen J, Saebe D, Parnas J. Exploring boundaries of schizophrenia: A comparison of ICD-10 with other diagnostic system in first-admitted patients. World Psychiatry 2002;1:109-114.
27. Kendell R, Jablensky A. Distinguishing between the validity and utility of psychiatric diagnoses. Am J Psychiatry 2003;160:4-12.