Volume 190, March 2007, pp.189-282

Psychiatry in pictures

Psychiatry in pictures


Highlights of this issue

Highlights of this issue
Sukhwinder S. Shergill

EDITORIALS

The bipolar spectrum
Jules Angst

Stigma: ignorance, prejudice or discrimination?
Graham Thornicroft, Diana Rose, Aliya Kassam, and Norman Sartorius

REVIEW ARTICLES

Bias in psychiatric case–control studies: Literature survey
Br J Psychiatry 2007 190: 204-209.

SPECIAL ARTICLES

Schizophrenia: a common disease caused by multiple rare alleles
Jon M. McClellan, Ezra Susser, and Mary-Claire King

Phenotypic and genetic complexity of psychosis: Invited commentary on ... Schizophrenia: a common disease caused by multiple rare alleles
Nick Craddock, Michael C. O'Donovan, and Michael J. Owen
Br J Psychiatry 2007 190: 200-203
**PAPERS**

**Diagnostic stability of psychiatric disorders in clinical practice**
Enrique Baca-Garcia, Maria M. Perez-Rodriguez, Ignacio Basurte-Villamor, Antonio L. Fernandez Del Moral, Miguel A. Jimenez-Arriero, Jose L. Gonzalez De Rivera, Jeronimo Saiz-Ruiz, and Maria A. Oquendo

**Using activity data to explore the influence of case-load size on care patterns**
Tom Burns, Jenny Yiend, Helen Doll, Tom Fahy, Matthew Fiander, and Peter Tyrer

**Patients discharged from medium secure forensic psychiatry services: reconvictions and risk factors**
Jeremy Coid, Nicole Hickey, Nadji Kahtan, Tianqiang Zhang, and Min Yang

**Anterior cingulate grey-matter deficits and cannabis use in first-episode schizophrenia**
Philip R. Szeszko, Delbert G. Robinson, Serge Sevy, Sanjiv Kumra, Claudia I. Rupp, Julia D. Betensky, Todd Lencz, Manzar Ashtari, John M. Kane, Anil K. Malhotra, Handan Gunduz-Bruce, Barbara Napolitano, and Robert M. Bilder

**Mortality in people with schizophrenia in rural China: 10-year cohort study**
Mao-Sheng Ran, Eric Yu-Hai Chen, Yeates Conwell, Cecilia Lai-Wan Chan, Paul S. F. Yip, Meng-ze Xiang, and Eric D. Caine

**Assessing insight in schizophrenia: East meets West**
Balasubramaninan Saravanan, K. S. Jacob, Shanthi Johnson, Martin Prince, Dinesh Bhugra, and Anthony S. David

**The Stigma Scale: development of a standardised measure of the stigma of mental illness**
Michael King, Sokratis Dinos, Jenifer Shaw, Robert Watson, Scott Stevens, Filippo Passetti, Scott Weich, and Marc Serfaty

**Risk of major adverse perinatal outcomes in women with eating disorders**
Nadia Micali, Emily Simonoff, and Janet Treasure

**Effect of antidepressant therapy on executive function after stroke**
Kenji Narushima, Sergio Paradiso, David J. Moser, Ricardo Jorge, and Robert G. Robinson
SHORT REPORTS

Effect of barriers on the Clifton suspension bridge, England, on local patterns of suicide: implications for prevention
Olive Bennewith, Mike Nowers, and David Gunnell

Impact of hospital consultants' poor mental health on patient care
Cath Taylor, Jill Graham, Henry Potts, Jenny Candy, Michael Richards, and Amanda Ramirez

Correspondence

Is DISC1 really a gene predisposing to psychosis?
T. J. Crow

Authors' reply
W. J. Muir, B. S. Pickard, and D. H. Blackwood

Hippocampal and amygdala volume reductions in first-episode schizophrenia
A. Vita and L. de Peri

Effective of cognitive–behavioural intervention by mental health nurses in schizophrenia
F. Alam

Authors' reply
D. Turkington

Contingency management for substance misuse
F. McQuaid, O. Bowden-Jones, and T. Weaver

Author's reply
N. M. Petry

Depression and anxiety after myocardial infarction
P. de Jonge and J. Ormel

Authors' reply
C. Dickens and F. Creed

High female suicide rates: ecological fallacy or sad reality?
E. Salib and G. Tadros

Authors' reply
K. Y. Liu and P. S. F. Yip

Self-poisoning with pesticides in India
R. Alex, J. Prasad, A. Kuruvilla, and K. S. Jacob
One hundred years ago

The increase of temperance
Henry Rollin

Book reviews

Surfacing Up: Psychiatry and Social Order in Colonial Zimbabwe, 1908–1968
Femi Oyebode

Bipolar Disorders: Mixed States, Rapid Cycling and Atypical Forms
Ingeborge Wilting

Alcohol, Gender and Drinking Problems: Perspectives from Low and Middle Income Countries
Oye Gureje

The science of well-being
Woody Caan

Social Capital and Mental Health
Scott Weich
Br J Psychiatry 2007 190: 278

The Cambridge Handbook of The Learning Sciences
Nick Rose

From the Editor's desk

From the Editor's desk
Peter Tyrer
ABCD: ALTERNATIVE GENETICS, BIAS, CANNABIS, AND DIMENSIONAL CLASSIFICATION

When, and how often, does depression become bipolar disorder? An editorial by Angst (pp. 189–191) offers a dimensional approach towards viewing these and related symptoms within a bipolar spectrum: a two-dimensional mood/affective spectrum. He discusses the additional value in taking this dimensional approach, which accounts for more recent clinical and genetic data indicating a large overlap, blurring the traditional distinction, between schizophrenia and bipolar disorder. McClellan et al (pp. 194–199) challenge the belief that schizophrenia arises partially as a consequence of people possessing an unfortunate selection of multiple common genes, each having a small effect. They suggest that the data can be interpreted to support the idea of schizophrenia stemming from a few rare genes with high penetrance. They examine the evidence supporting this hypothesis and conclude that a more informative approach may be to focus in more detail on the genome of individuals and families with schizophrenia or exposed cohorts, rather than current research designs that combine large samples of unrelated individuals, which will miss these rare genes of large effect. Case-control studies are one of the standard approaches to investigating interesting differences between patients and controls; the importance of ensuring that the design of the study minimises any systematic bias is often overlooked. Lee et al (pp. 204–209), in their literature survey of bias in psychiatric case-control studies, demonstrate that genetic studies were relatively poor in describing and controlling for selection bias. Neuroimaging studies showed good control of information bias, and one such study demonstrates that the volume of the grey matter within the anterior cingulate gyrus is reduced in patients with first-onset schizophrenia who used cannabis. Szeszko et al (pp. 230–236) suggest that this may be associated with the role of this region in decision-making and assessing risky outcomes.

STIGMA AND RISK FACTORS FOR REOFFENDING

Service users report that the rejecting behaviour of others may bring greater disadvantage than the primary condition itself. Thornicroft and colleagues (pp. 192–193) examine the underpinnings of stigma in their editorial, noting that this negative reaction is sufficiently powerful to cause significant autonomic arousal in people when they have been asked to imagine meeting a stigmatised individual. They emphasise that the focus needs to move from a more conceptual, or cognitive, appraisal of stigma to the functional consequences of stigma as evidenced by discriminatory practice. Continuing with this theme, King et al (pp. 248–254) describe the development of the Stigma Scale, a standardised instrument to measure stigma in mental illness. The three factors within the scale assess discrimination, disclosure and potential positive aspects of illness; it is of little surprise that the scores on the scale correlated negatively with self-esteem. One group that is arguably more stigmatised than others is patients who have been detained in secure forensic facilities. Coid et al (pp. 223–229) followed up a large group of these patients post-discharge and report that patients with two or more previous violent convictions, primary diagnosis of personality disorder, or comorbid antisocial personality disorder were more at risk of violent reoffending. They also suggest that sex offenders required particular vigilance in the 3–4 years after discharge, and that longer periods in security and restrictions on patients’ lifestyles following discharge were related to significant reductions in the risk of reoffending.

BACK TO BASICS: BRIDGES, EATING, CASE-LOADS AND ANTIDEPRESSANTS

Many acts of self-harm are impulsive in nature and restricting access to common methods, such as limiting the purchase of large numbers of tablets of commonly ingested analgesics, can result in reductions in suicides. Bennet et al (pp. 266–267) report that the number of suicides dropped from eight down to four per year following the installation of barriers on the Clifton suspension bridge. Although the immediate risks associated with eating disorders are well recognised, there is less awareness of the longer-term sequelae, particularly related to perinatal outcomes. Micali et al (pp. 255–259) examined a large longitudinal cohort and found bulimia to be associated with an increased rate of miscarriages, and anorexia nervosa to be associated with significantly lighter babies at birth. They comment that advising women with eating disorders about the possible effects on future fertility and adverse outcome for their offspring may provide an additional motivating factor to change their behaviour. Smaller case-loads are considered to offer improved service provision in psychiatric practice, but at least one recent study found no benefit of smaller (1:15) versus larger (1:35) case-loads. Burns et al (pp. 217–222) re-examine the data from this study and find a dose-response relationship between case-loads 1:10 and 1:20 and interpret this to support the value of smaller case-loads (below 1:20) for certain patient groups. Recent reports have indicated the value of antidepressants in reorganising neural circuits. Narushima et al (pp. 260–265) report that antidepressant treatment improved longer-term executive functioning following stroke. However, there were no improvements after acute treatment, and they interpret this later change as supporting the role of antidepressants in reorganising neural networks modulated by monoaminergic transmission.
Sujith Rathnayake (b. 1971). Picture selection and text by Ravimal Galappaththi

Sujith Rathnayake is an acclaimed artist from southern Sri Lanka. His life was thrown into turmoil by the tsunami disaster of 26 December 2004. He developed a psychotic illness and had flashbacks of his home being submerged by the tsunami waves. This set of untitled paintings reflects his inner conflict, restlessness, anxiety, disillusionment and confusion. He was admitted to a local psychiatric hospital and during his time there he produced many works of art. They were subsequently shown in an exhibition at the Lionel Wendt Art Gallery in Colombo. Painting and drawing, as well as religious rituals, provide an effective and culturally acceptable way of dealing with the loss of loved ones and material possessions.
The bipolar spectrum

JULES ANGST

Summary  The two-dimensional bipolar spectrum described here comprises a continuum of severity from normal to psychotic and a continuum from depression, via three bipolar subgroups to mania. This combination of dimensional and categorical principles for classifying mood disorders may help alleviate the problems of underdiagnosis and undertreatment of bipolar disorders.

Declaration of interest  None.

Depression is very distressful, prompts the depressed person to seek treatment and is relatively easy to diagnose. Hypomania, on the other hand, is often perceived as normal well-being and tends not to be reported. There is wide agreement as to the difficulty of identifying hypomania in bipolar II and minor bipolar disorders: patients with bipolar depression report that the recognition of their disorder was delayed by as much as 8–10 years. Unipolar depression, which is defined by the absence of hypomania or mania, remains an uncertain diagnosis lifelong. Over decades of recurrent depressive illness, bipolar disorder may manifest at any time: a lifelong follow-up of patients hospitalised showed a persistent risk of diagnostic change to bipolar disorder of 1.25% per year of observation.

Bipolar affective disorder is a more severe disorder than major depression, as measured by higher lifelong recurrence and greater comorbidity with psychiatric disorders, especially anxiety and secondary substance use disorders. In addition it is associated with serious somatic disorders such as diabetes, hypertension and cardiovascular disease. This explains the correspondingly higher mortality rates among people with bipolar disorder, although the suicide risk is lower in type I bipolar disorder than in depression (Osby et al, 2001).

Correct diagnosis of bipolar illness is essential for appropriate treatment, especially long-term secondary prophylaxis. As a consequence of their severity, unrecognised bipolar disorders lead, moreover, to higher costs than major depression, but these can be considerably reduced by early diagnosis and treatment, as recently shown by McCombs et al (2006).

The underdiagnosis of mood disorders, especially bipolar disorders, is not confined to the clinical setting but may also apply to traditional epidemiological studies, which found lifetime prevalence rates of 0.5–2%. Some recent studies (Kessler et al, 2003; Lieb, 2006) comprising two or three interview waves have described growing lifetime prevalence rates for both major depressive episodes (19% to 24%) and for bipolar disorder types I and II together (about 2% to 4%). An important question, then, is what proportion of patients with major depression should in fact be diagnosed as having bipolar disorder: is it one-fifth or one-tenth as generally reported, or as many as half, as we have found?

BIPOLAR SPECTRUM: A MODEL FOR RESEARCH AND CLINICAL PRACTICE

The development of a validated bipolar spectrum concept can provide a more differentiated research and treatment model for affective disorders and may help reduce the underrecognition of bipolarity.

A dimensional concept (from normal to pathological) was proposed by Kretschmer in 1921 for schizophrenia (schizothymic – schizoid – schizophrenic) and for affective disorders (cyclothymic temperament – cycloid ‘psychopathy’ – manic-depressive disorder) as well as by Bleuler (1922). The term ‘spectrum’ was first used in psychiatry in 1968 for the schizophrenia spectrum.

Fig. 1  Two-dimensional mood/affective spectrum (does not include schizoaffective disorder, as a transition to the schizophrenic spectrum). The precise relationship of personality disorders to the disease spectra is uncertain and an unsolved general problem of psychiatric classification. BP-I (I), bipolar-I disorder type I (II); D, major depression, d, minor depression; M, mania; m, hypomania; MDD, major depressive disorder; RBD, recurrent brief depression; sx, symptoms.

<table>
<thead>
<tr>
<th>SEVERITY SPECTRUM</th>
<th>MAJOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minor</td>
<td></td>
</tr>
<tr>
<td>Symptoms (normal)</td>
<td></td>
</tr>
<tr>
<td>No symptoms (supernormal)</td>
<td></td>
</tr>
<tr>
<td>Temperament (normal)</td>
<td></td>
</tr>
<tr>
<td>Affective personality disorders(^1)</td>
<td></td>
</tr>
<tr>
<td>Mood congruent or mood incongruent</td>
<td></td>
</tr>
<tr>
<td>Major mood disorders</td>
<td></td>
</tr>
<tr>
<td>Minor mood disorders</td>
<td></td>
</tr>
<tr>
<td>(sub-threshold)</td>
<td></td>
</tr>
<tr>
<td>Psychotic major mood disorders</td>
<td></td>
</tr>
<tr>
<td>Non-psychotic</td>
<td></td>
</tr>
<tr>
<td>Depressive personality disorder</td>
<td></td>
</tr>
<tr>
<td>Depressive temperament</td>
<td></td>
</tr>
<tr>
<td>Depressive temperament</td>
<td></td>
</tr>
<tr>
<td>MDD</td>
<td></td>
</tr>
<tr>
<td>BP-II</td>
<td></td>
</tr>
<tr>
<td>BP-I</td>
<td></td>
</tr>
<tr>
<td>MD</td>
<td></td>
</tr>
<tr>
<td>Hypomania</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SEVERITY SPECTRUM</th>
<th>MINOR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dysthymia</td>
<td></td>
</tr>
<tr>
<td>RBD</td>
<td></td>
</tr>
<tr>
<td>Minor depression</td>
<td></td>
</tr>
<tr>
<td>Depressive personality disorder</td>
<td></td>
</tr>
<tr>
<td>Depressive temperament</td>
<td></td>
</tr>
<tr>
<td>Minor BP</td>
<td></td>
</tr>
<tr>
<td>Cyclothymic disorder</td>
<td></td>
</tr>
<tr>
<td>Borderline/cyclot</td>
<td></td>
</tr>
<tr>
<td>Personality disorder</td>
<td></td>
</tr>
<tr>
<td>Cyclothymic temperament</td>
<td></td>
</tr>
<tr>
<td>MDD</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td></td>
</tr>
</tbody>
</table>

1. The precise relationship of personality disorders to the disease spectra is uncertain and an unsolved general problem of psychiatric classification.

Today the term ‘bipolar spectrum’ is mainly used in two complementary senses (Fig. 1):

(a) a spectrum of severity, which embraces psychotic and non-psychotic major and minor bipolar disorders (including bipolar dysthymia, recurrent brief and minor depressions), cyclothymic disorders, hypomania and, at its broadest, even borderline disorders and cyclothymic temperament;

(b) a proportional mood spectrum, which considers the two components mania and depression, first on the level of major mood disorders – major depression (D) bipolar I disorder (M), bipolar I disorder (MD), mania with mild depression (Md) and pure mania (M) (Angst, 1978) – and then on the level of minor mood disorders: the corresponding categories are mild depression (d), minor bipolar disorder (md) and hypomania (m).

This proportional model is an extension of Kleist’s concept of bipolar disorder as a combination of the two monopolar disorders depression and mania (Kleist, 1937, 1953). This model has proved fruitful not only in incorporating bipolar I and bipolar II disorders but also in differentiating mania with or without mild depression (Md/M) from bipolar I disorder (MD). Mania is not identical to bipolar I disorder in terms of family history, course and suicide risk, and on the sub-threshold level hypomania is not the same as minor bipolar disorder or cyclothymic disorder in terms of family history and temperament.

Both the above bipolar spectrum concepts are dimensional in nature, having no natural categorical subgroups. Epidemiological and clinical studies have demonstrated the continuous distribution of depressive and hypomanic/manic symptoms and syndromes from normal to pathological. Psychiatric symptoms in consulting populations have been shown to be dimensional (Goldberg, 2000). Moreover, in a 20-year follow-up, patients with type I and type II bipolar disorder were found to spend about half the time in sub-threshold affective conditions, and these were dimensional, involving the full range of symptom severity of depression and hypomania (Judd et al., 2003). Most healthy people report depressive and hypomanic symptoms and many are identifiable as manifesting depressive, hypomanic and cyclothymic temperaments, which appear to predispose to the respective affective disorders and personality disorders (see Fig. 1). Only about 15% of the population report no such symptom over their lifetime and are ‘super-normal’, with very low scores for vegetative lability and neuroticism.

**CASENESS**

The dimensional nature of the mood spectrum raises the question of the correct cut-off levels for caseness (Wing et al., 1978). The current concept of bipolar-II disorder requires a diagnosis of hypomania in addition to major depression. However, the definition of hypomania is the subject of much controversy and research. It is generally agreed that the DSM-IV criteria (American Psychiatric Association, 1994) are too strict (not sensitive) and not based on empirical evidence (not validated). All aspects of the definition are under discussion: the quality of the stem questions (criterion A), which are restricted to elevated and irritable mood (i.e. do not consider symptoms and signs of increased activity), the number and nature of symptoms required and the minimum duration of an episode. A duration of under 4 days (2 days or 1 day) and the presence of two or three instead of three or four of the seven symptoms of hypomania are now proposed and have been partially validated. As with depression, brief spells of hypomania (1–3 days) are far more common than manifestations lasting 4 days or 1 week. In order to improve the recognition of bipolarity, we have proposed a sub-diagnostic concept consisting of a few hypomanic symptoms of brief duration associated with a lifetime diagnosis of depression.

What we need today is an empirically validated, sensitive definition of hypomania, which will allow early recognition of major and minor bipolar disorders. Promising modern screening instruments for the self-assessment of hypomanic symptoms have now been developed such as the Mood Disorder Questionnaire (Hirschfeld et al., 2000) and the Hypomania Checklist–32 (Angst et al., 2003), but there is still no gold standard for valid cut-off points for caseness on these continuous measures: that would depend on a validated definition of hypomania, which is still lacking. A similar problem is present in measures of temperament, where clear distinctions between depressive, hyperthymic and cycloid (‘cyclothymic’) personality disorders are desirable.

**THE FUTURE**

It may take longer than hoped to develop better-validated diagnostic criteria: apart from genetic data, long-term follow-up studies over at least 10 years are needed in order to approach diagnostic classifications that can be used as gold standards. For this reason the forthcoming ICD–11 and DSM–V may again have to give us definitions with a limited half-life.

Too many studies, especially in epidemiology, have used methods tailored and restricted to the current DSM–IV diagnostic concepts and have not collected additional data which would have allowed these concepts to be questioned – and too many journals and reviewers hesitate to accept papers that deviate methodologically from the current diagnostic conservatism. At the other extreme, the promising bipolar spectrum concept can be discredited by uncritical generalisations and over-inclusiveness, for instance by taking for granted that cyclothymic personality or borderline disorders are validated elements of the bipolar spectrum. These hypotheses may be correct, but we need much more genetic and follow-up evidence to support them.

We can safely assume that the prevalence of bipolar disorders is seriously under-reported and that the burden of bipolar disorder, estimated by the World Health Organization to be much lower than that of depression, will as a consequence have to be reassessed.

The mood spectrum is also embedded in the spectrum of functional psychoses, including schizophrenia and schizoaffective and affective disorders. There is growing clinical evidence that the spectrum approach, with its dimensional nature, offers a real alternative to the traditional Kraepelinian dichotomy of schizophrenia vs. manic-depressive insanity (Marneros, 2006) and the unipolar–bipolar dichotomy. Moreover, in agreement with clinical genetic studies (Angst & Scharfetter, 1990), modern molecular genetic studies demonstrate that there is no clear-cut distinction between schizophrenia and bipolar affective disorder; both clinically and genetically they share many features (Craddock & Owen, 2005).

Diagnostic concepts of psychiatric syndromes are purely descriptive and cannot
constitute diseases (Kendell, 1999). The proposed mood spectrum model unifies categorical classification, which is essential, with a dimensional view, which is true to nature; both are needed and both are empirically testable.

REFERENCES


Stigma: ignorance, prejudice or discrimination?

GRAHAM THORNicroft, DIANA ROSE, ALIYA KASSAM and NORMAN SARTORIUS

Summary The term stigma refers to problems of knowledge (ignorance), attitudes (prejudice) and behaviour (discrimination). Most research in this area has been based on attitude surveys, media representations of mental illness and violence, has only focused upon schizophrenia, has excluded direct participation by service users, and has included few intervention studies. However, there is evidence that interventions to improve public knowledge about mental illness can be effective. The main challenge in future is to identify which interventions will produce behaviour change to reduce discrimination against people with mental illness.

Declaration of interest G.T. and A.K. undertake stigma-related research supported by an educational grant from Lundbeck UK Ltd.

Stigma is a mark or sign of disgrace usually eliciting negative attitudes to its bearer. If attached to a person with a mental disorder it can lead to negative discrimination. It is sometimes but not always related to a lack of knowledge about the condition that led to stigmatisation. There is now a voluminous literature on stigma (Link & Phelan, 2001; Corrigan, 2005), but this has largely been limited to attitude surveys rather than studies establishing an evidence base of effective interventions (Sartorius & Schulze, 2005). Stigma can therefore be seen as an overarching term that contains three elements: problems of knowledge (ignorance), problems of attitudes (prejudice), and problems of behaviour (discrimination).

SHORTCOMINGS OF WORK ON STIGMA

Five key features have limited the usefulness of stigma theories. First, although these processes are undoubtedly complex, academic writings on stigma (which in the field of mental health have almost entirely focused upon schizophrenia) have made relatively few connections with legislation concerning disability rights policy (Sayce, 2000) or clinical practice. For example, legislation such as the Americans with Disabilities Act of 1990 in the USA and the Disability Discrimination Act 1995 in the UK are now being applied to cases involving mental illness (23% of all Disability Discrimination Act cases in the UK). Second, most work on mental illness and stigma has been descriptive, overwhelmingly describing attitude surveys or the portrayal of mental illness by the media. Little is known about effective interventions to reduce stigma. Third, there have been notably few direct contributions to this literature by service users (Chamberlin, 2005). Fourth, there has been an underlying pessimism that stigma is deeply historically rooted and difficult to change. This has been one of the reasons for the reluctance to use the results of research in designing and implementing action plans. Fifth, stigma theories have de-emphasised cultural factors and paid little attention to issues related to human rights and social structures.

Recently there have been early signs of a developing focus upon discrimination. This can be seen as the behavioural consequences of stigma which act to the disadvantage of people who are stigmatised (Sayce, 2000). The importance of discriminatory behaviour has been clear for many years in terms of the personal experiences of service users, in terms of devastating effects upon personal relationships, parenting and childcare, education, training, work and housing (Thornicroft, 2006). Indeed, these voices have said that the rejecting behaviour of others may bring greater disadvantage than the primary condition itself.

IGNORANCE: THE PROBLEM OF KNOWLEDGE

At a time when there is an unprecedented volume of information in the public domain, the level of accurate knowledge about mental illnesses (sometimes called ‘mental health literacy’) is meagre (Crisp et al, 2005). In a population survey in England, for example, most people (55%) believed that the statement ‘someone who cannot be held responsible for his or her own actions’ describes a person who is mentally ill (Department of Health, 2003). Most (63%) thought that fewer than 10% of the population would experience a mental illness at some time in their lives. There is evidence that deliberate interventions to improve public knowledge about depression can be successful, and can reduce the effects of stigmatisation. In a campaign in Australia to increase knowledge about depression and its treatment, some states and territories received an intensive, coordinated programme while others did not. In the former, people more often recognised the features of depression, and were more likely to support help-seeking for depression or to accept treatment with counselling and medication (Jorm et al, 2005).

A series of government surveys in England between 1993 and 2003 revealed a mixed picture. On one hand there are some clear improvements: for example, the proportion thinking that people with mental illness can be easily distinguished from ‘normal people’ fell from 30% to 20% (Department of Health, 2003). On the other hand, views became significantly less favourable over this decade for several items: for example, the proportion believing that residents have nothing to fear from people coming into their neighbourhood to obtain mental health services decreased from 70% to 55%. An increase in knowledge about mental illness thus does not necessarily improve either attitudes or behaviour towards people with mental illness.

PREJUDICE: THE PROBLEM OF NEGATIVE ATTITUDES

Although the term ‘prejudice’ is used to refer to many social groups that experience disadvantage, for example minority ethnic groups, it is employed rarely in relation to people with mental illness. The reactions of a host majority to act with prejudice in
rej ecting a minority group usually involve not just negative thoughts but also emotions such as anxiety, anger, resentment, hostility, distaste or disgust. In fact, prejudice may more strongly predict discrimination than do stereotypes.

Interestingly, there is almost nothing published about emotional reactions to people with mental illness apart from that describing a fear of violence. One fascinating exception to this is work carried out in south-eastern USA, in which students were asked to imagine meeting people who either did or did not have a diagnosis of schizophrenia. All three physiological measures of stress (brow muscle tension, palm skin conductance and heart rate) were raised during imaginary meetings with ‘labelled’ compared with ‘non-labelled’ individuals. Such tension also associated with self-reported negative attitudes of stigma towards people with schizophrenia. The authors concluded that one reason why individuals avoid people with mental illness is physiological arousal, which is experienced as unpleasant feelings (Graves et al, 2005).

DISCRIMINATION: THE PROBLEM OF REJECTING AND AVOIDANT BEHAVIOUR

Attitude and social distance surveys usually ask either students or members of the general public what they would do in imaginary situations or what they think ‘most people’ would do, for example, when faced with a neighbour or work colleague with mental illness. Important lessons have flowed from these findings. This work has emphasised what ‘normal’ people say without exploring the actual experiences of people with mental illness themselves about the behaviour of normal people toward them. Further, it has been assumed that such statements (usually on knowledge, attitudes or behavioural intentions) are congruent with actual behaviour, without assessing such behaviour directly. Such research has generally focused on hypothetical rather than real situations, neglecting emotions and the social context, thus producing very little guidance about interventions that could reduce social rejection. In short, most work on stigma has been beside the point.

CONSEQUENCES FOR ACTION

Experience and evidence gained so far indicates that the time has come to shift the focus of research and action from stigma to discrimination, Thus, instead of asking an employer whether he or she would hire a person with mental illness, we should assess whether he or she actually does. This would allow an evaluation of our interventions by measuring whether and how they change behaviour towards people with mental illness, without necessarily assessing changes of knowledge or feelings. Finally – and most importantly – such a shift of focus would make it possible for people with mental illness to expect to benefit from relevant anti-discrimination policies and laws in their country or jurisdiction, on a basis of parity with people with physical disabilities (Thornicroft, 2006). In sum, this means sharpening our focus upon human rights, upon injustice and discrimination as actually experienced by people with mental illness, and upon adding to our knowledge about interventions that society should undertake to reduce both stigmatisation and its consequences.

REFERENCES


Schizophrenia: a common disease caused by multiple rare alleles†

JON M. McCLELLAN, EZRA SUSSER and MARY-CLAIRE KING

Summary  Schizophrenia is widely held to stem from the combined effects of multiple common polymorphisms, each with a small impact on disease risk. We suggest an alternative view: that schizophrenia is highly heterogeneous genetically and that many predisposing mutations are highly penetrant and individually rare, even specific to single cases or families. This ‘common disease – rare alleles’ hypothesis is supported by recent findings in human genomics and by allelic and locus heterogeneity for other complex traits. We review the implications of this model for gene discovery research in schizophrenia.

Declaration of interest  None.

Funding detailed in Acknowledgements.

Current research in the genetics of schizophrenia is guided primarily by the ‘common disease – common alleles’ model (Chakravarti, 1999). This model originated from the hypothesis that the illness results from the cumulative impact of multiple common small-effect, genetic variants, interacting with environmental exposures to exceed a biological threshold (Gottesman & Shields, 1982). The ‘common disease – common alleles’ model for schizophrenia is heuristically appealing. The illness is relatively frequent and is found worldwide. Thus common susceptibility alleles shared across populations are plausible. The ‘common disease – common alleles’ model has also been posited to explain the variable and inconsistent results of linkage studies devoted to finding genes of large effect responsible for schizophrenia, and the weak associations of various candidate genes with schizophrenia. Furthermore, mathematical modelling suggested that the observed decline in recurrence risk of disease with increased genetic distance from affected individuals is inconsistent with monogenic inheritance of large-effect alleles (Risch, 1990). As pointed out by Risch, these models were based on the assumption that the illness was genetically homogeneous throughout the population for which the recurrence risks were calculated (Risch, 1990). Increasing evidence suggests that schizophrenia is genetically heterogeneous (Fanous & Kendler, 2005). If so, then recurrence risk data are also consistent with monogenic inheritance of large-effect alleles in a proportion of people with schizophrenia, with different alleles for different families.

We suggest that the ‘common disease – rare alleles’ model explains many cases of schizophrenia. Our hypothesis is that many mutations predisposing to schizophrenia are highly penetrant and individually rare, even specific to single patients or families. In this model, different families harbour different mutations, either in the same gene or in different genes, but any one family carries only one or two mutations. Many different disease-associated mutations may occur in the same gene.

The ‘common disease – common alleles’ and ‘common disease – rare alleles’ models are not mutually exclusive (Goldstein & Chikhi, 2002). Rare severe mutations may occur in genes that also harbour more common variants with modest effects on disease risk. However, the two models have distinctly different implications for gene-finding strategies. Most current psychiatric genetic research is designed to identify common alleles or haplotypes associated with increased risk of disease and shared by large numbers of patients compared with appropriate controls (Merikangas & Risch, 2003). If many cases of schizophrenia stem from individually rare large-effect alleles, current approaches – even if executed perfectly – will fail to identify critical genes.

We argue that current observations from epidemiology and genetics of schizophrenia are consistent with the influence of a large number of individually rare deleterious mutations, many of which have occurred in the present or recent generations. Several features of schizophrenia support this view:

(a) Schizophrenia is familial, i.e. close relatives of affected persons are at increased risk of the illness. Occasional families are very severely affected (Gottesman & Shields, 1982), but most patients have no close affected relative. Taken together, these observations are consistent with a subset of families harbouring high-penetrance, recently occurring alleles predisposing to schizophrenia, with different alleles present in different families. Unless caused by detectable chromosomal alterations (in the case of DISC1), such alleles have therefore been difficult to find because individual families are not sufficiently informative for single-family linkage analyses.

(b) Paternal age is consistently associated with increased risk of schizophrenia (Brown et al, 2002; Dalman & Allebeck, 2002; Malaspina et al, 2002; Byrne et al, 2003; El Saad et al, 2004; Sipos et al, 2004; Tsuchiya et al, 2005). Paternal age is also associated with increased rates of several types of de novo germ-line mutations (Crow, 2003).

(c) The illness is associated with decreased fertility (Nimgaonkar, 1998; Haukka et al, 2003). If this had been the case over long periods, then the frequencies of any ancient common alleles associated with schizophrenia would be reduced. An ongoing contribution of new and therefore individually rare risk alleles could explain the persistence of the disorder.

To explain our reasoning, we first describe recent findings in genomics and genetics of other complex human disorders. Then we show that the results of schizophrenia research are consistent with the existence of multiple individually rare alleles of large effect. Finally, we consider the implications of this model for future schizophrenia research.

The human genome

The dynamics of the human genome are proving more complex than anticipated, revealing more mechanisms by which
genetic changes may lead to human disease. Particularly relevant to our argument are the following observations (International Human Genome Sequencing Consortium, 2001a,b, 2004).

(a) Only about 2% of the genome consists of protein-coding genes. There are approximately 20,000 protein-coding genes, far fewer than the 80,000–100,000 hypothesised a decade ago. However, the human proteome is enormously complex. At most genes, variable transcription leads to multiple transcripts, and thus multiple proteins, derived from the same locus but with different amino acid sequences. Variable transcription is frequently tissue-specific. As a result, the consequences of a mutation may also be tissue-specific.

(b) Germ-line mutations occur more commonly than previously thought. Potentially deleterious new mutations may occur at a rate as high as three per zygote (Eyre-Walker & Keightley, 1999; Crow, 2000). Rates of occurrence of different classes of de novo mutations are differently influenced by parent of origin and parental age (Crow, 2000, 2003). The increased mutation rate associated with greater paternal age is particularly relevant, given that risk of schizophrenia is also associated with paternal age.

(c) Epigenetic alterations – stable changes in gene expression that do not depend on changes in DNA sequence (Jaenisch, 1999; Crow, 2000). Observations of possible epigenetic effects related to development include phenotypic variability in monozygotic twins and imprinting effects on neurodevelopmental disorders (Fraga et al., 2005; Wong et al., 2005).

GENETIC HETEROGENEITY IN HUMAN DISEASE

The genetic heterogeneity of complex illnesses is the natural result of the origins of human genetic variation. The oldest human alleles originated in Africa millions of years before people first migrated out of Africa some 50,000 years ago (Cavalli-Sforza et al., 1994). These ancient variants are found in all human populations, are the most common worldwide, and account for approximately 95% of human variation. Yet the exponential growth of the human population has resulted in many new alleles, each individually rare and each specific to one population (or even one family). Most alleles are of this sort. Thus the paradox: most human variation is ancient and shared; most alleles are recent and individually rare. Given the size of the present human population and the rate of occurrence of new mutations, all mutations compatible with life have probably already occurred and will occur again. However, mutations with deleterious effects before or during the reproductive years will be less frequently transmitted to subsequent generations owing to their adverse impact on fertility or viability. Therefore, mutations with large effects on disease may be disproportionately of recent origin and individually very rare.

To the extent that any class of mutation – point mutations, copy number errors or abnormalities of chromosome number – occur spontaneously, they appear at similar rates in all human populations. All humans share the same basic genomic architecture, including the same genomic regions vulnerable to mutations. The incidence of schizophrenia does not appear to vary substantially across populations by virtue of genetic ancestry. This pattern is consistent with a disease due to multiple, independent de novo mutations that arise in many different vulnerable genes and genomic regions. Furthermore, environmental exposures with mutagenic consequences may lead to high rates of new mutations among exposed individuals. For example, environmental factors such as maternal starvation that are associated with disease (Cannon et al., 2003) may mediate their effects through de novo genetic or epigenetic mutations. We explore this theme in more detail below.

Complex illnesses are almost universally characterised by allelic heterogeneity (multiple different mutations in the same gene leading to disease) and locus heterogeneity (mutations in multiple different genes leading to the same disease) (Botstein & Risch, 2003; Goldstein et al., 2003). We propose that both are characteristic of schizophrenia. To understand the potential implications of genetic heterogeneity for schizophrenia, we briefly consider other complex disorders for which genes have been identified.

Deafness

To date, nearly a hundred genes have been identified that harbour inherited mutations leading to hearing loss (Petit et al., 2001; Friedman & Griffith, 2003). All mutations are recent and all but one are individually rare. The one frequent mutation, 30delG in connexin 26, is the exception that proves the rule, in that the same mutation has occurred independently numerous times in a mutational hot-spot.

Epilepsy

The inherited forms of epilepsy are characterised by allelic and locus heterogeneity (Meisler et al., 2001). Mutations in any of several genes involved with neuronal signalling can lead to broadly defined epilepsy. Rare mutations in three different sodium channel genes lead to one more narrowly defined form of epilepsy (generalised epilepsy with febrile seizures plus).

Alzheimer’s disease

Alzheimer’s disease illustrates the ‘common disease – common allele’ and ‘common disease – rare allele’ models need not be mutually exclusive. The common e4 allele of APOE (apolipoprotein E) is associated with a threefold to fourfold increased risk in individuals of European descent of developing common, late-onset Alzheimer’s disease (Bird, 2005). On the other hand, multiple rare mutations in genes encoding amyloid precursor protein (APP), presenilin 1 and 2 (PS1 and PS2) and ubiquitin 1 (UBQLN1) are responsible for familial early-onset Alzheimer’s disease. Therefore, both common modest-effect alleles and rare large-effect alleles have a role in Alzheimer’s disease. The role of APOE4 is an excellent example of the ‘common disease – common allele’ model. However, the effect of APOE4 on Alzheimer’s disease risk is substantially larger than the effect sizes of 2 or less that are typically estimated for schizophrenia susceptibility genes.

Inherited predisposition to cancer

In each of the two major genes for inherited breast and ovarian cancer, BRCA1 and BRCA2, more than a thousand different pathogenic mutations have been found (Walsh et al., 2006). Large genomic rearrangements account for about 10% of these mutations. All inherited BRCA1 and BRCA2 mutations are individually rare. Both locus and allelic heterogeneity are also characteristic of inherited colon cancer and the rarer cancer syndromes (Vogelstein & Kinzler, 2004).
Lipid metabolic pathways

Rare variants in genes related to lipid metabolism are associated with low levels of high-density lipoprotein cholesterol (HDL–C) (Cohen et al, 2004) and low-density lipoprotein cholesterol (LDL–C) (Cohen et al, 2006). Although each variant is individually rare, in total these variants are found in a substantial portion of individuals at the far end of the spectrum in terms of levels of HDL–C or LDL–C respectively.

These examples illustrate two ways in which mutations of large effect are important for understanding human disease. First, the collective effect of individually rare mutations in the same gene may explain a considerable proportion of an illness. Second, rare mutations in genes of large effect can reveal pathways critical to disease development.

IMPLICATIONS OF GENETIC HETEROGENEITY FOR GENE DISCOVERY

All complex illnesses evaluated thus far are characterised by locus and allelic heterogeneity. Disease genes for these illnesses have been identified primarily by positional cloning in large kindreds. Although subsequent association studies confirmed their role, the original gene discoveries were dependent upon individual highly informative families. Such large informative kindreds are extremely rare in schizophrenia. Linkage studies of schizophrenia based on single gene models have not been successful at identifying causal mutations (Owen et al, 2004). Because pedigrees with schizophrenia have not been large enough to be individually informative, studies generally pool data from different families. If many different genes were responsible for the illness in different families, pooling results across families would preclude identification of any of them.

Currently, most gene-discovery strategies for schizophrenia research – case–parent triad studies, candidate gene studies and haplotype association studies – are designed to identify alleles or haplotypes that appear more frequently among affected individuals than among appropriate controls (Cannon et al, 2003). These designs are not robust to either allelic heterogeneity or locus heterogeneity. Sib-pair linkage analyses are designed to detect genomic regions consistently shared by affected siblings. Sib-pair analyses are robust to allelic heterogeneity but not to locus heterogeneity. For each of these designs, hundreds or thousands of rigorously diagnosed cases of unrelated affected and unaffected individuals are examined. If a substantial portion of schizophrenia stems from different individually rare alleles, increasing the number of cases also increases the number of different disease risk mutations represented among them. As a result, increasing sample size does not confer a corresponding increase in statistical power. In the most extreme scenario, in which every case results from a different mutation, an increase in sample size would not lead to any increase in the power to detect any one mutation. Consequently, even very large studies may fail to detect individually rare disease risk mutations.

In addition, genetic analyses that focus only on single nucleotide polymorphisms (SNPs), either individually or in haplotypes, rather than fully sequenced DNA, will inevitably miss rare disease alleles and thus fail to detect critical genes harbouring such alleles. Association and linkage studies generally assume that individuals sharing the same SNP-defined haplotype share the entire region, including any hypothetical embedded disease alleles. This assumption is reasonable for ancient alleles and nearly always true for related individuals for whom there is direct inheritance of the haplotype. However, this assumption is not reasonable for a study of unrelated individuals who carry disease alleles of recent origin. Rare recent mutations causing schizophrenia within the same haplotype will differ among unrelated individuals, diluting any association.

TWO CANDIDATE GENES FOR SCHIZOPHRENIA: DYSBINDIN AND DISCI

The familial nature of schizophrenia is well established. Large collaborative linkage studies have suggested multiple candidate chromosomal regions that may harbour genes associated with the illness. Regions best supported by genome-wide scans include 6p22–p24 (Straub et al, 1995), 1q21–q22 (Brustowicz et al, 2000) and 13q32–q34 (Blouin et al, 1998). Other regions with positive linkage findings include 1q42, 5q21–q33, 6q21–q25, 8p21–p22, 10p15–p11 and 22q11–q12 (Owen et al, 2004). These regions combined represent a substantial portion of the genome.

Candidate genes have been suggested in several of these regions, including dysbindin on 6p22, neuregulin on 8p22, G72 on 13q34, COMT on 22q11, RGS4 on 1q21 and GRM3 on 7q21 (see reviews by Blouin et al, 1998; Harrison & Weinberger, 2005). Each of these genes is biologically plausible (Owen et al, 2004). However, for each candidate gene, both positive and negative associations have been reported with the same SNPs; strengths of effects are generally weak; the specific allele or haplotype associated with the illness varies across studies; and definitive causative mutations have not been identified.

To illustrate the implications of the common allele v. rare allele models, we will review two promising susceptibility genes, dysbindin (DTNBP1) and DISCI. Very different study designs revealed these genes, with correspondingly different results to date. Linkage, association and functional studies all support some role for dysbindin in schizophrenia. Several studies involving different populations have found positive associations of dysbindin alleles or haplotypes with schizophrenia (Straub et al, 2002; Schwab et al, 2003; Funke et al, 2004; Kirov et al, 2004; Kohn et al, 2004; Numakawa et al, 2004; Williams et al, 2004; Bray et al, 2005; Gornick et al, 2005). Dysbindin is widely expressed in brain, and appears to play a part in cognitive functioning and capacity (Owen et al, 2004). Post-mortem studies suggest that brain levels of dysbindin may be reduced in individuals with schizophrenia (Weickert et al, 2004). However, no variant of dysbindin has been specifically linked to schizophrenia. Across different studies, the risk conferred by any dysbindin variant is small, with effect sizes of about 2.0 or less. Among positive association studies, the specific alleles associated with the disease differ. Moreover, an allele may be associated with increased disease risk in some studies and decreased risk in others (Owen et al, 2004). In general, the variants of interest (defined by SNPs) are common and without known functional significance. An exception is SNP rs1047631, which has been associated with differences in the expression of dysbindin in brain (Funke et al, 2004). However, the frequency of the haplotype with this SNP was similar between cases (45.6%) and controls (40.4%). Thus far, resequencing efforts have not revealed any coding sequence mutations in dysbindin among individuals with schizophrenia (Liao & Chen, 2004).
Therefore, at present the evidence supporting dysbindin is mixed. Variable associations with different alleles have been attributed to allelic heterogeneity; yet allelic heterogeneity refers to different disease-causing mutations in the same gene, not to the same allele reducing risk in some cases and increasing risk in others. There are at least three possible interpretations of these data. The most favoured in the literature is that dysbindin variants mediate disease risk as part of a complex interaction with other genes and environmental factors. This is possible, in principle, although difficult to test given the challenge of establishing the role of a mediating factor of small effect on a complex illness of unknown cause.

A second possibility is that relatively rare, as yet unidentified variants in the dysbindin locus are embodied in illness-associated haplotypes in some, but not all, cases (including those potentially located in non-coding regulatory regions). Such alleles could have substantial effects on the phenotype, but would be masked by studying only the common haplotype. The third possibility is that many, most, or all of the various positive associations with dysbindin are false positives. The number of different positive association studies (albeit with different variants) is taken as prima facie evidence that the gene must be involved with the disorder. However, dysbindin, like many genes involved with brain development, is large (> 140 kb). The dysbindin locus includes at least 363 SNPs (Hinrichs et al., 2006), from which various candidates are selected for association studies. Incorporating linkage disequilibrium across the locus, many thousands of SNP and haplotype combinations appear in different populations. The potential for false positives is enormous. Unless negative and positive studies were published with equal frequency, this possibility is also nearly impossible to test.

In contrast, DISC1 and its associated non-coding antisense RNA DISC2 were originally identified by a balanced translocation involving chromosome 1q42 which segregated with schizophrenia (and other major psychiatric disorders) over four generations in a large Scottish kindred (St Clair et al., 1990; Millar et al., 1998, 2000). Sachs et al. (2005) found that a frameshift mutation that abnormally truncates DISC1 co-segregated with schizophrenia in three siblings. However, this mutation is also rarely found in controls with unknown diagnostic status (Green et al., 2006). The gene PDEAB (phosphodiesterase 4B), which interacts with DISC1 in the neuronal cyclic adenosine monophosphate (AMP) pathway, was disrupted by a balanced translocation in two related individuals with chronic psychotic illnesses (Millar et al., 2005). Finally, mice with a deletion variant that disrupts the DISC1 protein have working memory deficits (Koike et al., 2006). These findings suggest a role of rare large-effect mutations of DISC1 and of genes involved in DISC1 pathways in the development of schizophrenia.

Not surprisingly, DISC1 became the subject of association studies. Haplotypes of DISC1 were associated with schizophrenia and other mental illness in European and North American populations (Ekelund et al., 2001; Hennah et al., 2003; Hodgkinson et al., 2004; Callicott et al., 2005; Cannon et al., 2003; Hamshere et al., 2005) but not in Japanese or Scottish populations (Devon et al., 2001; Kockelkorn et al., 2004; Zhang et al., 2005). Within populations with positive associations, DISC1 haplotypes were also associated with putative endophenotypes, including neuroanatomical and/or neurocognitive profiles (Hodgkinson et al., 2004; Burdick et al., 2005; Cannon et al., 2005). However, disease-risk haplotypes vary across populations and effect sizes are small. Therefore, although there is compelling evidence that rare large-effect mutations in DISC1 are associated with schizophrenia, it is not clear whether common polymorphisms play a part.

IMPlications for genetIc research In schizophrenia

The ‘common disease – rare allele’ model has important implications for gene-finding strategies. A current mantra in schizophrenia genetic research is the need for ever-larger sample sizes in order to obtain adequate statistical power to detect common small-effect variants (Devon et al., 2001). These designs are dependent upon the existence of disease-risk alleles that are shared across large numbers of related cases. These strategies will be inadequate if schizophrenia in large part stems from individually rare disease-risk mutations in a large number of different genes.

We propose an alternative strategy in selecting cases for study. Rare cases of schizophrenia with mutations that can be individually detected using current genomic technologies are extremely valuable. It is worthwhile devoting resources to finding them. It has been recognised for decades that such cases would include any large kindred with a number of well-diagnosed individuals or cases with identifiable genomic events of recent origin (e.g. balanced translocations). The emerging story of DISC1 highlights this strategy, since the gene was originally identified by a balanced translocation on chromosome 1q42 co-inherited with schizophrenia (Millar et al., 1998). Similar promising findings have been noted for specific genes in the 22q11 region (Maynard et al., 2002), stemming in part from the recognised association between deletions at 22q11 (i.e. the velo-cardio-facial syndrome) and schizophrenia.

Current genomic technology now enables the identification of an increasingly large number of classes of mutations. Here-tofore, identifiable genomic events have been limited to chromosomal abnormalities such as translocations or deletions. However, as the resolution of genome-wide mutation screening technologies improves, smaller genomic events in informative cases or families can be detected. Once identified, a gene altered by a single chromosomal event becomes a candidate to be screened for other (typically smaller) mutations in other cases. Advanced resequencing technology allows for more rapid identification of mutations in candidate genes. The occurrence of multiple deleterious mutations among unrelated cases provides both biological evidence and epidemiological support for the causal role of the gene, using gene-based hypothesis testing strategies (Chen et al., 2006).

Individuals who develop schizophrenia following a known environmental exposure are also potentially informative. Such exposures may focus gene discovery in two ways. First, genomic approaches (e.g. resequencing efforts) can focus on candidate genetic pathways relevant to the suspected exposure, screening for otherwise benign variants that are deleterious given the exposure. For example, associations between schizophrenia and in utero exposure to maternal starvation (Susser et al., 1996; St Clair et al., 2005), and associations between schizophrenia and genes in the folate metabolic pathway (Lewis et al., 2005; Picker & Coyle, 2005) suggest that mutations in genes in the folate metabolism network could be linked to the illness. Second, the mutagenic effects of the environmental exposure can be evaluated. Following the same example, gestational folate deficiency may be mutagenic, in that
it leads to an increase in the rate of mutation genome-wide (McClellan et al., 2006). Among such cases, genome-wide mutational screening may detect de novo mutations. If disease-associated alleles are identified, other mutations within the same gene may confer some risk for other individuals with the same exposure. Furthermore, severe mutations in the same gene may lead to the disorder without the exposure.

As informative cases are identified, genomic technologies are needed for efficient screening for potential disease-associated mutations. Effective transcriptome- and proteome-based tools are needed to characterise such variants. These technologies are under rapid development. It is already possible to detect deletions, duplications and other chromosomal aberrations of multiple kilobases anywhere in the genome, and the sensitivity of these methods for detecting smaller mutations is improving (Sebat et al., 2004; Sharpe et al., 2005). Resequencing tools are increasingly efficient, so large numbers can be screened for rare events in candidate genes first identified by rare, individually detectable events. User-friendly bioinformatics resources now exist to help characterise the structure and function of potential candidate genes. It is increasingly possible to characterise not only mutations in protein coding sequence but also mutations and epigenetic changes in regulatory regions, in non-coding RNA and in transposable elements.

To summarise, we propose that individually rare alleles with large effect, many of which are recent in origin, have a substantial role in causing schizophrenia. Current research designs that focus on collecting large samples of unrelated individuals for analysis of shared alleles or haplotypes are not suitable for detecting such disease alleles. In contrast, rare disease mutations may be revealed by studies of individuals and families that harbour informative genomic events, and by studies of exposed cohorts. A gene harbouring one mutation predisposing to schizophrenia is likely to harbour more than one, with frequencies ranging from relatively common to rare, and effects ranging from modest to severe. To see one is not to see them all.

ACKNOWLEDGEMENTS

Support for this work was provided by grants from the National Institutes of Health (PH010120) and the Stanley Medical Research Foundation to JMM., the Lieber Center for Schizophrenia Research to E.S. and a National Association for Research on Schizophrenia and Affective Disorders Distinguished Investigator Award to M.-C.K.

REFERENCES

Burdick, K. E., Hodgkinson, K. A., Szczesniak, P. R., et al. (2005) DISC1 and neurocognitive function in schizophrenia. Neuropathol. 13, 1399–1402.
Association study of polymorphisms in the 5′ upstream region of human DISC1 gene with schizophrenia and bipolar disorder. American Journal of Human Genetics, 75, 862–872.


Phenotypic and genetic complexity of psychosis
Invited commentary on . . . Schizophrenia: a common disease caused by multiple rare alleles†

NICK CRADDOCK, MICHAEL C. O’DONOVAN and MICHAEL J. OWEN

Summary  Psychosis, like other major psychiatric disorders, is both genetically and clinically complex. Increasingly powerful molecular genetic studies have the potential to identify DNA variation that influences susceptibility to genetically complex disorders. There is a need to use a range of genetic approaches appropriate to identifying a spectrum of risk variants from the common through to the rare. Some variants might have large effects at the level of the individual but most are likely to have modest or small effects at both population and individual level. Extensive clinical heterogeneity is likely to have a significant impact on the power of even the largest studies and, more importantly, will lead to extensive variability between studies and hamper attempts at replication. If we are to realise the potential of molecular genetics, we need to overcome the major limitations imposed by current psychiatric diagnostic classifications and identify clinical phenotypes that reflect the presence of underlying entities with biological validity.

Declaration of interest  None.

In this month’s Journal McClellan et al contrast two models of the genetic architecture of schizophrenia. Here we provide a context for that paper by considering more widely the genetic and phenotypic complexity of psychosis and how this has an impact on genetic research.

COMPLEX GENETIC DISEASES

The term ‘complex genetic diseases’ refers to common familial illnesses that do not show a simple Mendelian pattern of inheritance (Lander & Schork, 1994). Examples include coronary heart disease, hypertension, rheumatoid arthritis, type I and type II diabetes mellitus, asthma, many cancers and most psychiatric disorders. In terms of their genetic properties and complexity, psychiatric disorders, including schizophrenia and bipolar disorder, are very similar to the non-psychiatric common familial disorders. In fact, perhaps surprisingly, genetic susceptibility to risk is substantially higher for the major psychiatric illnesses than for most of the non-psychiatric diseases (Plomin et al, 1994). What makes the study of psychiatric genetics substantially more difficult than investigation of the complex non-psychiatric diseases is the lack of biologically valid measures for phenotype definition.

RARE VARIANTS OF LARGE EFFECT AND COMMON VARIANTS OF SMALL EFFECT

Theory
McClellan et al consider two distinct genetic models that can explain the transmission of common familial disorders that show non-Mendelian inheritance. They conclude that most cases of schizophrenia are likely to be explained by genetic variants of large effect that, although individually rare, in their totality account for the majority of cases in the population. According to this view, only one rare variant of large effect is involved in each family, but different variants, which may be in the same or in other genes, operate in other families. This is sometimes called the ‘common disease–rare variant’ model. It can be contrasted with the ‘common disease–common variant’ model which forms the rationale for the large-scale genetic association studies that are ongoing in many centres around the world. In the latter model, a common disease, such as schizophrenia, results from the co-action of multiple (ranging in principle from a few to many thousand) common variants (‘polymorphisms’), each of which has a small effect on illness susceptibility. When an individual inherits several, or many, susceptibility variants together, they have a sizable influence on disease risk. This is essentially the traditional ‘multifactorial’ model that assumes the action of multiple genes and environmental risk factors (Falconer & MacKay, 1995).

Schizophrenia
McClellan et al argue strongly against the common disease–common variant model but argue in favour of rare variants of large effect. Since we still do not know the true genetic architecture of schizophrenia, challenges to widely held assumptions and discussions of the possible impact on gene discovery are welcome. However, a considerable body of genetic epidemiological and molecular data relating to schizophrenia as well as population genetic findings do allow some inferences to be drawn and constrain the nature of plausible models. We agree with McClellan et al that rare mutations are likely to be important in some cases of schizophrenia; there are indeed examples in which schizophrenia is related to chromosomal abnormalities. However, it is our contention that key genetic epidemiological and molecular genetic observations are inconsistent with the hypothesis that rare variants of large effect can explain the majority of cases of schizophrenia. Further, the dismissal by McClellan and colleagues of the importance of variants of modest or small effect is not well founded. Important pieces of evidence that contradict their assertions are given below.

Families with clear Mendelian inheritance patterns are rare
Under the model of rare variants of major effect, even allowing for a high proportion of new mutations, it would be expected that there would be many families with clear-cut single gene inheritance. However, as experienced clinical psychiatrists will know, such families are rare.

Single genes of major effect have not been found
Over the past 20 years hundreds of diseases with Mendelian inheritance have been subjected to genetic analysis (‘positional cloning’) which has allowed detection of mutations of major effect using only a few
families or just one large pedigree (Collins, 1992; Botstein & Risch, 2003), some of which are cited by McClelland et al. Although rare, extended pedigrees multiply affected by psychosis do exist and have been studied genetically. If single genes of major effect explained illness in such pedigrees, the genetic methods used should have identified them, or at least unambiguously defined a chromosomal location, as has been done successfully for the many disorders with Mendelian inheritance. However, when extended pedigrees with multiple cases of illness that are consistent with simple Mendelian inheritance patterns have been subjected to intensive molecular genetic study, not only have mutations of major effect not been identified, the evidence for linkage is much weaker than one would expect if the families were segregating a single cause of the disorder. Rather, findings to date are consistent with multiple variants of modest effect (see Chumakov et al., 2002; Stefansson et al., 2002; Straub et al., 2002).

**Mathematical modelling of familial risk is inconsistent with single genes of large effect**

McClelland et al cite Risch's studies modelling the way risk of illness changes as a function of genetic relatedness to a sufferer (Risch, 1990). For both schizophrenia and bipolar disorder there is a very rapid, non-linear decrease of risk when moving from a genetically identical individual (i.e. monozygotic co-twin where the risk is 50–60%), to an individual who shares half the genes (e.g. sibling, parent, dizygotic co-twin where risk is around 10%). Contrary to the assertion by McClelland et al., mathematical modelling demonstrates that this pattern cannot be explained by a collection of genes of large effect that act on their own, even if a sizable proportion are de novo mutations. For illnesses where one mutation is a sufficient cause of illness in each family (whether or not there is a different mutation or gene involved in different families) there is a more gradual (linear) decrease of risk (McCue & Gottesman, 1989). In contrast, the rapid, non-linear decrease of risk is compatible with multiple interacting risk factors, albeit of unknown frequency, that individually have modest effects (Risch, 1990; Craddock et al., 1995).

**Molecular genetic findings are consistent with multiple risk alleles of modest effect**

Several genes have been implicated repeatedly as conferring risk for schizophrenia or bipolar disorder. These include dysbindin (DTNBP1) (Straub et al., 2002; Williams et al., 2005), neuregulin 1 (NRG1; Stefansson et al., 2002; Tosato et al., 2005; Munafò et al., 2006) and D-amino acid oxidase activator (DAOA, G72/G30; Chumakov et al., 2002; Detera-Wadleigh & McMahon, 2006). The patterns of effect sizes and allele frequencies are consistent with the common disease–common variant model and with the positive findings that have been emerging in studies of non-psychiatric complex genetic diseases (Todd, 2006). Estimated effect sizes are all modest, with estimated relative risks (or odds ratios) typically below 2.0. In contrast, no rare alleles of large effect have yet been unequivocally identified, although a few rare chromosomal aberrations have been shown to dramatically increase risk (Craddock et al., 2005).

**GENETIC COMPLEXITY OF SCHIZOPHRENIA**

As is often the case with dichotomous decisions, choosing between either rare variants of large effect or common variants of small effect is almost certainly oversimplistic. Instead, it is probably more reasonable to assume that the spectrum of mutations for common disease is similar to that of all variants in the human genome. This leads to the expectation of a spectrum of risk variants of varying effect sizes, including both common and rare alleles (Wang et al., 2005). Note that the spectrum of likely risk alleles also includes rare variants of small or modest effect size, the existence of which might well prove to be a far greater obstacle to gene discovery than rare alleles of large effect size.

It is important to acknowledge that, in addition to the models already discussed, several molecular genetic mechanisms are known that result in complex, non-Mendelian patterns of inheritance for a disorder or trait. Examples include: dynamic mutations (e.g. the expanding trinucleotide repeats that underlie fragile X disorder); genomic imprinting (e.g. Prader–Willi syndrome); and mitochondrial inheritance (e.g. some optic atrophies) as well as other mechanisms involving deletion, insertion or variable repetition of stretches of DNA. Such mechanisms might contribute to the genetic complexity of psychiatric illness and need to be considered in the search for genetic factors that influence susceptibility to schizophrenia (see Margolis et al., 1999; Malaspina, 2001; Singh et al., 2002; Ben-Shachar & Laifenfeld, 2004).

---

**PHENOTYPIC COMPLEXITY IN PSYCHIATRIC ILLNESS**

For most complex genetic diseases, although pathophysiology is incompletely understood, there are biological measures that can be used to define the phenotype (e.g. blood glucose in diabetes, blood pressure in hypertension or biopsy and histology for cancers). These measures are typically reliable and, importantly, have biological validity. The fundamental importance of phenotype definition and measurement for the success of gene identification in human genetics has long been appreciated (Lander & Schork, 1994). It is surprising, therefore, that the vast majority of psychiatric genetic studies continue to rely on DSM–IV (or ICD–10) diagnostic categories as if they were proven, valid disease entities. Many researchers have assumed that the effects of heterogeneity would be overcome as the technical advances of molecular genetics allowed increasingly large and powerful studies. However, this might not occur if, as well might be the case, researchers tend to adopt less restrictive, or more ‘pragmatic’, inclusion criteria to facilitate the assembly of larger samples. Genetic research might benefit more from the use of smaller, more homogeneous than larger, more heterogeneous samples (Craddock et al., 2006). This principle is clearly demonstrated by the example of the D-amino acid oxidase activator gene (DAOA) which has been implicated in some, but not all, studies of both schizophrenia and bipolar disorder (Chumakov et al., 2002; Hattori et al., 2003; Detera-Wadleigh & McMahon, 2006). In a large study we found evidence that the gene confers risk for episodes of pathologic mood disturbance irrespective of diagnostic category (Williams et al., 2006). We found significant association in 706 individuals meeting DSM–IV criteria for bipolar disorder. Among those meeting criteria for schizophrenia there was significant association in 112 who had also experienced major mood episodes but not in 597 without major mood episodes. Further, no association was detectable if the schizophrenia sample was treated as a single homogeneous entity (as is the case in most studies). This suggests that results of studies of DAOA in categorically defined ‘schizophrenia’ samples are dependent upon the proportion of people in the sample that have experienced mood disorder – information that is not usually provided or considered by researchers.
A further striking example is provided by the gene Disrupted in Schizophrenia 1 (DISC1). This was identified by studies of an extended Scottish pedigree in which a spectrum of psychiatric illness, including mood and psychotic diagnoses, co-segregated with a chromosomal translocation (St Clair et al, 1990; Millar et al, 2000; Muir et al, 2006). Although the name given to the gene by the research team explicitly refers to schizophrenia, major disorder is actually more strongly linked to the translocation (Blackwood et al, 2001). We have provided independent evidence from a linkage study of families with schizoaffactive disorder that variation at the DISC1 locus influences susceptibility to psychopathology involving disturbances in both mood and psychotic domains (Hamshere et al, 2005).

All the molecular evidence suggests that genetic susceptibility does not respect current operational diagnostic boundaries (Craddock & Owen, 2005). This will not surprise psychiatrists. ‘Schizophrenia’ is so broad that it is possible for one sample to be composed of individuals with chronic disability involving cognitive impairment, marked negative features and minimal affective or positive psychotic symptoms whereas another sample could include individuals who are able to function relatively well, with an episodic course and marked affective and positive psychotic symptoms. Self-evidently, unless clinical variation is the consequence of chance or of environmental risk factors, illness in each of the above two samples will reflect the operation of at least some susceptibility alleles not held in common at the same frequency in each group. Since the key to unambiguously identifying a risk factor is replication across different samples, we must move beyond diagnostic categories for describing and analysing samples and routinely consider more detailed measures of lifetime psychopathology. There are substantial theoretical benefits of using endophenotypes (intermediate phenotypes) such as neuroimaging or tests of cognitive function to define more homogeneous groups or to access more directly abnormalities that mediate the effects of genes on psychopathology (Jablensky, 2006; Braff et al, 2007). However, these approaches are not without difficulties (Owen et al, 2005) and most samples collected to date for genetic studies have clinical data rather than these extended measures. It is therefore relatively simple and inexpensive to make more effective use of the clinical data that can be used to characterise individuals.

### ISSUES THAT AFFECT PHENOTYPIC COMPLEXITY

Much attention has been devoted to genetic issues that can affect comparisons between studies. These include genetic differences between different geographical or ethnic populations (so-called population stratification or structure; see Cardon & Palmer, 2003), methods of ascertainment (see McCarthy et al, 1998), approaches to dealing with unknown genetic models (see Risch, 2000), and genotyping error (see Moskvin et al, 2006). In contrast, similar issues contributing to phenotypic heterogeneity have been less widely considered, although they can cause substantial clinical variability between samples (Table 1). These include:

(a) Geographical origin. In addition to the likelihood that genetic contribution to illness varies between populations, there will be differences resulting from varying environmental exposures, sociocultural factors, service provision, etc.

(b) Ascertainment method. The spectrum of clinical features (symptoms, severity, functioning, illness course, etc.) of individuals recruited depends upon the mode of ascertainment. For example, inpatients at a tertiary referral centre differ from out-patients in secondary care.

(c) Unknown phenotypic model. Reliance on DSM-IV or ICD–10 categories obscures enormous clinical variability within categories. Perhaps of even greater concern, similarities across different categories of disorder are hidden.

(d) Measurement issues. Standardised methodologies for lifetime assessment of psychopathology are available. The same attention that is routinely given to technical issues of laboratory measurement must be given to correct, reliable and consistent use of phenotype measurement.

To maximise the potential of molecular genetic studies we need to pay much more attention to these phenotypic methodological issues than has recently been the norm.

### CONCLUSION

There is a need to use a range of genetic approaches that are appropriate to identifying a spectrum of risk variants from the common through to the rare. McClellan et al (2007) argue for approaches targeting risk alleles of large individual effect but low population effect size. Although data from genetic epidemiology and molecular genetics support the existence of some rare chromosomal abnormalities of large effect size, the evidence suggests rare variants of large effect size do not account for the majority of cases of schizophrenia. However, one unanswered possibility is that most genetic risk results from rare alleles of moderate effect size. Since for the next few years common alleles of modest effect size are likely to be more tractable than are rare
alleles, it seems appropriate that the focus in the immediate future will be on large samples and molecular genetic methods powered to detect the common alleles of modest effect size. Such approaches have only become available in the past 1–2 years; it is far too early to judge whether they have been successful or not; indeed, at the time of writing, no whole-genome-based surveys for common alleles of moderate effect size have been published. However, we predict that the interpretation of the data from such studies will be impeded by clinical variability across samples.

Replication of novel findings is essential. However, if as we suspect genetic variants that influence risk for psychiatric disorder influence aspects of the phenotype across DSM–IV/ICD–10 categories, and also influence only some aspects of the phenotype within these diagnostic categories, replication in psychiatric genetics will require close attention to both clinical psychiatric methodology and genetic methodology. We will need to become more sophisticated in our phenotypic thinking and move beyond studies and analyses based mainly on the traditional descriptive diagnostic categories (see Craddock & Owen, 2003; Marneros, 2007). We need to ensure that clinical psychiatry is placed very firmly at the heart of psychiatric genetics.

REFERENCES


NICK CRADDOCK, PhD, FRCPsych, MICHAEL C. O’DONOVAN, PhD, FRCPsych, MICHAEL J. OWEN, PhD, FRCPsych, Department of Psychological Medicine, Wales School of Medicine, Cardiff University, Heath Park, Cardiff CF4 4XN, UK.

Correspondence: Professor Nick Craddock, Department of Psychological Medicine, Henry Wellcome Building, Wales School of Medicine, Cardiff University, Heath Park, Cardiff CF4 4XN UK.

Email: craddockn@cardiff.ac.uk.

(First received 23 November 2006, final revision 17 January 2007, accepted 18 January 2007)
Bias in psychiatric case–control studies

Literature survey

WILLIAM LEE, JONATHAN BINDMAN, TAMSIN FORD, NICK GLOZIER, PAUL MORAN, ROBERT STEWART and MATTHEW HOTOPF

Background: Case–control studies are vulnerable to selection and information biases which may generate misleading findings.

Aims To assess the quality of methodological reporting of case–control studies published in general psychiatric journals.

Method All the case–control studies published over a 2-year period in the six general psychiatric journals with impact factors of more than 3 were assessed by a group of psychiatrists with training in epidemiology using a structured assessment devised for the purpose. The measured study quality was compared across type of exposure and journal.

Results The reporting of methods in the 408 identified papers was generally poor, with basic information about recruitment of participants often absent. Reduction of selection bias was described best in the ‘pencil and paper’ studies and worst in the genetic studies. Neuroimaging studies reported the most safeguards against information bias. Measurement of exposure was reported least well in studies determining the exposure with a biological test.

Conclusions Poor reporting of recruitment strategies threatens the validity of reported results and reduces the generalisability of studies.

Declaration of interest None.

Many studies in psychiatry compare biological, psychological, or social variables between healthy controls and individuals with psychiatric disorder. These studies are conceptually similar to the case–control study of epidemiology in that the participants are selected according to the presence or absence of a disorder. The two main sources of bias in case–control studies are selection bias and information bias. Selection bias exists where exposure status has a non-random effect on the selection of either cases or controls. The choice of the control group is crucial in this respect, since it functions to represent the level of exposure within the general population from which the cases have been identified. Information bias includes recall bias (where the participants’ illness experience systematically affects recall) and observer bias (where the knowledge the investigator has about the study hypothesis and of participants’ case or control status influences the assessment of the parameter under study). Case–control studies are an important source of evidence for many areas of mental health research. In a survey of papers published in leading non-specialist psychiatric journals, we evaluated the reported quality of the methods of case–control studies in psychiatry and evaluated the extent to which measures were taken to avoid these potential biases.

METHOD

Identification of studies

We hand-searched general psychiatric journals with an impact factor greater than 3 in 2001 from January 2001 to December 2002 inclusive. Studies were included if they compared participants with psychiatric disorders with healthy controls on any variable. Post-mortem studies were excluded, as were twin, co-twin and affected sibling designs.

Assessment of studies

We devised a data extraction form to describe the general characteristics of the paper, the selection of cases and controls, and the methods used to reduce information bias. We recorded the parameter compared between groups, the type and number of cases and the type and number of controls. If more than two diagnoses were studied we assigned the most numerous group to the cases, and did not collect details of other diagnostic groups. We also recorded details of individual matching and, if matching was performed, whether a matched analysis was used.

To examine selection bias we recorded details of the clinical setting where recruitment took place and whether the denominator from which cases were selected was described. For example, studies that reported recruiting patients with a specific diagnosis from consecutive series of referrals to a service, and gave details of the total number of patients eligible, would score for both items. We collected information on whether new (incident) cases were used, descriptions of the duration of illness, and the use of medication for disorders in which these data are relevant. We focused on the process by which recruitment was undertaken – in particular whether information was supplied on the total number of potential participants who were approached, the numbers of participants and non-participants, and whether differences between participants and non-participants were described. We also assessed whether inclusion and exclusion criteria were described in sufficient detail for the study to be replicated by other researchers. We recorded whether controls were recruited from students or employees of the organisation where the research was performed; whether they were selected from a defined population; whether they were recruited from advertisements; how many were approached; whether the differences between participant and non-participant controls were described; and whether similar exclusion criteria were applied to both cases and controls.

To assess information bias, we recorded whether the determination of exposure status had been carried out in a comparable way for both cases and controls and whether the investigators performing ratings had been masked to the participants’ illness status.

We piloted the rating scale by testing the interrater reliability of each item for
RESULTS

Interrater reliability
Twenty-two (5%) of the 408 papers were rated by all six of the raters. Seven of the papers were neuroimaging papers, eight were biological, six were pencil and paper, and one was a genetics paper. Of the 17 questions answered by the raters, three had a kappa value of greater than 0.8, five had kappa values between 0.6 and 0.8, two had kappa values between 0.4 and 0.6 and seven had kappa values of less than 0.4. (All but one of the questions had a percentage agreement in excess of 70% and many of those with the lowest kappa values had the highest percentage agreements. Even highly reliable measures show low kappa values when the expected frequency is low, as in this case; Altman, 2003). For each item on the questionnaire, a consensus answer was reached at a meeting of the raters. A manual was devised such that the raters using the manual gave the consensus answer on retesting.

Sample
The six journals that met the inclusion and exclusion criteria are listed in Table 1. From these journals 408 papers were identified. Eligible studies represent between 2% (Journal of Clinical Psychiatry) and 55% (Archives of General Psychiatry) of all published research. Papers reporting neuroimaging studies accounted for the largest number of papers in four of the six journals, with papers involving paper and pencil tests being the most frequent in the remaining two journals (Psychological Medicine and Journal of Clinical Psychiatry). Genetic papers were the least numerous in the sample (Table 1). Table 2 shows the study sample sizes by research area and journal. In general sample sizes were small, with a median group size of 23.5 (interquartile range 15.0–43.3). The groups were particularly small in biological and neuroimaging studies.

Selection bias
The questionnaire items concerning the clinical setting from which participants were recruited and medication use were described the most adequately, with 61% and 68% of papers respectively providing satisfactory information. Approximately half of the papers performed satisfactorily on the items concerning the use of similar exclusion criteria for cases and controls (57%) and the description of inclusion and exclusion criteria (50%). However, the reporting was particularly poor in four of the items: few of the papers fully described participants and non-participating potential cases (5%), or the differences between them (2%); similarly, information on the number of potential controls approached was rarely provided (5%), and only 1% of papers described the differences between participating controls and those who were approached to be controls but declined (Table 3). Two items (the use of students or employees of the research institution and the use of advertising for recruitment) were very frequently rated as ‘unclear’, indicating insufficient information was available to make a judgement. However, at least a third of all studies used advertisements to recruit controls, and at least 15% used staff or students from the research institution as controls.

Table 1 Distribution of the included case-control studies between journals and areas of research

<table>
<thead>
<tr>
<th>Journal</th>
<th>Research area, n (%)</th>
<th>Case–control studies n (%)</th>
<th>Total published studies n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neuroimaging</td>
<td>Biological</td>
<td>Pencil and paper</td>
</tr>
<tr>
<td>American Journal of Psychiatry</td>
<td>66 (47)</td>
<td>37 (26)</td>
<td>34 (24)</td>
</tr>
<tr>
<td>Archives of General Psychiatry</td>
<td>26 (46)</td>
<td>18 (32)</td>
<td>9 (16)</td>
</tr>
<tr>
<td>Biological Psychiatry</td>
<td>57 (47)</td>
<td>43 (35)</td>
<td>13 (11)</td>
</tr>
<tr>
<td>British Journal of Psychiatry</td>
<td>10 (42)</td>
<td>5 (21)</td>
<td>9 (38)</td>
</tr>
<tr>
<td>Journal of Clinical Psychiatry</td>
<td>0 (0)</td>
<td>1 (20)</td>
<td>4 (80)</td>
</tr>
<tr>
<td>Psychological Medicine</td>
<td>10 (17)</td>
<td>11 (18)</td>
<td>39 (65)</td>
</tr>
<tr>
<td>Total</td>
<td>169 (41)</td>
<td>115 (28)</td>
<td>108 (26)</td>
</tr>
</tbody>
</table>

1. Values in parentheses are percentages of row totals.
2. Values in parentheses are percentages of column total.
3. Values in parentheses are the proportion of papers published in each journal that are case-control studies.
Information bias

Most (93%) papers reported that they assessed exposure status in a sufficiently similar way for cases and controls (Table 3), but only 25% indicated that the investigators were 'masked' to the illness status of the participants, and in 70% of the papers it was impossible to determine whether the investigators were 'masked' or not.

### Table 2 Median size and interquartile range of the largest case group in each study

<table>
<thead>
<tr>
<th>Research area</th>
<th>Studies, n (%)</th>
<th>Sample size, Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neuroimaging</td>
<td>169 (41)</td>
<td>18 (12–30)</td>
</tr>
<tr>
<td>Biological</td>
<td>115 (28)</td>
<td>21 (15–38)</td>
</tr>
<tr>
<td>Pencil and paper</td>
<td>108 (26)</td>
<td>38.5 (23–88.5)</td>
</tr>
<tr>
<td>Genetic</td>
<td>16 (4)</td>
<td>108 (36–177.5)</td>
</tr>
</tbody>
</table>

Journal

<table>
<thead>
<tr>
<th>Journal</th>
<th>Median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Journal of Psychiatry</td>
<td>22.5 (15–41)</td>
</tr>
<tr>
<td>Archives of General Psychiatry</td>
<td>22 (15–40)</td>
</tr>
<tr>
<td>Biological Psychiatry</td>
<td>21 (14–42)</td>
</tr>
<tr>
<td>British Journal of Psychiatry</td>
<td>25.5 (17–88)</td>
</tr>
<tr>
<td>Journal of Clinical Psychiatry</td>
<td>34 (20–40)</td>
</tr>
<tr>
<td>Psychological Medicine</td>
<td>26 (19.5–49)</td>
</tr>
</tbody>
</table>

Total

| Total                          | 408 (100) | 23.5 (15–43.5) |

IQR, interquartile range.

**Information bias**

**Overall quality of the papers**

Studies that used pencil and paper tests showed significantly more desirable methodological features in the selection of both cases and controls than the studies in other research areas. Genetic studies were rated poorest in the selection of cases. Neuroimaging studies showed most desirable features.

### Table 3 Answers to items in the questionnaire used to evaluate the methodological quality of the case–control studies.

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes n (%)</th>
<th>No n (%)</th>
<th>Unclear n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Selection bias</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cases</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was the clinical setting used for recruitment made clear?</td>
<td>248 (61)</td>
<td>160 (39)</td>
<td></td>
</tr>
<tr>
<td>Was the denominator from which cases were recruited described?</td>
<td>96 (24)</td>
<td>301 (74)</td>
<td>11 (3)</td>
</tr>
<tr>
<td>Were incident cases used?</td>
<td>44 (11)</td>
<td>344 (84)</td>
<td>20 (5)</td>
</tr>
<tr>
<td>Was duration of illness adequately described?</td>
<td>174 (43)</td>
<td>212 (52)</td>
<td>22 (5)</td>
</tr>
<tr>
<td>Was medication use adequately described?</td>
<td>277 (68)</td>
<td>86 (21)</td>
<td>45 (11)</td>
</tr>
<tr>
<td>Was adequate information given on the total number of patients approached?</td>
<td>43 (11)</td>
<td>357 (88)</td>
<td>8 (2)</td>
</tr>
<tr>
<td>Was information given on participants and non-participants?</td>
<td>20 (5)</td>
<td>379 (93)</td>
<td>9 (2)</td>
</tr>
<tr>
<td>Was information given on the differences between participants and refusers?</td>
<td>9 (2)</td>
<td>390 (96)</td>
<td>9 (2)</td>
</tr>
<tr>
<td>Were the inclusion and exclusion criteria described well enough to be replicable?</td>
<td>203 (50)</td>
<td>205 (50)</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Did the study use controls who were students/employees of the research institution?</td>
<td>56 (14)</td>
<td>125 (31)</td>
<td>227 (56)</td>
</tr>
<tr>
<td>Were controls selected from an explicit sampling frame?</td>
<td>67 (16)</td>
<td>332 (81)</td>
<td>9 (2)</td>
</tr>
<tr>
<td>Did the study recruit through advertisements?</td>
<td>143 (35)</td>
<td>46 (11)</td>
<td>219 (54)</td>
</tr>
<tr>
<td>Were similar exclusion criteria applied for controls as for cases?</td>
<td>231 (57)</td>
<td>32 (8)</td>
<td>145 (36)</td>
</tr>
<tr>
<td>Was information given on number of controls approached?</td>
<td>21 (5)</td>
<td>375 (92)</td>
<td>12 (3)</td>
</tr>
<tr>
<td>Was adequate information given on differences between controls refusing and agreeing?</td>
<td>3 (1)</td>
<td>395 (97)</td>
<td>10 (2)</td>
</tr>
<tr>
<td><strong>Information bias</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were 'exposure' status performed in a sufficiently similar way?</td>
<td>381 (93)</td>
<td>8 (2)</td>
<td>19 (5)</td>
</tr>
<tr>
<td>Were the investigators who rated the exposure masked to participants' status?</td>
<td>95 (25)</td>
<td>16 (4)</td>
<td>265 (70)</td>
</tr>
</tbody>
</table>

1. Values in parentheses are row percentages; the denominator in each row is the 408 papers included in the study.
2. Denotes that 'no' is the answer indicative of good methodological practice.
3. Thirty-two studies were removed from the denominator in this item because no human decision was required to rate the exposure.
in the elimination of information bias (Table 4).

Papers published in *Biological Psychiatry* were rated as showing fewest desirable features in the recruitment of cases and controls. Papers published in *Archives of General Psychiatry* showed significantly superior methodology in reducing selection bias of controls compared with papers published in other journals (Table 4).

The data from our three quality rating scales are shown in histogram form in Figs 1–3.

**DISCUSSION**

The case–control study design is common across many areas of psychiatric research, as it is a cost-effective study design, especially for relatively rare psychiatric outcomes such as psychotic illness. In this review, we found that the general level of methodological description was poor and many of the papers failed to include sufficient information to allow a judgement to be made about the impact of selection or information biases on the findings of the studies. Genetic studies achieved the poorest ratings in reducing selection bias, whereas pencil and paper studies achieved the best. Neuroimaging studies gave the most complete information relevant to information bias. There were few differences between journals in the reporting of measures to reduce information biases.

The recruitment of participants was not described well in most of the studies examined. This means that the generalisability of the findings arising from these studies cannot be assessed, and that accurate replication of the study in a different population or time period becomes impossible. In case–control studies the control group functions to represent the level of exposure within the general population from which the cases have been identified, and researchers should ensure that the selection of cases and controls takes place within a defined population in as transparent and reproducible a manner as possible (Wacholder, 1995). The practice of advertising within a research institution to recruit controls – who are frequently students or staff members of that organisation – is widespread and is likely to introduce biases which may be difficult to quantify. It is not improbable that the often subtle experimental...

---

### Table 4

<table>
<thead>
<tr>
<th>Research area</th>
<th>Selection bias (cases)</th>
<th>Selection bias (controls)</th>
<th>Information bias</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0–9)</td>
<td>(0–6)</td>
<td>(0–2)</td>
</tr>
<tr>
<td></td>
<td>Median (IQR)</td>
<td>Median (IQR)</td>
<td>Median (IQR)</td>
</tr>
<tr>
<td>Neuroimaging</td>
<td>3 (2–4)</td>
<td>1 (0–1)</td>
<td>1 (1–2)</td>
</tr>
<tr>
<td>Biological</td>
<td>3 (2–4)</td>
<td>1 (0–2)</td>
<td>1 (1–1)</td>
</tr>
<tr>
<td>Pencil and paper</td>
<td>3 (2–5)</td>
<td>1 (0–2.5)</td>
<td>1 (1–1)</td>
</tr>
<tr>
<td>Genetic</td>
<td>2 (1–3)</td>
<td>1 (0–1)</td>
<td>1 (1–1)</td>
</tr>
<tr>
<td>American Journal of Psychiatry</td>
<td>3 (2–4)</td>
<td>1 (0–1)</td>
<td>1 (1–1)</td>
</tr>
<tr>
<td>Archives of General Psychiatry</td>
<td>3 (2–4)</td>
<td>1 (1–2)</td>
<td>1 (1–2)</td>
</tr>
<tr>
<td>Biological Psychiatry</td>
<td>3 (1–4)</td>
<td>1 (0–1)</td>
<td>1 (1–1)</td>
</tr>
<tr>
<td>British Journal of Psychiatry</td>
<td>3.5 (1.5–5)</td>
<td>1 (1–2.5)</td>
<td>1 (1–2)</td>
</tr>
<tr>
<td>Journal of Clinical Psychiatry</td>
<td>4 (3–5)</td>
<td>2 (0–2)</td>
<td>1 (1–1)</td>
</tr>
<tr>
<td>Psychological Medicine</td>
<td>3 (2–4)</td>
<td>1 (0.5–2)</td>
<td>1 (1–1)</td>
</tr>
<tr>
<td>Total</td>
<td>3 (2–4)</td>
<td>1 (0–2)</td>
<td>1 (1–1)</td>
</tr>
</tbody>
</table>

IQR, interquartile range.

1. The median of the group is statistically significantly different from the median of the other groups making up the entire sample, using the Kruskal–Wallis test with a *p* value required for significance corrected for multiple comparisons to 0.0125 for the research areas and 0.00833 for the journals.
conditions devised in functional brain imaging studies may be influenced by educational level or motivation to participate in research. Further, the poor quality of reporting of the selection of cases suggests that many studies use what are effectively ‘convenience’ samples, which will tend to comprise the most severe and treatment-resistant cases in a service. These two opposing factors – ‘super-healthy’ controls and unrepresentatively ill cases – are likely to lead to an overestimate of effect sizes (Lewis & Pelosi, 1990).

The masking of raters was generally poorly reported. There are, no doubt, situations in which a parameter can be estimated without any risk of observer bias and therefore with no theoretical need for masking. However, it is difficult to determine when these situations are present. Many apparently ‘hard’ outcomes – such as volume of brain structures or concentrations of immune parameters – involve a good deal of measurement performed by humans and are therefore open to observer bias (Sackett, 1979). It is hard to envisage a situation where masking of those performing such ratings is not feasible, and we can think of no situation where to attempt masking would be harmful. We therefore suggest that authors have a duty either to report that masking took place or the reasons why this was unnecessary. In the majority of papers we assessed, this information was not available. Those reading the papers without a detailed knowledge of the techniques used have no idea whether observer bias is a possible explanation of the reported findings.

Unlike chance and confounding, bias cannot be readily quantified, may not be detectable and cannot be taken into account in data analysis. This means that the only opportunity to reduce the influence of bias on the results of a study is at the design phase. Problems with the methodology and reporting of randomised controlled trials were observed in the 1990s (Schulz, 1995a,b,c, 1996; Hotopf et al., 1997; Ogundipe et al., 1999). An outcome of this was the Consolidated Standards of Reporting Trials (CONSORT) statement, in which authors are required to describe their methodology according to a 22-item checklist (Altman et al., 2001). This has unified clinicians, academics, policy makers and the pharmaceutical industry, and is now a mandatory part of submissions of randomised controlled trials to major journals.

A number of reviews have documented many areas of scientific research where the findings of case–control studies have not been replicated in methodologically superior prospective cohort studies (Mayes et al., 1988; Pocock et al., 2004; von Elm & Egger, 2004). In psychiatry, the emerging finding that large, population-based case–control neuroimaging studies in psychosis (Dazzan et al., 2003; Busatto et al., 2004) have failed to replicate the multitude of small, clinic-based case–control studies that preceded them (Shenton et al., 2001) suggests that the findings of the latter may owe much to the processes involved in selecting cases and controls.

The Strengthening the Reporting of Observational studies in Epidemiology (STROBE) initiative is an attempt to bring about improvements to the methodology and reporting of observational studies, by publishing a checklist with which it is intended all observational research reports will have to comply as a condition of publication (Altman et al., 2005). We are optimistic that efforts such as this will improve the standard of reporting and methodology in psychiatric case–control studies in future years.

Although the main aim of our review was to assess potential sources of bias in case–control studies, we noted that many studies had very small sample sizes, with a quarter of all studies having no more than 15 cases. Small sample sizes lead to type II error – when a genuine difference between groups is not detected. We also noted that sample sizes varied to a large extent according to the parameter under study. Neuroimaging and ‘biological’ studies generally had much smaller sample sizes than did genetic and ‘pencil and paper’ studies. It is difficult to make a general recommendation about the sample size required for the question under study, and variation between methods may be owing to differences in what investigators perceive to be an effect size worth detecting. Differences may also arise because the parameter under study may be measured as a continuous variable (e.g. the volume of a brain structure) or a categorical variable (e.g. the presence of a specific genotype); the use of continuous variables improves power, and therefore smaller sample sizes can be used. However, we also suspect that the expense of performing complex neuroimaging studies or biological assays might mean that these studies are particularly prone to be underpowered.

We were surprised that many studies were individually matched without it being clear that a matched analysis was executed, as this practice results in the needless loss of statistical power (Miettinen, 1970). This and the prevalence of non-equal group sizes in ‘matched’ studies illustrate some of the many problems with individual matching and explain why this technique has largely been superseded in epidemiology by the use of the more flexible multivariable statistical methods (Prentice, 1976; Rosner & Hennekens, 1978).

This review has several limitations. We undertook to examine studies published only in the highest-impact general psychiatric journals; this was done over a limited period; we only examined one case group and one control group from each study, and the rating scales were simply constructed. We chose the journals with high impact factors to target studies likely to represent accepted practice, where one might expect only examples of good methodology to be accepted, and therefore papers published in less prestigious journals may have even poorer reporting of methodology. The 2-year period we chose was the most recent period for which we had impact factors when the hand-searching was started. We only chose one case group and one control group from each study to simplify our method and analyses. We believe this made little difference to our findings, as most of the studies had only two groups, and in studies with more the methods of selection and reporting of the other groups tended to be similar. Our sampling frame was explicit and representative, including journals from the UK and the USA, and our inclusion and exclusion criteria were predetermined. We feel that the results of this review are likely to represent the standard of global English-language accepted practice of the reporting of psychiatric case–control studies in 2001 and 2002, and we suspect that the standards of reporting of case–control studies are unlikely to have improved markedly since then. The construction of the three rating scales, simply adding the number of questions answered to indicate good practice within the three sections of the questionnaire, was chosen as the most straightforward method of indicating the general quality of the studies. The authors believe that although equating the methodological characteristics of the papers may seem arbitrary, all the items on the questionnaire are important, so none should be deemed less important than any
BIAS IN CASE-CONTROL STUDIES

WILLIAM LEE, MBChB, MRCPsych, Academic Department of Psychological Medicine; JONATHAN BINDMAN, MSc, PhD, MRCPsych, Health Services Research; TAMSIN FORD, MSc, PhD, MRCPsych, Child and Adolescent Psychiatry; NICK GLOZIER, MSc, PhD, MRCPsych, PAUL MORAAN, MSc, MD, MRCPsych, Health Services Research; ROBERT STEWART, MSc, MD, MRCPsych, Section of Epidemiology; MATTHEW HOTOPF, MSc, PhD, MRCPsych, Academic Department of Psychological Medicine, Institute of Psychiatry, King’s College London, UK

Correspondence: Professor Matthew Hotopf, Department of Psychological Medicine, King’s College London, Institute of Psychiatry, Weston Education Centre, 10 Cutcombe Rd, London SE5 9RJ, UK. Email: m.hotopf@iop.kcl.ac.uk

(First received 8 June 2006, accepted 1 September 2006)

REFERENCES


other. The number of questions in each of the rating scales was small (9, 6 and 2 respectively) which could leave the results vulnerable to floor and ceiling effects, potentially not detecting true associations. Although the numbers are small, on inspection of the data (see Figs 1–3) the authors do not think that large effects are likely to have been undetected.

We have shown that there is a tendency for psychiatric researchers to ignore the potential impact of bias on their results. It is impossible to determine whether the studies we included simply reported their methods inadequately or used inadequate methods. We suggest that researchers have a responsibility to reassure readers that appropriate steps have been taken to eliminate bias, and at present this is not happening.
Diagnostic stability of psychiatric disorders
in clinical practice

ENRIQUE BACA-GARCIA, MARIA M. PEREZ-RODRIGUEZ,
IGNACIO BASURTE-VILLAMOR, ANTONIO L. FERNANDEZ DEL MORAL,
MIGUEL A. JIMENEZ-ARRIERO, JOSE L. GONZALEZ DE RIVERA,
JERONIMO SAIZ-RUIZ and MARIA A. OQUENDO

Background  Psychiatric disorders are among the top causes worldwide of disease burden and disability. A major criterion for validating diagnoses is stability over time.

Aims  To evaluate the long-term stability of the most prevalent psychiatric diagnoses in a variety of clinical settings.

Method  A total of 34,368 patients received psychiatric care in the catchment area of one Spanish hospital (1992–2004). This study is based on 10,025 adult patients who were assessed on at least ten occasions (360,899 psychiatric consultations) in three settings: in-patient unit, 2000–2004 (n=546); psychiatric emergency room, 2000–2004 (n=1408); and out-patient psychiatric facilities, 1992–2004 (n=10,016). Prospective consistency, retrospective consistency and the proportion of patients who received each diagnosis in at least 75% of the evaluations were calculated for each diagnosis in each setting and across settings.

Results  The temporal consistency of mental disorders was poor, ranging from 29% for specific personality disorders to 70% for schizophrenia, with stability greatest for in-patient diagnoses and least for out-patient diagnoses.

Conclusions  The findings are an indictment of our current psychiatric diagnostic practice.

Declaration of interest  None.

Diagnosis is essential in clinical practice, research, training and public health. Definitions for psychiatric diagnoses are derived from expert opinion rather than the biological basis of the disorder. The modest knowledge base regarding the causation of disease has hindered the use of aetiological factors in psychiatric classification systems. The current classifications (World Health Organization, 1992; American Psychiatric Association, 2000) were designed to achieve high interrater reliability of diagnostic assessment. It is widely believed that if future editions of the DSM and the ICD are to be a significant improvement on their predecessors, the validity of the diagnostic concepts they include will have to be enhanced (Kendell & Jablensky, 2003). Follow-up studies including evidence of diagnostic stability and diagnostic consistency over time have traditionally been proposed to test the validity of psychiatric diagnoses (Robins & Guze, 1970; Kendler, 1980; Andreasen, 1995). However, several authors have noted that as longitudinal data become available, significant fluctuations in diagnostic stability and changes in clinical presentation are seen (Krishnan, 2005).

The aim of our study was to evaluate the long-term stability of the most prevalent chronic psychiatric diagnoses according to ICD–10 in a range of clinical settings.

METHODS

Participants

In total 34,368 patients received psychiatric care in the catchment area of Fundacion Jimenez Diaz General Hospital, Madrid, between 1 January 1992 and 31 December 2004. This hospital is part of the Spanish national health services and provides free medical coverage to a catchment area of 280,000 people. There were 449,317 psychiatric consultations in a variety of clinical settings, including visits to out-patient psychiatric facilities (438,622), emergency visits (9101) and admissions to the psychiatric brief hospitalisation unit (1594). The current study is based on 10,025 patients aged 18 years and over who were assessed on at least ten occasions during the period studied. These patients had 360,899 psychiatric consultations, including visits to out-patient psychiatric facilities (355,166), psychiatric emergency visits (4628) and admissions to the psychiatric brief hospitalisation unit (1105).

Individual service users are reliably identified in the database used for our analyses because each patient is given an identifying number (a numeric code is used to ensure patient anonymity), which remains the same throughout all contacts with psychiatric services within the study area. To ensure that no patient had been assigned more than one identifier, we reviewed all the cases in the database and removed any duplicates we found. We defined duplicates as ‘patients with identical first name, family name, gender and year of birth’, ‘patients with identical first name, family name, gender and street address’, or ‘patients with identical first name, family name, gender and hospital/ambulatory record number’. We deleted any cases with significant suspicion of duplication.

Settings

Participants (n=10,025) were assessed in three different clinical settings: in-patient unit (psychiatric brief hospitalisation unit), 2000–2004 (n=546); psychiatric emergency room, 2000–2004 (n=1408); and out-patient psychiatric facilities (mental health care centres) within the catchment area of the Fundacion Jimenez Diaz General Hospital, 1992–2004 (n=10,016).

Diagnostic procedures

Procedure during ambulatory visits

Since 1986 public mental health centres within the province of Madrid have had to record all ambulatory visits in a regional registry, the Registro Acumulativo de Casos de la Comunidad de Madrid. All diagnoses in this registry must be coded according to the ICD–9 (World Health Organization, 1978). Since 1992 diagnoses have been assigned according to ICD–10 (World Health Organization, 1992) criteria and recorded with the appropriate ICD–9 coding numbers; ICD–10 codes were converted to ICD–9 codes using the guidelines published by the World Health Organization (Organización Mundial de la Salud, 1993). The

Procedure during emergency visits

The emergency diagnoses were taken from the emergency medical records. Emergency diagnoses were assigned by clinical psychiatrists after reviewing all available information, including data from clinical interviews with the patient and relatives.

Procedure during admissions to the in-patient unit

Clinical diagnoses during admissions are the result of an intensive diagnostic and treatment process by physicians with speciality training in psychiatry, including data from medical records, other research assessments and clinical interviews. The psychiatrists assigned the clinical diagnoses were not aware of the study in process.

Diagnostic groups included in analysis

Among all chronic psychiatric diagnoses, we selected those diagnoses assigned to more than 500 patients in our sample (prevalence higher than 5%). According to data from naturalistic studies like ours, the frequency and use of the ICD–10 two-digit, three-digit and four-digit diagnostic categories show significant variations. Some categories are not used at all, and others represent less than 0.1% of the sample studied (Mussigbrodt et al., 2000). In the latter study of a sample of 33857 treated cases from 19 departments of psychiatry in ten different countries, ‘on a four-character level (Fxx.x), the ten most often used diagnostic categories represented 40% of all main diagnoses, and 70% on a three-character level (Fxx.)’ (Mussigbrodt et al., 2000). The diagnoses analysed here (with ICD–10 codes) are:

(a) schizophrenia, schizotypal and delusional disorders (F20–29), including individual diagnoses of schizophrenia (F20), paranoid schizophrenia (F20.0), residual schizophrenia (F20.5) and persistent delusional disorders (F22);
(b) mood (affective) disorders (F30–39), including individual diagnoses of bipolar affective disorder (F31), bipolar affective disorder, current episode mild or moderate depression (F31.3), recurrent depressive disorder (F33), persistent mood (affective) disorders (F34), and dysthymia (F34.1);
(c) obsessive–compulsive disorder (F42);
(d) eating disorders (F50);
(e) disorders of adult personality and behaviour (F60–69), including the individual diagnoses of specific personality disorders (F60) and other specific personality disorders (F60.8).

Data extraction and analysis

Diagnostic stability through all the evaluations is calculated according to Schwartz et al. (2000). Three measures of stability are presented for each diagnosis. The first, ‘prospective consistency’, is the proportion of individuals in a category at the first evaluation who retain the same diagnosis at their last evaluation. This would correspond to positive predictive value if the last diagnosis were the gold standard. The second measure, retrospective consistency, is the proportion of individuals with a diagnosis assigned at the last evaluation who had received the same diagnosis at the first evaluation; this is conceptually similar to sensitivity. The third measure is the proportion of patients who received the same diagnosis in at least 75% of the evaluations. The agreement between diagnoses at the first and the last evaluations was calculated by the kappa coefficient, which measures the agreement correcting the effect of chance.

Using the Statistical Package for the Social Sciences, version 13.0 for Windows, we performed four different analyses: three separate analyses for each clinical setting (psychiatric emergencies, out-patient visits and hospitalisations) to control for influences of the setting on the stability of diagnoses; and a fourth analysis of the combined data from the three clinical settings to reflect the evolution of diagnoses through the clinical process.

RESULTS

The socio-demographic characteristics of the sample are presented in Table 1.

Stability of diagnoses

Data about the prospective and retrospective consistency of the diagnoses across settings, in the out-patient setting, in the emergency setting and in the in-patient setting are presented in Tables 2–5 and graphically in a data supplement to the online version of this paper. The percentages of patients who received the same diagnosis in at least 75% of their evaluations, across settings, in the out-patient setting, in the emergency setting and in the in-patient setting are presented in Table 6.

Across clinical settings

Prospective consistency ranged from 28.7% for other specific personality disorders to 69.6% for schizophrenia, (Table 2).

Table 1: Socio-demographic characteristics of the sample (n=10 025)

<table>
<thead>
<tr>
<th>Variable</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3752 (37.4)</td>
</tr>
<tr>
<td>Female</td>
<td>6186 (61.7)</td>
</tr>
<tr>
<td>Transsexual</td>
<td>6 (0.1)</td>
</tr>
<tr>
<td>Missing data</td>
<td>81 (0.8)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>5281 (52.7)</td>
</tr>
<tr>
<td>Married</td>
<td>2923 (29.2)</td>
</tr>
<tr>
<td>Divorced</td>
<td>320 (3.2)</td>
</tr>
<tr>
<td>Widow</td>
<td>620 (6.2)</td>
</tr>
<tr>
<td>Missing data</td>
<td>881 (8.8)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Illiterate</td>
<td>88 (0.9)</td>
</tr>
<tr>
<td>Never went to school</td>
<td>533 (5.3)</td>
</tr>
<tr>
<td>Primary school</td>
<td>2401 (24.0)</td>
</tr>
<tr>
<td>High school</td>
<td>3617 (36.1)</td>
</tr>
<tr>
<td>University</td>
<td>2491 (24.8)</td>
</tr>
<tr>
<td>Other education</td>
<td>49 (0.5)</td>
</tr>
<tr>
<td>Missing data</td>
<td>846 (8.4)</td>
</tr>
<tr>
<td>Accommodation</td>
<td></td>
</tr>
<tr>
<td>Alone</td>
<td>1907 (19.0)</td>
</tr>
<tr>
<td>With partner</td>
<td>3352 (33.4)</td>
</tr>
<tr>
<td>With parents</td>
<td>2755 (27.4)</td>
</tr>
<tr>
<td>With children</td>
<td>675 (6.7)</td>
</tr>
<tr>
<td>With other family members</td>
<td>530 (5.3)</td>
</tr>
<tr>
<td>In an institution</td>
<td>86 (0.8)</td>
</tr>
<tr>
<td>Adopted</td>
<td>280 (2.8)</td>
</tr>
<tr>
<td>Missing data</td>
<td>440 (4.4)</td>
</tr>
<tr>
<td>Current working status</td>
<td></td>
</tr>
<tr>
<td>Military service</td>
<td>14 (0.1)</td>
</tr>
<tr>
<td>Employed</td>
<td>3617 (36.1)</td>
</tr>
<tr>
<td>Looking for first job</td>
<td>92 (0.9)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>1221 (12.2)</td>
</tr>
<tr>
<td>Retired</td>
<td>1133 (11.3)</td>
</tr>
<tr>
<td>Student</td>
<td>1243 (12.4)</td>
</tr>
<tr>
<td>Homemaker</td>
<td>1058 (10.6)</td>
</tr>
<tr>
<td>Transient disability</td>
<td>425 (4.2)</td>
</tr>
<tr>
<td>Permanent disability</td>
<td>186 (1.9)</td>
</tr>
<tr>
<td>Missing data</td>
<td>1036 (10.3)</td>
</tr>
</tbody>
</table>
prospective consistency of the three most prevalent diagnoses at first evaluation was 44.7% for dysthymia, 69.6% for schizophrenia and 49.4% for bipolar affective disorder (see Table 2). Retrospective consistency at the last evaluation ranged from 23.4% for bipolar affective disorder, current episode mild or moderate depression (F31.3) to 29.4% for eating disorders (F50). Patients who received the same diagnosis during at least 75% of their evaluations ranged from 9.8% for other specific personality disorders (F60.8) to 58.0% for eating disorders (F50).

### Table 2: Prospective and retrospective consistency of ICD-10 diagnoses across settings (n=10 025)

<table>
<thead>
<tr>
<th>Diagnosis (ICD-10 code)</th>
<th>First evaluation</th>
<th>Last evaluation</th>
<th>First v. last evaluation</th>
<th>Prospective consistency</th>
<th>Retrospective consistency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>n</td>
<td>k1</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Schizophrenia, schizotypal and delusional disorders (F20–F29)</td>
<td>878</td>
<td>1103</td>
<td>0.6</td>
<td>68.6</td>
<td>54.6</td>
</tr>
<tr>
<td>Schizophrenia (F20)</td>
<td>540</td>
<td>819</td>
<td>0.5</td>
<td>69.6</td>
<td>45.9</td>
</tr>
<tr>
<td>Paranoid schizophrenia (F20.0)</td>
<td>292</td>
<td>427</td>
<td>0.4</td>
<td>50.0</td>
<td>34.2</td>
</tr>
<tr>
<td>Residual schizophrenia (F20.5)</td>
<td>148</td>
<td>304</td>
<td>0.3</td>
<td>49.3</td>
<td>24.0</td>
</tr>
<tr>
<td>Persistent delusional disorders (F22)</td>
<td>148</td>
<td>155</td>
<td>0.3</td>
<td>34.5</td>
<td>23.9</td>
</tr>
<tr>
<td>Mood (affective) disorders (F30–F39)</td>
<td>2204</td>
<td>2322</td>
<td>0.4</td>
<td>54.9</td>
<td>52.2</td>
</tr>
<tr>
<td>Bipolar affective disorder (F31)</td>
<td>342</td>
<td>443</td>
<td>0.4</td>
<td>49.4</td>
<td>38.1</td>
</tr>
<tr>
<td>Bipolar affective disorder, current episode mild or moderate depression (F31.3)</td>
<td>127</td>
<td>192</td>
<td>0.3</td>
<td>35.4</td>
<td>23.4</td>
</tr>
<tr>
<td>Recurrent depressive disorder (F33)</td>
<td>268</td>
<td>267</td>
<td>0.4</td>
<td>40.3</td>
<td>40.4</td>
</tr>
<tr>
<td>Persistent mood (affective) disorders (F34)</td>
<td>1424</td>
<td>1457</td>
<td>0.3</td>
<td>44.6</td>
<td>43.6</td>
</tr>
<tr>
<td>Dysthymia (F34.1)</td>
<td>1397</td>
<td>1429</td>
<td>0.4</td>
<td>44.7</td>
<td>43.7</td>
</tr>
<tr>
<td>Obsessive–compulsive disorder (F42)</td>
<td>157</td>
<td>212</td>
<td>0.4</td>
<td>46.5</td>
<td>34.4</td>
</tr>
<tr>
<td>Eating disorders (F50)</td>
<td>195</td>
<td>188</td>
<td>0.6</td>
<td>55.9</td>
<td>58.0</td>
</tr>
<tr>
<td>Disorders of adult personality and behaviour (F60–F69)</td>
<td>378</td>
<td>471</td>
<td>0.3</td>
<td>34.7</td>
<td>27.8</td>
</tr>
<tr>
<td>Specific personality disorders (F60)</td>
<td>352</td>
<td>457</td>
<td>0.3</td>
<td>34.1</td>
<td>26.3</td>
</tr>
<tr>
<td>Other specific personality disorders (F60.8)</td>
<td>136</td>
<td>148</td>
<td>0.3</td>
<td>28.7</td>
<td>26.4</td>
</tr>
</tbody>
</table>

1. All kappa statistics are significant (P < 0.001).

### Table 3: Prospective and retrospective consistency of ICD-10 diagnoses in the out-patient setting (n=10 016)

<table>
<thead>
<tr>
<th>Diagnosis (ICD-10 code)</th>
<th>First evaluation</th>
<th>Last evaluation</th>
<th>First v. last evaluation</th>
<th>Prospective consistency</th>
<th>Retrospective consistency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>n</td>
<td>k1</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Schizophrenia, schizotypal and delusional disorders (F20–F29)</td>
<td>875</td>
<td>1088</td>
<td>0.6</td>
<td>68.3</td>
<td>55.0</td>
</tr>
<tr>
<td>Schizophrenia (F20)</td>
<td>538</td>
<td>809</td>
<td>0.5</td>
<td>69.1</td>
<td>46.0</td>
</tr>
<tr>
<td>Paranoid schizophrenia (F20.0)</td>
<td>290</td>
<td>427</td>
<td>0.4</td>
<td>49.3</td>
<td>33.5</td>
</tr>
<tr>
<td>Residual schizophrenia (F20.5)</td>
<td>148</td>
<td>304</td>
<td>0.3</td>
<td>50.7</td>
<td>24.7</td>
</tr>
<tr>
<td>Persistent delusional disorders (F22)</td>
<td>148</td>
<td>158</td>
<td>0.3</td>
<td>35.1</td>
<td>32.9</td>
</tr>
<tr>
<td>Mood (affective) disorders (F30–F39)</td>
<td>2203</td>
<td>2343</td>
<td>0.4</td>
<td>55.6</td>
<td>52.2</td>
</tr>
<tr>
<td>Bipolar affective disorder (F31)</td>
<td>342</td>
<td>440</td>
<td>0.4</td>
<td>50.6</td>
<td>39.3</td>
</tr>
<tr>
<td>Bipolar affective disorder, current episode mild or moderate depression (F31.3)</td>
<td>127</td>
<td>194</td>
<td>0.3</td>
<td>35.4</td>
<td>23.2</td>
</tr>
<tr>
<td>Recurrent depressive disorder (F33)</td>
<td>268</td>
<td>270</td>
<td>0.4</td>
<td>40.3</td>
<td>40.0</td>
</tr>
<tr>
<td>Persistent mood (affective) disorders (F34)</td>
<td>1424</td>
<td>1496</td>
<td>0.4</td>
<td>45.8</td>
<td>43.6</td>
</tr>
<tr>
<td>Dysthymia (F34.1)</td>
<td>1397</td>
<td>1464</td>
<td>0.4</td>
<td>45.7</td>
<td>43.6</td>
</tr>
<tr>
<td>Obsessive–compulsive disorder (F42)</td>
<td>157</td>
<td>213</td>
<td>0.4</td>
<td>47.1</td>
<td>34.7</td>
</tr>
<tr>
<td>Eating disorders (F50)</td>
<td>194</td>
<td>189</td>
<td>0.6</td>
<td>56.2</td>
<td>57.7</td>
</tr>
<tr>
<td>Disorders of adult personality and behaviour (F60–F69)</td>
<td>375</td>
<td>456</td>
<td>0.3</td>
<td>35.7</td>
<td>29.4</td>
</tr>
<tr>
<td>Specific personality disorders (F60)</td>
<td>351</td>
<td>442</td>
<td>0.3</td>
<td>35.6</td>
<td>28.3</td>
</tr>
<tr>
<td>Other specific personality disorders (F60.8)</td>
<td>136</td>
<td>156</td>
<td>0.3</td>
<td>29.4</td>
<td>25.6</td>
</tr>
</tbody>
</table>

1. All kappa statistics are significant (P < 0.001).
Out-patient setting

Prospective consistency ranged from 29.4% for other specific personality disorders to 69.1% for schizophrenia. The prospective consistency of the three most prevalent specific diagnoses at the first evaluation was 45.7% for dysthymia, 69.1% for schizophrenia and 50.6% for bipolar affective disorder (see Table 3). Retrospective consistency at the last evaluation ranged from 23.2% for bipolar affective disorder, current episode mild or moderate depression (F31.3), to 57.7% for eating disorders; it was 43.6% for dysthymia, 46.0% for schizophrenia and 39.3% for bipolar affective disorder (see Table 3). The proportion of patients who received the same diagnosis...
during at least 75% of the evaluations ranged from 10.7% for other specific personality disorders to 49.6% for schizophrenia, schizotypal and delusional disorders (see Table 6).

**Emergency department setting**
Prospective consistency ranged from 44.4% for other specific personality disorders to 81.1% for bipolar affective disorder. The prospective consistency of the three most prevalent specific diagnoses at the first evaluation was 79.2% for schizophrenia, 81.1% for bipolar affective disorder and 62.5% for dysthymia (see Table 4). Retrospective consistency at the last evaluation ranged from 41.7% for obsessive–compulsive disorder to 80.0% for recurrent depressive disorder; it was 67.0% for schizophrenia, 70.6% for bipolar affective disorder and 69.0% for dysthymia (see Table 4).

The proportion of patients who received the same diagnosis during at least 75% of the evaluations ranged from 19.5% for residual schizophrenia to 54.6% for schizophrenia, schizotypal and delusional disorders (see Table 6).

**In-patient setting**
Prospective consistency ranged from 66.7% for recurrent depressive disorder to 100.0% for obsessive–compulsive disorder and eating disorders. The prospective consistency of the three most prevalent specific diagnoses at the first evaluation was 90.9% for schizophrenia, 91.5% for bipolar affective disorder and 81.8% for dysthymia (see Table 5). Retrospective consistency at the last evaluation was between 63.1% for specific personality disorders and 100.0% for recurrent depressive disorder and obsessive–compulsive disorder; it was 91.5% for schizophrenia, 89.3% for bipolar affective disorder and 75.0% for dysthymia (see Table 5).

The proportion of patients who received the same diagnosis during at least 75% of the evaluations ranged from 37.5% for bipolar affective disorder, current episode mild or moderate depression, to 100.0% for obsessive–compulsive disorder and other specific personality disorders (see Table 6).

## DISCUSSION

The main variable influencing diagnostic stability for the most prevalent chronic psychiatric diagnoses was the clinical setting in which the patients were assessed. The inpatient setting showed the highest diagnostic stability, followed by the emergency and outpatient settings. The temporal consistency of psychiatric disorders was lower than that found in other studies.

**Strengths and weaknesses of the study**

The main strengths of this study are the large, representative sample, the length of follow-up (up to 12 years) and the large number of evaluations. Moreover, although most previous studies focused on one psychiatric diagnosis assessed in a single clinical setting, we assessed the stability of all psychiatric diagnoses naturally presenting in clinical practice. Psychiatric diagnoses were evaluated in three different clinical settings, using the same diagnostic procedure that is used during routine clinical practice. Clinicians who assigned the diagnoses were masked to the study process. Other work has used semi-structured interviews and other diagnostic instruments not used ordinarily in clinical practice. The results of our study may more accurately reflect the real use of diagnostic classifications in psychiatric practice and may be more useful in estimating the clinical utility of current psychiatric classification systems.

Diagnostic changes over time may reflect the evolution of an illness, the emergence of new information or unreliability of measurement (Schwartz et al, 2000). Spitzer et al (1978) divided the sources of unreliability that lead to diagnostic disagreement among clinicians into categories...
(sources of variance): subject variance, occasions variance (e.g. different episodes of bipolar disorder), information variance (e.g. the differences across settings and informants), observation variance (e.g. differences among clinicians) and criterion variance. Our study has limitations that may reflect the influence of these sources of unreliability. The stability of bipolar disorder may be affected by the occasions variance, particularly the diagnostic category of bipolar affective disorder, current episode mild or moderate depression (ICD-10 F31.3). Information and observation variances can be significantly reduced by training clinicians in interviewing techniques and observational skills, and by the use of structured or semi-structured clinical interviews. Because of the naturalistic nature of our research, structured or semi-structured clinical interviews were not used in the study. This might have increased the criterion variance. The clinicians who assigned the diagnoses were not specifically trained to improve interrater reliability, which might have influenced the consistency of the analyzed diagnoses. Psychiatrists used different diagnostic classifications to code the diagnoses throughout the study period.

Other research

The stability of chronic psychiatric diagnoses has been evaluated in a number of studies (Tsuang et al, 1981; Schwartz et al, 2000; Lieb et al, 2002; Shea et al, 2002; Mojtabai et al, 2003; Barkow et al, 2004; Grilo et al, 2004; Veen et al, 2004; Culverhouse et al, 2005; Kessing, 2005a,b; McGlashan et al, 2005; Rufino et al, 2005; Schimmelmann et al, 2005). Most of these studies have focused on one diagnostic cluster, mainly psychoses (schizophrenia spectrum and mood psychoses; Schwartz et al, 2000; Mojtabai et al, 2003; Veen et al, 2004; Kessing, 2005b; Rufino et al, 2005; Schimmelmann et al, 2005) and personality disorders (Shea et al, 2002; Grilo et al, 2004; McGlashan et al, 2005). These studies usually have a small number of evaluations – two or three in most of them (Schwartz et al, 2000; Lieb et al, 2002; Barkow et al, 2004; Grilo et al, 2004; Schimmelmann et al, 2005) – and the follow-up period is usually under 3 years (Schwartz et al, 2000; Shea et al, 2002; Barkow et al, 2004; Grilo et al, 2004; Veen et al, 2004; McGlashan et al, 2005; Rufino et al, 2005; Schimmelmann et al, 2005) with a few exceptions (Tsuang et al, 1981; Lieb et al, 2002; Mojtabai et al, 2003; Culverhouse et al, 2005; Kessing, 2005a,b). Kessing (2005b) recently pointed out that no study has investigated the diagnostic stability of the most common ICD-10 psychiatric diagnoses given under ecological clinical conditions.

Other authors have reported rates of consistency that are much higher than the ones found in our study (Tsuang et al, 1981; Schwartz et al, 2000; Veen et al, 2004; Kessing, 2005b; Schimmelmann et al, 2005). However, most studies that have evaluated the stability of chronic psychiatric diagnoses have shorter follow-up periods than in our study and have focused on a single clinical setting (mainly the in-patient setting). Schwartz et al (2000) reported that rates of consistency of some diagnoses decreased as the follow-up period increased. For example, the retrospective consistency of schizophrenia was 73.1% in a comparison of 6-month and 24-month diagnoses, but fell to 55% (similar to the figure of 45.9% obtained in our study across clinical settings) when baseline and 24-month diagnoses were compared. However, the retrospective consistency of bipolar disorder remained high: 84.8% (6-month and 24-month diagnoses) and 73% (baseline and 24-month diagnoses).

Compared with the data from the study by Schwartz et al (2000), the retrospective consistency of bipolar disorder across clinical settings in our study (38.1%) is strikingly low. The third measure of stability that we calculated (the percentage of patients who received the same diagnosis in at least 75% of the evaluations) may more accurately reflect the diagnostic process through different evaluations, and was also strikingly low in our study. Some examples of low values are bipolar affective disorder (23.1%) and specific personality disorders (12.7%), whereas schizophrenia (42.4%) and eating disorders (43.9%) showed the highest rates of stability.

The very low consistency for the category 'bipolar affective disorder, current episode mild or moderate depression' may be explained by the fact that this diagnosis is inherently expected to change, since it represents an episode rather than a disorder. Perhaps the use of semi-structured interviews would have enhanced reliability and therefore stability. A structured interview, the Structured Clinical Interview for DSM-III–R was used to provide DSM-III–R psychiatric diagnoses in the study by Schwartz et al (2000).

Explanations and implications for clinicians and policy makers

There may be several explanations for the differences in diagnostic stability across clinical settings. First, it may be easier to diagnose a disorder correctly when symptom severity is at its highest, as in hospital admissions and emergency visits. We did not have data regarding illness severity; however, it would be interesting to conduct a similar study controlling for symptom severity. Second, during hospitalisations, round-the-clock surveillance and symptom observation may increase the accuracy of the diagnoses. In addition, during hospitalisations, clinicians can more easily interview the patient’s family, and there is more time for thorough diagnostic assessment and questioning about areas of functioning and symptoms. According to Spitzer et al (1978), this may contribute to information variance, and may partially explain the differences in diagnostic stability across clinical settings. Third, the duration of the follow-up period was much longer in the out-patient setting (1992–2004) than in the emergency and hospitalisation settings (2000–2004). Finally, the number of psychiatric contacts was different in each setting (data not shown). Some authors have suggested that the causall relationship between diagnostic stability and the number of psychiatric contacts is unknown:

'Patients who have many psychiatric contacts may present with more unstable psychiatric illness leading to more diagnostic variation. On the other hand, it may be that clinicians have problems with diagnosing some patients accurately and that this may lead to less effective treatment and more psychiatric contacts for these patients.' (Kessing, 2005b).

It is surprising that diagnostic stability was higher in the emergency department setting than in the out-patient setting. Other authors (Segal et al, 1995; Rufino et al, 2005) have noted that psychiatric diagnoses assigned in an emergency department may be less accurate than diagnoses assigned in other settings. In emergency department settings, time is usually limited, frequently there is no additional information from relatives, and in most cases, there is a need for immediate intervention (Segal et al, 1995; Rufino et al, 2005).

The temporal consistency of mental disorders in our study is lower than that found in other longitudinal studies. The relative lack of diagnostic stability over time is striking given that there is likely to be a bias towards maintaining the same diagnosis
over time. Psychiatrists treating the patients in this study often had access to past records and diagnoses, and may have been inclined to keep the previous diagnosis rather than assign a different one. It should be noted that the view that disorders may not be discrete ‘disease entities’ but rather dimensions of continuous variations has gained currency (Kendell & Jablensky, 2003). The categorical approach to psychiatric diagnostic classification has been criticized in favour of other classification systems, such as symptom-cluster dimensions (Kendell & Jablensky, 2003). The possibility of alternative approaches to diagnoses also raises questions about the value of diagnostic stability as an indicator of the validity of the diagnoses. Krishnan (2005) has recently stated that ‘the limits of the nominalist tradition have been reached’ and has suggested four criteria for defining disease: clinical symptoms; course and outcome; familial pattern; and treatment response.

The results of our investigation raise worrisome concerns regarding the validity of results of epidemiological, clinical and pharmacological psychiatric research, particularly in studies of chronic disorders with short follow-up periods that may not allow enough time to reach the right diagnosis or in studies that do not take setting into account. This underscores the inherent weaknesses in our diagnostic system, leading to instability of diagnoses which could reflect limitations of the nosology and result in inappropriate treatment recommendations or interventions.

Future research
It is likely that psychiatric diagnostic categories require revision. This can only be determined definitively with a large-scale study using structured or semi-structured interviews. Such a project may be feasible, but we believe that it might not accurately reflect the conditions of psychiatric practice in the real world.

REFERENCES


Using activity data to explore the influence of case-load size on care patterns

TOM BURNS, JENNY YIEND, HELEN DOLL, TOM FAHY, MATTHEW FIANDER and PETER TYRER

Background  A limited case-load size is considered crucial for some forms of intensive case management and many countries have undertaken extensive reorganisation of mental health services to achieve this. However, there has been limited empirical work to explore this specifically.

Aims  To test whether there is a discrete threshold for changes in intensive case management practice determined by case-load size.

Method  ‘Virtual’ case-load sizes were calculated for patients from their actual contacts over a 2-year period and were compared with the proportions of contacts devoted to medical and non-medical care (as a proxy for a more comprehensive service model).

Results  There were 39 025 recordings for 545 patients over 2 years, with a mean rate of contacts per full-time case manager per month of 48 (range 35–60). There was no variation in the proportion of non-medical contacts when case-load sizes were over 1:20 but there was a convincing linear relationship when sizes were between 1:10 and 1:20.

Conclusions  Case-load size between 1:10 and 1:20 does affect the practice of case management. However, there is no support for a paradigm shift in practice at a discrete level.

Declaration of interest  PT. is editor of the British Journal of Psychiatry but had no part in the evaluation of this paper for publication.

Mental health practice is increasingly driven by national policy initiatives which stipulate care structures in considerable detail (Department of Health, 1999, 2001). These prescribed service models draw on international examples of best practice (Stein & Test, 1980; Edwards et al, 2000) which have generally been associated with a range of desirable outcomes such as reduced in-patient care, reduced loss to follow-up and increased engagement. Reduced hospitalisation is the most commonly quoted outcome and the one most used for comparison of models (Marshall & Lockwood, 1998). There is a growing dissatisfaction, however, with the use of purely administrative or symptomatic outcome, particularly in long-term and disabling mental illnesses such as psychoses where there is a drive for a broader range of outcome dimensions incorporating social functioning, quality of life and satisfaction with services (Artkisson et al, 1992). This more comprehensive or ‘holistic’ approach to assessing outcomes has paralleled a call for an equally comprehensive approach to treatment, with an emphasis on the provision of a range of psychosocial interventions in addition to pharmacological treatment (National Institute for Clinical Excellence, 2002). Smaller case-loads are proposed as the foundation of this more holistic approach.

Small case-loads (e.g. 1:10, 1:12) are a feature of most current service prescriptions such as assertive outreach, crisis resolution teams and early intervention teams (Department of Health, 2001). The success of some of these service models in reducing the need for hospital care in several influential trials (Stein & Test, 1980; Hoult et al, 1983) has led to their extensive replication (Marshall & Lockwood, 1998) although not always with the same success (Thornicroft et al, 1998; Burns et al, 1999; Catry et al, 2002). Despite the absence of any evidence that very small case-load sizes themselves are closely associated with improved outcomes (as distinct from the comprehensive approach embodied in such model teams) (Wright et al, 2004), they are still strongly endorsed and precisely stipulated (Stein & Santos, 1998).

First attempts to explain the variation in outcome in these studies of ostensibly similar interventions explored the impact of varying model fidelity (McHugo et al, 1999; Fiander et al, 2003) and yielded mixed results. One criticism of the model-fidelity approach is that it focuses predominately on structural and organisational aspects of the services and less so on day-to-day practice. Assessments of model fidelity are also generally based on self-report rather than direct measurement. The one published study using prospectively collected data (Fiander et al, 2003) did not find a strong association with improved outcome. A criticism of this prospective study, which had drawn its UK data from the UK700 study (Burns et al, 1999), is that its negative result could indicate either that there was no association between the factors examined, or simply that the levels of case-loads tested were badly chosen.

The UK700 trial was the first in this field to test the impact of varying only one feature between experimental and control conditions – in this instance a comparison of case-load sizes of 1:12–15 and 1:30–35. The trial was a large multisite randomised controlled trial of case management in psychosis and failed to find any impact of case-load size on hospitalisation or clinical outcomes. It has been proposed (Gournay & Thornicroft, 2000) that the experimental case-load sizes were too high and had they been smaller, as in the original study (Stein & Test, 1980), a positive outcome would have been found.

This issue is of fundamental importance. In the absence of major differences in hospitalisation rates, case-load size is the major cost driver in such services. However, a series of adequately powered trials using differing case-load thresholds is hardly feasible. Alternative methods of identifying a critical case-load size need to be considered, either to inform service provision or as the basis for a definitive trial.

Data collected in the UK700 trial have previously been used to explore the effects of case-load size on process of care of patients with severe psychotic illness (Burns et al, 2000), with the balance of medical to non-medical interventions as a proxy.
indicator for holistic care. The proportion of non-medical contacts was only increased when rates of contact were above about one per week and medical contacts comprised the majority when frequency was less than this. As with the original UK700 trial this process of care study was limited to two pre-set case-load levels.

In the current study we test for a relationship between the balance of medical and non-medical contacts and contact frequency to explore the impact of varying case-load sizes in the community care of individuals with severe mental illness.

**METHOD**

We constructed ‘virtual’ case-load sizes for each patient based on actual contact frequency and compared this level with the proportion of contacts devoted to non-medical activities (taken to indicate that some of the goals of the new intensive service to provide more comprehensive care were being achieved).

### Constructing ‘virtual’ case-loads from service data

The UK700 study collected detailed, prospective data on staff activity and this confirmed that the two treatment arms did provide different patterns of care despite the absence of an outcome difference (Burns et al, 2000). There were a total of 39 025 recordings for 545 patients over 2 years. However, the data indicated a wide variation in the levels of activity between individual patients within each treatment group. There were some patients within the group with standard case management (case-load 1:30–35) who had more frequent contact than some patients in the intensive case management group (case-load 1:12–15). Using individual patient-level data it is possible to derive a ‘virtual case-load’ size for each patient by dividing their mean contacts per month over the 2 years of follow-up by the mean monthly contacts achieved by the average case manager.

#### Choice of service measure

The prospective service recording in the UK700 study included five categories (face-to-face contacts, telephone contacts, carer contacts, failed contacts, care coordination). The content of face-to-face contacts was classified into 11 event types based on the focus of therapeutic activity (housing, occupation and leisure, finance, daily living skills, criminal justice system, carer issues, engagement, physical health, specific medical intervention/assessment, medication, case conference). These were derived using a modified Delphi approach to achieving consensus with clinicians (Burns et al, 2000). Activity rates for each category were calculated per patient per 30 days for the 2 years of the study.

We chose face-to-face contact as the service measure to construct ‘virtual’ case-loads. This measure was responsible for over 80% of all recorded activities and was the most consistently recorded across the sites. Face-to-face contacts were also the only service category where the focus of the event was recorded.

### Calculation of case manager activity

Not all case managers were full-time and some also dedicated time to patients not in the study. In order to calculate the ‘virtual’ case-load it is necessary first to decide the routine number of contacts per week or month made by an average full-time member of staff. Information on this fundamental aspect of community mental healthcare is surprisingly hard to obtain. Two local surveys of contact frequency yielded levels that were considerably lower than expected (Greenwood et al., 2000; Kent et al., 2003).

#### Development of a proxy for change in clinical practice

In the previous study (Burns et al, 2000) the proportion of ‘medical’ contacts (where the focus was either ‘medication’ or ‘specific medical intervention/assessment’) to ‘non-medical’ contacts (the focus was any of the other nine categories listed previously). We have used the same proxy measure in this study.

### Statistical analyses

To generate graphical representations patients were categorised according to their notional allocation to intensive case or standard case management as determined by study design. Calculated (‘virtual’) case-loads were categorised by dividing consecutive values into 13 samples of equal sizes that reflected differing case-load ranges. Proportions of patients in various categories were compared using χ² tests.

Correlations were assessed using Spearman’s method owing to non-normality of the distributions. Stepwise linear regression was used to assess relationships between model of care, calculated case-loads and proportion of non-medical contacts. The proportion of non-medical contacts was

---

**Table 1** Contact frequencies for patients in the intensive and standard case management groups of the UK700 trial

<table>
<thead>
<tr>
<th></th>
<th>Intensive case management</th>
<th>Standard case management</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>St George’s</td>
<td>St Mary’s</td>
<td>Kings</td>
</tr>
<tr>
<td>Nominal case-load</td>
<td>12</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Total patients</td>
<td>97</td>
<td>98</td>
<td>77</td>
</tr>
<tr>
<td>Total staff</td>
<td>8</td>
<td>6.5</td>
<td>5</td>
</tr>
<tr>
<td>Total face-to-face contacts (per 30 days)</td>
<td>476</td>
<td>230</td>
<td>206</td>
</tr>
<tr>
<td>Mean face-to-face contacts per staff member per 30 days</td>
<td>60</td>
<td>35</td>
<td>41</td>
</tr>
</tbody>
</table>
the dependent factor, with care model (intensive or standard management) and calculated case-load entered as fixed factors. Stepwise linear regression was used to assess the effect of the calculated case-load size on primary and secondary outcomes, controlling for baseline variables (as specified in the original UK700 report) and baseline levels of the tested outcome variable.

RESULTS

Virtual case-load size

Figure 1 shows the distribution of patients according to their notional allocation (either intensive or standard case management) and their calculated ‘virtual’ case-loads. Based on recorded activity the average case-loads were 14 and 33 for intensive and standard case management respectively.

Most patients in the standard group were receiving the levels of care that would be predicted, with only 2% (6 of 267) receiving care equivalent to a case-load of 15 or under. However, only 40% of patients in the intensive management group were receiving care equivalent to a case-load of 15 and under, and 21% (57 of 272) were receiving care equivalent to a case-load size of 30 and above. The difference in the distribution is highly statistically significant ($\chi^2 = 113, P < 0.0001$), suggesting that patients in the two treatment groups really did receive distinctly different services.

‘Virtual’ case-load size and non-medical contacts

Figure 2a–c shows scatterplots of ‘virtual’ case-load in relation to proportion of non-medical contacts. Estimated case-load sizes are limited to 1:100 (because some patients could only be contacted once or twice during the 2 years they generate spuriously high virtual case-load sizes). Spearman’s correlation demonstrates a small but statistically significant negative relationship between virtual case-load size and the proportion of non-medical contacts ($r = -0.138, P < 0.005$, two-tailed). Separate analyses showed a significant relationship for the group with intensive case management ($r = -0.231, P < 0.001$) but not for the standard management group ($r = 0.108, P = 0.1$). However, linear regression analysis with the proportion of non-medical contacts as the dependent variable and care model and grouped virtual case-load size as fixed factors revealed no significant interaction term (care model $\times$ virtual case-load size, $P > 1$).

Figure 3 presents the mean proportion of non-medical contacts according to ‘virtual’ case-load size. The range of these steps is unequal as comparative numbers of results in each bin are required for analysis. Analysis by each individual case-load size (e.g. 10, 11, 12) was not possible because of empty cells. There was a steady increase in the proportion of non-medical contacts as case-load sizes fell from 1:19–21 to 1:9–11. The proportion of non-medical contacts was around 50% for case-load sizes below 9. The proportion of non-medical contacts varied in a rather irregular manner for case-load sizes between 1:22 and 1:34 and for sizes of 1:35 and above the proportion remained essentially stable.

Case-load size and patient outcomes

The outcomes tested were the same as in the original UK700 study – days in hospital (primary outcome) and psychiatric symptoms (Comprehensive Psychiatric Rating Scale (CPRS; Asberg et al., 1978); an adapted form of the Disability Assessment Schedule (DAS; World Health Organization, 1998)); quality of life (Lancashire Quality of Life Profile; Oliver et al., 1997); and patients’ satisfaction (Camberwell Assessment of Need; Phelan et al., 1995) (secondary outcomes). Analyses were adjusted for baseline levels of the corresponding outcome variable and for other baseline variables (e.g. age, months since onset) as in the original report (Burns et al., 1998). Results showed no significant relationship between ‘virtual’ case-load size and primary outcome. One secondary outcome, DAS score, was significantly predicted by ‘virtual’ case-load size ($\beta = -0.086$, $P < 0.005$). Larger case-loads predicted an average decrease in social disability.


**DISCUSSION**

Given the controversy that the UK700 study generated (Gournay & Thornicroft, 2000; Smyth & Hoult, 2000) and the emphasis placed on case-load sizes by commissioners and policy makers it is remarkable how little research has been conducted into the effects of varying case-load size. In the UK case-load sizes have been explicitly prescribed and linked to funding for all the new teams recommended in the NHS Plan (assertive outreach, crisis resolution/home treatment, first-onset) (Department of Health, 2000, 2002). Similarly in the USA, Canada, several European countries and Australia adherence to case-load sizes is a requirement for the funding of specialised mental health teams. For commissioners the issue is decisive as case-load size, after duration of in-patient stay, is the major determinant of the cost of mental healthcare.

**Assertive community treatment**

The insistence on an absolute threshold for case-loads reflects a consistently expressed belief that there is a qualitative shift in practice – that the assertive community treatment model is ‘all or nothing’ (Allness & Knoedler, 1998). This insistence drew its legitimacy from the series of studies indicating that assertive community treatment teams were routinely associated with a reduction in bed usage (Marshall & Lockwood, 1998). However there have been important service changes in mental healthcare in the USA over the past two decades which have involved more actively managed in-patient care and the development of a clearer community focus. These have led to a marked decrease in the potential for reduction in bed usage as a consequence of assertive community treatment and few modern studies can hope to achieve the dramatic reductions found by Stein & Test (1980) or Rosenheck et al (1995). Essock and colleagues (2006) recently failed to demonstrate a significant overall reduction in hospitalisation when comparing assertive community treatment with standard case management in two urban populations of American patients with mental illness complicated by unstable housing and substance misuse. Overall, patients in both groups improved but the rate of institutionalisation, reflecting the European experience (Burns et al, 2002).

**Models of care**

However, there is evidence that resource enhancement alone may fail to change practice without an explicit change in model of care. Kent et al (2003) found no increase in psychosocial interventions used by community mental health teams who had expressed a wish to do so despite the provision of substantial extra clinical time. The impact of these findings is limited, however, by the absence of evidence for an optimal, or critical, case-load size. It could be argued that the teams studied by Kent et al (2003) were so underresourced that their enhancement only permitted adequate medical-model care to all patients or, conversely, that they were already sufficiently resourced, the extra clinical time was not needed and the level of non-medical care had been clinically appropriate. This is similar to the criticism of the UK700 trial – that both arms of the trial lay on one side of this crucial threshold.

**Main results**

The contact frequencies reported in this trial are lower than many clinicians would have expected or wished and there is a clear difference in frequency between sites. However, there is no published evidence that they are lower overall than frequencies in previously reported studies and there is some evidence that they broadly reflect clinical practice in these teams (Fiander & Lockwood, 1998). Why is there such a range of contact frequency in similarly staffed teams is an interesting question and one for which carefully targeted studies will be needed (Weaver et al, 2003). It is, however, beyond the scope of this paper.

Our results give little support for the importance of a clear-cut and crucial case-load threshold to dismiss the findings of the UK700 study. Figure 3 does not demonstrate a step-wise change in practice at any case-load size, but rather a dose-response curve between case-load sizes of 1:10 and 1:20. Thus the patients in these ‘virtual’ case-loads appeared to receive steadily increasing non-medical (taken here to indicate comprehensive) care as the case-load fell. This would support the value of small case-loads (i.e. below 1:20) for the community care of individuals with severe psychotic illnesses. The ‘dose-response’ character indicates how clinicians may be able to use extra contact time creatively. However, the argument for smaller case-loads must rest on what is going to be delivered in terms of treatments – there is no support for the idea that a certain case-load threshold triggers a quite different way of working.

Interpreting the results for case-loads above 1:21 or below 1:9 is difficult. Above 1:33 the curve is essentially flat and there is no identifiable influence of case-load size, with two-thirds of contacts being explicitly medical. However, these larger ‘virtual’ case-loads reflect increased difficulties in maintaining contact with patients rather than planned clinical activity – what contact could be achieved, not what was considered appropriate. Limitations of the data and statistical methodology prevent us from further testing of case-loads below 1:9.

The range of case-load sizes between 1:21 and 1:35 contains an uncertain mixture of patients receiving intensive and standard case management and shows no simple consistent trend. It is difficult, and probably unwise, to try to draw conclusions from these results. Our scatterplots further support this interpretation that it is only with small case-loads that this shift in the balance of activity is demonstrated. The weak association found in the scatterplot for all patients is entirely accounted for by patients receiving intensive case management.

**Case-load threshold**

Burns et al (2000) found no difference in the mean number of medical contacts per patient per 30 days between teams with case-load sizes of 1:12 and 1:15. The difference between the teams was that the team with a case-load of 1:12 was using most of their ‘extra’ contacts for non-medical activity. Burns et al speculated that teams might be prioritising medical contacts, that there could be a clinically determined ‘ceiling’ for such contacts in this patient group and that once this level (approximating to 1 visit per 3 weeks) was reached all further activity would be devoted to a broader range of non-medical interventions.

Our current findings do not support such a ‘ceiling’ effect for medical contacts. When the proportions of medical contacts at the different ‘virtual’ case-load sizes were translated into absolute frequencies they rose steadily across the range. At case-load sizes of 36–44 a mean of 0.78 medical
contacts were made per patient per 30 days; case-loads of 30–35 yielded 1.1 medical contacts, at 19–21 the frequency was 1.85 and by 9–11 it had risen to 2.6 per 30 days.

However, our findings should not be taken as a rejection of the importance of a fixed case-load. The emphasis placed on case-load size by assertive teams may be more related to the need for greater autonomy and an internal locus of control for the team than for perceived fidelity to the assertive approach. One of the attractions of working in an assertive outreach team is the guarantee of a limited case-load. Control over case-load size has been associated with less burnout in personnel compared with equivalent staff in community mental health teams where case-load sizes are bigger (Billings et al, 2003). Greater latitude in decision-making and lower job demands have also been associated with higher levels of job satisfaction and performance (Evans et al, 2006). By setting a limit to case-load size this control can be exercised unambiguously and transparently. What that limit needs to be remains, however, open to local consideration based on the clinical goals of the team and local needs and services.

Limitations

There are a number of obvious limitations to this exploratory study. We report here analyses of data collected from a study designed to answer a different question. The most severe limitation is that this study is built on two artificially constructed proxies – a ‘virtual’ case-load derived from contact frequency and a rough measure of comprehensive care based on the proportion of ‘medical’ and ‘non-medical’ activities. The problem for the ‘virtual’ case-loads is that they were not predetermined and reflect clinical need. Any conclusions about causality (i.e. that small case-loads are responsible for, rather than associated with, a more comprehensive approach) can only be speculative.

Both of these measures are based on self-report by case managers. Although extensive verifications of contact frequency were conducted in the original study (Burns et al, 2000), no audits of activity or reliability exercises were conducted into the allocation of contacts to medical and non-medical categories other than to check that visits at which depot medication was administered were classified as medical.

Conclusions

Our study does not support a threshold effect for a case-load size which significantly alters clinical practice but confirms that distinctions between types of community services for this patient group (e.g. assertive community treatment, intensive case management, ‘standard’ case management) are more likely to be differences of degree than of fundamentally different practices (Catty et al, 2002). Case-load sizes vary but generally sizes of 1:20 and below seem to be characteristic of sustained intensive care in this patient group (Wright et al, 2004). Our study indicates a ‘dose response’ within this range.

The UK700 study concluded with a request for less attention to precise definitions of care structures and more focus on the content of care (Burns et al, 1999). There has, however, been very little empirical investigation of what a smaller case-load would permit that a larger one would not. Presumably this is because it is considered self-evident – more care, higher quality care, a broader range of care. Weaver’s qualitative approach to understanding the possible mechanisms of the impact of smaller case-loads on the process of care is a notable exception (Weaver et al, 2003). Our findings should alert researchers, clinicians and policy makers to the need for a careful critical approach to interpreting health service trials of complex mental health interventions. How extra resources are used is more important than how it is organised.

REFERENCES


Patients discharged from medium secure forensic psychiatry services: reconvictions and risk factors

JEREMY COID, NICOLE HICKEY, NADJI KAHTAN, TIANQIANG ZHANG and MIN YANG

Background  Treatment within medium secure forensic psychiatry services is expected to reduce risk to the public.

Aims  To measure the period prevalence and incidence of offending following discharge and identify associated risk factors.

Method  Follow-up of patients from 7 of 14 regional services in England and Wales who spent time at risk (n=1344) for a mean of 6.2 years. Outcome was obtained from offenders index, hospital case-files and the central register of deaths.

Results  One in 8 men and 1 in 16 women were convicted of grave offences. Incidence rates indicated low density and most patients were not subsequently convicted. Offence predictors included gender, younger age, early-onset offending, previous convictions and a comorbid or primary diagnosis of personality disorder. Longer in-patient stay and restriction on discharge were protective.

Conclusions  Risks of reoffending remain for a subgroup of discharged patients. Future research should aim to improve their identification and risk management following discharge.

Declaration of interest  None.

Concern over public safety has resulted in proposals for new services and new mental health legislation for high-risk psychiatric patients (Home Office, 1999; Department of Health & Home Office, 2001; National Institute for Mental Health in England, 2003; Department of Health, 2004), with the requirement that health services work with the criminal justice system to reduce reoffending (Home Office, 1998). Services for offender patients in the UK are the outcome of earlier recommendations (Butler Committee, 1975) for a network of regional secure units at a medium level of security between ordinary psychiatric hospitals and the special (maximum security) hospitals. Although subsequent service development has been uneven (Coid et al, 2001), all health regions have now provided these services, in which treatments are expected to reduce dangerousness of patients. It has been argued that criminal recidivism is of greater importance when assessing clinical effectiveness than clinical relapse (MacCulloch & Bailey, 1991), although this emphasis has been disputed (Robertson, 1989; Friendship et al, 1999).

Previous follow-up studies of forensic patients in the UK have limitations. Reports on those discharged from special hospitals to psychiatric hospitals or the community include samples discharged more than 20 years ago (Tong & Mackay, 1959; Gathercole et al, 1968; Acres, 1975; Black, 1982; Tennent & Way, 1984; Bailey & MacCulloch, 1992; Buchanan, 1998; Jamieson & Taylor, 2004). Most now undergo rehabilitation and gradual community leave through the medium secure services (Coid & Kahtan, 2000). Furthermore, apparent improvement in rates of reoffending over time (Buchanan, 1998) may result from changing populations, specifically a decline in admissions of those with a primary diagnosis of personality disorder (Coid et al, 1999), rather than improved after-care. Follow-up studies of patients discharged from medium secure services have been limited by small numbers and unrepresentative samples (McMurran et al, 1998; Baxter et al, 1999; Falla et al, 2000): selection of patients with a single diagnosis (Baxter et al, 1999; Halstead et al, 2001); restriction to a single unit or small geographical area (Baxter et al, 1999; Friendship et al, 1999; Maden et al, 1999; Castro et al, 2002; Edwards et al, 2002); or follow-up over a 2-year period (Maden et al, 2004). None has used directly comparable outcome measures of re offending, or controlled for time at risk.

Large studies are needed which include all subgroups of patients and have sufficient statistical power to quantify the long-term risks of reoffending following psychiatric treatment in medium secure services, and identify those posing the highest risk to public safety. We followed a large, nationally representative sample of patients discharged from medium secure units to the community to examine the incidence of reoffending, to identify risk factors for reoffending and to explore the implications for future risk management.

METHOD

Patients  Patients were included who had been admitted to medium secure forensic psychiatry services in 7 of the 14 (prior to reorganisation) regional health authorities between 1989 and 1993. These form a representative range of geographical areas, including large urban, small town and rural areas, characterised by different levels of socio-economic deprivation. This was an original admission cohort from the North West Thames, North East Thames, South Western, West Midlands, Merseyside, North Western and East Anglian Regional Health Authority catchment areas and is described in previous publications (Coid & Kahtan, 2000). Patients admitted to these services during the study period, but placed in private sector or other National Health Service (NHS) secure units as extra-contractual referrals, were included so as not to underrepresent the catchment areas.

The follow-up period was calculated from date of discharge to the end of the study period (31 December 1998), or date of death or leaving the country, whichever occurred first. Time at risk of reconviction was defined as any time spent in the community during the follow-up period. The
original admission cohort consisted of 2085 patients over the 5-year period. A total of 472 (23%) were excluded from follow-up because hospital case-files were unavailable or there was insufficient information to complete coding schedules. Subsequent comparison revealed no statistically significant differences between this group and those included according to demography, previous convictions, previous hospitalisation for psychiatric illness and age at admission to medium secure services. However, significantly more excluded patients were admitted because of non-criminalised behaviour and detention under a civil order of the Mental Health Act 1983, and were admitted from a psychiatric hospital or directly from the community. A further 269 patients (13%) were excluded as they did not enter the community during the follow-up period, and therefore did not enter a period of ‘time at risk’ of reconviction in the community.

Patients initially transferred from medium secure services to a local psychiatric hospital were only considered to enter ‘time at risk’ once they had been discharged to the community. Those who died during the follow-up period but who had spent some time at risk were included.

The project was approved by the East London and City Health Authority Ethics Committee.

Data sources
Data for each patient were obtained from a range of sources and different sites. Medical records files from the medium secure units were examined in the medical records office at each location. These included pre-admission psychiatric reports, case conference reports, social histories, general correspondence and discharge summaries. The Mental Health Unit at the Home Office, which is responsible for monitoring the progress of patients subject to restriction orders under sections 41 and 49 of the Mental Health Act 1983, also gave access to their files. Discharge under restrictions (section 37/41) was included as a risk factor in our analysis. The medical records departments in all relevant general psychiatric hospitals, including out-patient departments, and special hospitals were also contacted for information on any in-patient and out-patient contacts after discharge from medium secure services.

Lifetime diagnostic data on categories of mental illness were included and assessed from case notes by a trained psychiatrist (N.K.) using ICD–10 criteria (World Health Organization, 1992). Personality disorder was also included but sub-categories were considered to be infrequently and inaccurately specified in case notes, therefore the researcher made a diagnostic decision based on available information using DSM–III-R Axis II criteria (American Psychiatric Association, 1987). Comorbid diagnoses of lifetime alcoholism and alcohol misuse, drug dependence and drug misuse, and sexual deviation were obtained from case notes. Categories of mental disorder included in the analysis described the primary psychopathology and included mutually exclusive categories of schizophrenia or schizoaffective disorder, delusional disorder, personality disorder, mania or hypomania, depression and organic brain syndrome. Comorbid categories included alcoholism/alcohol misuse and drug dependence/drug misuse. Anti-social personality disorder could be a primary diagnosis within the category of personality disorder or a comorbid diagnosis with other conditions.

Outcome data
The Offenders Index at the Home Office provided outcome data on convictions for standard list offences in England and Wales up to the end of the study period (31 December 1998). For the purposes of analysis, outcome measures included offences of violence against the person; sexual offences; arson; acquisitive offences of burglary, theft, fraud and deception, and robbery; and any conviction for ‘grave’ offences. The Home Office defines ‘grave’ offences as homicide, serious wounding, rape, buggery, arson, robbery and aggravated burglary. The NHS Central Register, which is administered by the Office for National Statistics was searched to determine whether any people who had not been traced at the end of the follow-up period had died.

Statistical analysis
Prevalence of conviction by offence category over the follow-up period was calculated using descriptive statistics. Incidence rates, based on the number of reconvictions and total person-years of time at risk, were also calculated. Confidence intervals of incidence rates and incidence rate ratios between men and women were estimated using Stata version 7 for Windows based on Poisson distribution. Offence-specific incidence rates and their confidence intervals were calculated to show the degree of risk for specific offences. Survival curves plotted for first reconviction for each offence type and based on time at risk were estimated using SPSS version 11 for Windows. Cox regression models for each type of offence were fitted separately to estimate the hazard rates for associated risk factors.

RESULTS
The follow-up period was a mean of 6.2 years (s.d.=2.1) with a range of less than a month to 9.9 years. Of the 1613 patients in the original admission cohort, 1344 (83.3%) spent some time at risk and were therefore included in the subsequent analyses. Most were men, a large proportion were Black or from minority ethnic groups, were not born in the UK, had a diagnosis of psychotic illness with a comorbid lifetime history of substance misuse or dependence, and 1 in 5 had anti-social personality disorder (Table 1). Most were detained under the legal category ‘mental illness’ of the Mental Health Act 1983. A large subgroup had no previous convictions.

More than a third of men and nearly 1 in 7 women were convicted of a criminal offence during the follow-up period, more than 1 in 6 men, but only 1 in 20 women, for violence against the person (Table 2). Nearly 1 in 8 men and 1 in 16 women were convicted of a grave offence. Few people were subsequently convicted of sexual offences or arson. However, the true risk of any conviction following release was 46.8 and 16.3 offences per 100 patients discharged per year among men and women respectively (Table 2). Hazard rates were much lower for violent and grave offences, with very low risks for sexual and arson offending. Table 2 shows that incidence rates of subsequent conviction were significantly higher among men for all offence categories except arson. However, incidence rate ratios demonstrated that men were no more likely than women to be convicted of grave offences, and women were significantly more likely to be convicted of arson.

The subgroup of 250 patients admitted as a result of non-criminalised behavioural disorder had much lower rates of conviction than those admitted following criminal behaviour, with a prevalence of 20.8% and an incidence rate of 16.9 (95% CI
9.9 years. 9.9 years.
criminalised sexual behaviour and arson. criminalised sexual behaviour and arson.
or arson following admission for non-or arson following admission for non-
lowing non-criminalised violent behaviour, lowing non-criminalised violent behaviour,
for other offences. For those admitted fol-for other offences. For those admitted fol-
grave offences; and 15.6%, 8.6 (7.0–10.0) for grave offences; and 15.6%, 8.6 (7.0–10.0)
0.2) for arson; 7.2%, 1.9 (1.2–2.7) for arson; 7.2%, 1.9 (1.2–2.7) for
for sexual offences; 9.6%, 4.1 (3.0–5.1) for sexual offences; 9.6%, 4.1 (3.0–5.1)
(2.7–4.7) for violence; 1.6%, 0.5 (0.1–0.9) for violence; 1.6%, 0.5 (0.1–0.9)

Table 1  Characteristics of the follow-up sample  
(n=1344)1

<table>
<thead>
<tr>
<th>Demographic variables, n, (%)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>1167 (86.6)</td>
</tr>
<tr>
<td>White</td>
<td>1033 (76.9)</td>
</tr>
<tr>
<td>African–Caribbean</td>
<td>218 (16.2)</td>
</tr>
<tr>
<td>South-Asian</td>
<td>33 (2.5)</td>
</tr>
<tr>
<td>Other</td>
<td>60 (4.5)</td>
</tr>
<tr>
<td>Born outside UK</td>
<td>192 (14.3)</td>
</tr>
<tr>
<td>Primary diagnosis, n (%)</td>
<td></td>
</tr>
<tr>
<td>Schizophrenia/schizoaffective</td>
<td>761 (59.5)</td>
</tr>
<tr>
<td>Delusional disorder</td>
<td>72 (5.6)</td>
</tr>
<tr>
<td>Mania/hypomania</td>
<td>104 (8.1)</td>
</tr>
<tr>
<td>Depression</td>
<td>92 (7.2)</td>
</tr>
<tr>
<td>Organic brain syndrome</td>
<td>63 (4.9)</td>
</tr>
<tr>
<td>Personality disorder</td>
<td>188 (14.0)</td>
</tr>
<tr>
<td>Comorbid diagnosis, n (%)</td>
<td></td>
</tr>
<tr>
<td>Alcoholism/alcohol misuse</td>
<td>325 (24.3)</td>
</tr>
<tr>
<td>Drug dependence/misuse</td>
<td>378 (28.3)</td>
</tr>
<tr>
<td>Antisocial personality disorder</td>
<td>302 (22.6)</td>
</tr>
<tr>
<td>Sexual deviation</td>
<td>34 (2.5)</td>
</tr>
</tbody>
</table>

Category under Mental Health Act
1983, n (%)
| Mental illness              | 957 (71.2) |
| Psychopathic disorder       | 74 (5.5) |
| Mental illness and psychopathy | 14 (1.0) |
| Mental impairment           | 9 (0.6) |
| Other                       | 121 (9.0) |
| Not applicable              | 164 (12.2) |

Admission
Age, years: mean (s.d.) range 31.6 (10.1) 16–81
Previous convictions, n:
mean (s.d) range 8 (11) 1–114
No previous conviction, n (%) 397 (29.5)
Non-crime admission, n (%) 250 (18.6)
Stay in medium secure unit, years: mean (s.d) range 0.8 (1.1) 0.01–9.6

1. Mean follow-up 6.2 years (s.d.--2.1), range 1 month to 99 years.

Furthermore, their cumulative risk of reoffending increased over time. Similar patterns were demonstrated for grave offending. A further increase in risk of grave reoffending emerged at 8 years post-discharge from medium secure services among those originally admitted for a grave offence.

Admission to medium secure services following either an index offence or previous convictions for sexual or acquisitive offences also substantially increased the probability of reconviction for similar offences. The risks were greatest for those with the same index offence. However, the probability of subsequent convictions for sexual offences or arson remained relatively low for each of the three groups.

Risk factors
Table 3 shows the independent risk factors for the range of convictions following discharge. The risk of conviction for violence against the person was increased among

14.7–19.0) for any offending; 11.6%, 3.7 (2.7–4.7) for violence; 1.6%, 0.5 (0.1–0.9) for sexual offences; 9.6%, 4.1 (3.0–5.1) for acquisitive offences; 0.4%, 0.07 (0.0–0.2) for arson; 7.2%, 1.9 (1.2–2.7) for grave offences; and 15.6%, 8.6 (7.0–10.0) for other offences. For those admitted following non-criminalised violent behaviour, the prevalence for a subsequent conviction for violence was 13.2%, incidence rate 4.3% (95% CI 3.0–5.6). There were no subsequent convictions for sexual offences or arson following admission for non-criminalised sexual behaviour and arson.

Fig. 1  Cumulative probability of conviction post-discharge according to offence. Sexual offence is for men only. —→, All patients; —, index offence of same category; —, pre-admission convictions for the same offence.
Table 2  Gender difference in incidence of reconviction according to type of offence

<table>
<thead>
<tr>
<th>Type of offence</th>
<th>Men (n = 1167)</th>
<th>Women (n = 177)</th>
<th>Incidence rate ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patients convicted, n</td>
<td>Prevalence, %</td>
<td>Offences, n</td>
</tr>
<tr>
<td>Violence¹</td>
<td>211</td>
<td>18.1</td>
<td>437</td>
</tr>
<tr>
<td>Acquisitive²</td>
<td>232</td>
<td>19.9</td>
<td>1347</td>
</tr>
<tr>
<td>Grave³</td>
<td>142</td>
<td>12.2</td>
<td>248</td>
</tr>
<tr>
<td>Sexual</td>
<td>27</td>
<td>2.3</td>
<td>44</td>
</tr>
<tr>
<td>Arson</td>
<td>15</td>
<td>1.3</td>
<td>17</td>
</tr>
<tr>
<td>Other⁴</td>
<td>251</td>
<td>21.5</td>
<td>918</td>
</tr>
<tr>
<td>Any</td>
<td>400</td>
<td>34.3</td>
<td>2764</td>
</tr>
</tbody>
</table>

1. Total person-years at risk = 5901.
2. Total person-years at risk = 925.
3. Number of offences/number patients convicted.
4. Number of offences/total person years at risk/100.
5. Includes conviction for homicide, attempted murder, threat/conspiracy to murder, wounding, malicious wounding, assaults and weapon offences.
6. Includes conviction for robbery, burglary, aggravated burglary, theft, fraud and forgery.
7. Includes conviction for homicide, attempted murder, wounding, malicious wounding, robbery, aggravated burglary, rape and arson.
8. Includes conviction for drug offence, criminal damage, absconding, breach, firearms, abduction and other offences.

men, younger patients, Black patients and those from other minority ethnic groups, those younger when first appearing in court and those with a higher number of previous convictions for violence. The risk of violent convictions was also increased among patients with a primary diagnosis of personality disorder, those with a primary or comorbid diagnosis of antisocial personality disorder, and those originally admitted under the legal category 'psychopathic disorder'. Risk of violent conviction was reduced among those who had stayed 2 years or more in medium secure services.

Risk of sexual reoffending was substantially increased among patients with primary diagnoses of affective disorder and those with comorbid diagnoses of sexual deviation. Risks were also increased for those of younger age, who were younger when first in court, and were from 'other' ethnic subgroups. Moreover, risks progressively increased the higher the number of previous sexual convictions. Risks of subsequent convictions for arson were increased among women patients, those with one or more previous arson convictions and those with a history of alcohol dependence/alcohol misuse.

Risk of acquisitive offending was increased among younger patients, among male patients, among those younger when first in court, among those with a primary diagnosis of personality disorder, and among those detained under the legal category 'psychopathic disorder'. Risk of acquisitive convictions progressively increased the higher the number of previous acquisitive convictions, and were reduced among those who had spent 2 years or more in medium secure services. Previous substance misuse, antisocial personality disorder and ethnicity were not predictive of subsequent acquisitive convictions in this sample.

There were no differences between men and women in their risk of convictions for grave offences following discharge. However, younger patients, those younger when first in court, Black patients and those with a primary diagnosis of personality disorder demonstrated increased risks. Risk of grave convictions progressively increased the higher the number of previous convictions for grave offences. Discharge subject to section 37/41 restrictions reduced the risk of subsequent grave offending. There were no independent associations between risk of grave offences following discharge and length of stay in medium secure services, substance misuse or dependence, or antisocial personality disorder.

**DISCUSSION**

**Level of risk**

The acceptability of the risk of subsequent offending posed by patients discharged from medium secure services will ultimately be determined by the public and policy makers. Our findings indicate that these patients continued to present risks, with over a third of men receiving subsequent convictions, nearly 1 in 5 for violence. Nevertheless, despite a true risk of 47 offences per year for every 100 male patients discharged, only 7 of these were violent offences, of varying levels of severity, in a population originally admitted for violent and criminal behaviour as a result of mental disorder. Whether a lower hazard rate of 4 serious or 'grave' offences per year for every 100 patients (men and women) discharged is acceptable would also require a consensus view. However, it is important when considering these findings that they are not perceived as a measure of the performance of medium secure services but the criminal careers of patients discharged from these services and their risks of reoffending. Furthermore, it is questionable whether treatment in these services had a bearing on offending several years after discharge.

Our findings indicate that in-patient treatment programmes and subsequent supervision following discharge should be better targeted at preventing similar reoffending by identifying those at highest risk of recidivism. Those with previous convictions for violence, arson and grave offences were clearly at greatest risk of reconviction for these offences post-discharge. However, violent offending appeared for the first time in a subgroup post-discharge, indicating particular difficulties in accurate prediction of future violence among some patients admitted to medium secure services. On the other hand, this subgroup may have demonstrated previous non-criminalised violence that was
Table 3  Time to first reconviction after discharge according to prognostic risk factors from Cox regression analysis

<table>
<thead>
<tr>
<th>Prognostic factor</th>
<th>n</th>
<th>Violence</th>
<th>Sexual offence</th>
<th>Acquisitive offence</th>
<th>Arson</th>
<th>Grave offence</th>
<th>Any</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>1167</td>
<td>3.8 (1.6–8.7)</td>
<td>NA</td>
<td>2.6 (1.3–5.2)</td>
<td>0.22 (0.07–0.70)</td>
<td>1.3 (0.63–2.5)</td>
<td>2.3 (1.4–3.7)</td>
</tr>
<tr>
<td>Age</td>
<td>0.96 (0.93–0.99)</td>
<td>0.94 (0.87–1.0)</td>
<td>0.94 (0.92–0.97)</td>
<td>0.98 (0.90–1.1)</td>
<td>0.95 (0.92–0.99)</td>
<td>0.96 (0.94–0.98)</td>
<td></td>
</tr>
<tr>
<td>Born in UK</td>
<td>1152</td>
<td>1.3 (0.70–2.3)</td>
<td>2.1 (0.25–18.2)</td>
<td>1.1 (0.56–1.8)</td>
<td>0.0 (0.0)</td>
<td>1.0 (0.51–2.1)</td>
<td>1.1 (0.72–1.6)</td>
</tr>
<tr>
<td>Ethnicity v. White</td>
<td>1033</td>
<td>1.0 (0.72–1.6)</td>
<td>1.0 (0.72–1.6)</td>
<td>1.0 (0.72–1.6)</td>
<td>1.0 (0.72–1.6)</td>
<td>1.0 (0.72–1.6)</td>
<td></td>
</tr>
<tr>
<td>Previous convictions of the same category v. none 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.1 (0.69–1.7)</td>
<td>7.2 (2.3–23.9)</td>
<td>0.89 (0.35–2.3)</td>
<td>15.4 (4.5–52.4)</td>
<td>2.2 (1.3–3.6)</td>
<td>2.4 (1.4–3.9)</td>
<td></td>
</tr>
<tr>
<td>2–5</td>
<td>1.8 (1.2–2.6)</td>
<td>10.3 (1.9–56.8)</td>
<td>1.9 (1.0–3.6)</td>
<td>6.9 (1.6–29.6)</td>
<td>2.6 (1.6–4.2)</td>
<td>1.5 (0.94–2.4)</td>
<td></td>
</tr>
<tr>
<td>≥ 6</td>
<td>2.6 (1.6–4.5)</td>
<td>0.0 (0.0)</td>
<td>6.1 (3.6–10.6)</td>
<td>0.0 (0.0)</td>
<td>3.1 (1.4–6.6)</td>
<td>3.1 (2.1–4.6)</td>
<td></td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>551</td>
<td>1.1 (0.76–1.5)</td>
<td>0.59 (0.23–1.6)</td>
<td>1.2 (0.84–1.6)</td>
<td>3.6 (1.2–11.0)</td>
<td>1.3 (0.83–1.9)</td>
<td>1.3 (1.0–1.6)</td>
</tr>
<tr>
<td>Drug dependence</td>
<td>712</td>
<td>1.3 (0.91–1.9)</td>
<td>0.71 (0.27–1.8)</td>
<td>1.3 (0.91–1.9)</td>
<td>1.1 (0.36–3.1)</td>
<td>1.4 (0.91–2.3)</td>
<td>1.3 (0.99–1.7)</td>
</tr>
<tr>
<td>Primary diagnosis v. schizophrenia/schizoaffective disorder</td>
<td>761</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Personality disorder</td>
<td>188</td>
<td>2.4 (1.6–3.6)</td>
<td>3.0 (0.90–10.3)</td>
<td>2.4 (1.7–3.5)</td>
<td>2.0 (0.65–6.2)</td>
<td>1.7 (1.0–2.9)</td>
<td>2.6 (2.0–3.6)</td>
</tr>
<tr>
<td>Delusional disorder</td>
<td>72</td>
<td>1.5 (0.83–2.8)</td>
<td>0.0 (0.0)</td>
<td>1.0 (0.53–2.0)</td>
<td>0.0 (0.0)</td>
<td>0.81 (0.32–2.0)</td>
<td>1.3 (0.81–2.1)</td>
</tr>
<tr>
<td>Mania/hypomania</td>
<td>104</td>
<td>1.7 (0.99–2.9)</td>
<td>5.5 (1.4–22.6)</td>
<td>1.5 (0.85–2.8)</td>
<td>0.0 (0.0)</td>
<td>1.5 (0.79–2.9)</td>
<td>1.9 (1.3–2.8)</td>
</tr>
<tr>
<td>Depression</td>
<td>92</td>
<td>0.68 (0.24–1.9)</td>
<td>8.2 (1.7–39.7)</td>
<td>1.3 (0.63–2.8)</td>
<td>0.0 (0.0)</td>
<td>0.73 (0.22–2.4)</td>
<td>1.3 (0.75–2.4)</td>
</tr>
<tr>
<td>Organic brain disorder/other</td>
<td>63</td>
<td>1.4 (0.79–2.7)</td>
<td>0.0 (0.0)</td>
<td>1.6 (0.89–2.7)</td>
<td>0.0 (0.0)</td>
<td>0.95 (0.40–2.2)</td>
<td>1.5 (0.91–2.4)</td>
</tr>
<tr>
<td>Restriction order under section</td>
<td>247</td>
<td>0.52 (0.26–1.07)</td>
<td>1.9 (0.39–9.7)</td>
<td>0.50 (0.24–1.1)</td>
<td>0.11 (0.01–1.3)</td>
<td>0.33 (0.13–0.86)</td>
<td>0.45 (0.28–0.75)</td>
</tr>
<tr>
<td>37/41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sexual deviation2</td>
<td>34</td>
<td>0.41 (0.1–1.7)</td>
<td>10.5 (2.6–41.9)</td>
<td>1.3 (0.57–3.0)</td>
<td>0.0 (0.0)</td>
<td>0.37 (0.05–2.7)</td>
<td>0.96 (0.49–1.9)</td>
</tr>
<tr>
<td>Antisocial personality disorder2</td>
<td>302</td>
<td>1.6 (1.1–2.3)</td>
<td>2.0 (0.75–5.4)</td>
<td>1.1 (0.80–1.6)</td>
<td>0.86 (0.27–2.7)</td>
<td>1.2 (0.81–1.9)</td>
<td>1.4 (1.1–1.8)</td>
</tr>
<tr>
<td>Admission category v. mental illness2</td>
<td>957</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychopathic disorder</td>
<td>74</td>
<td>2.4 (1.2–4.8)</td>
<td>3.0 (0.70–13.2)</td>
<td>2.6 (1.4–4.8)</td>
<td>2.5 (0.58–10.6)</td>
<td>1.7 (0.74–4.0)</td>
<td>2.6 (1.7–4.1)</td>
</tr>
<tr>
<td>Mental impairment</td>
<td>9</td>
<td>0.0 (0.0)</td>
<td>0.0 (0.0)</td>
<td>1.6 (0.21–11.7)</td>
<td>0.0 (0.0)</td>
<td>0.0 (0.0)</td>
<td>0.89 (0.12–6.5)</td>
</tr>
<tr>
<td>Mental illness plus psychopathy</td>
<td>14</td>
<td>1.2 (0.27–4.9)</td>
<td>4.2 (0.49–36.2)</td>
<td>0.0 (0.0)</td>
<td>7.6 (0.83–69.4)</td>
<td>2.3 (0.69–7.8)</td>
<td>1.6 (0.60–4.4)</td>
</tr>
<tr>
<td>Others</td>
<td>285</td>
<td>1.7 (1.2–2.3)</td>
<td>0.96 (0.32–2.9)</td>
<td>1.4 (0.98–1.9)</td>
<td>1.5 (0.50–4.7)</td>
<td>1.4 (0.91–2.1)</td>
<td>1.4 (1.1–1.8)</td>
</tr>
</tbody>
</table>

1. Number of cases varies according to the type of offences.
2. Each of these was a substitution for primary diagnosis with all the above covariates adjusted.

not measured in this study and which should be included in future studies. The modus operandi of the index offence leading to admission may also have been important in the prediction of reconviction in this sample, but could not be identified from this study.

Compared with our sample of patients discharged from medium secure services, reoffending was two and a half times more prevalent, and violent reoffending five times more prevalent, among a cohort of released prisoners (National Offender Management Service, 2004). However, our findings cannot be directly compared with criminal recidivism among released prisoners. The criminal careers of our patients, most of whom had psychotic illness, differed and their mean age was greater. Before it can be concluded that factors such as the presence of mental
disorder, the effects of treatment interventions, or subsequent after-care, are associated with reduced offending, it will be necessary to match samples of patients and prisoner controls. However, our findings can be compared with those of previous follow-up studies from medium secure services and high-security hospitals. Taking length of follow-up into account, where patients progressively accrue further convictions over time, our findings in Fig. 1 appear to be broadly similar to previous reports. Studies with follow-up periods ranging from 1 to 5 years have demonstrated prevalence rates for ‘all offending’ of 11–16% (Falla et al., 2000; Edwards et al., 2002; Maden et al., 2004), rising to 30% for those with a mean follow-up of 6 years (Friendship et al., 1999). The more recent special hospital cohorts demonstrated higher prevalence rates (e.g. 34–38%) (Buchanan, 1998; Jamieson & Taylor, 2004) for longer follow-up periods (9–10.5 years), but these included larger proportions of high-risk patients, including more with personality disorder.

Methodological considerations

In the present study, the follow-up period was longer, and the sample larger than previous studies of patients discharged from medium secure services. However, the study was subject to the same limitations of the Offenders Index. This has a small source of error among patients with long follow-up periods, but more serious limitations for recent discharges. The time lag in the criminal justice system between charges and conviction in some cases can be over 2 years. This means that the true cut-off point is earlier in some cases. Criminal convictions are recorded using an offender’s name and some offenders change their names frequently. An estimate of missing data is that 9% of criminal records will be missing from the Offenders Index (Buchanan, 1998). Searching of multiple sources was not carried out. These limitations therefore indicate that rates of offending by this cohort are likely to be higher than we have reported.

This study included more outcome categories of offending than previous studies, and has avoided the difficulties posed for future replication by the use of different permutations of offence categories. Future studies should include measures of specific categories of reoffending in addition to rates of subsequent offending for entire samples. These are more representative of the effectiveness of interventions and do not obscure the specific risks of small groups such as sex offenders. An additional strength was inclusion of data on time at risk of reoffending. However, time at risk did not include offending while in secure services.

Some hospital case files were unavailable or access was denied. We were also unable to ensure for all patients that re-admission to a hospital setting had not occurred at some time during the follow-up period when estimating time at risk. Finally, findings may differ among more recent cohorts of discharged patients. For example, patients now spend longer in medium secure services than during the years of our study and few have diagnoses of personality disorder (Maden et al., 2004).

Risk factors and risk management

Our findings correspond to a meta-analysis of previous studies of risk factors for offending by mentally disordered offenders (Bonta et al., 1998) which demonstrated that major predictors of recidivism are the same as those for non-disordered offenders. Criminal history and actuarial measures were the best predictors, whereas clinical variables showed the smallest effect sizes. A history of previous convictions for the same offence was among the largest and most consistent predictors in our study. Primary diagnoses of personality disorder and previous detention under the legal category ‘psychopathic disorder’ have been consistently associated with a greater risk of reoffending in previous UK follow-up studies, except that of Phillips et al. (2005). However, increased risks for patients with mania/hypomania have not previously been reported, suggesting that this subgroup has specific characteristics which require further study to improve their clinical management. Patients with depressive disorder were no less likely than those with schizophrenia to offend and their risks were greatly increased for sexual offending. Male gender, younger age when first appearing in court and being younger when discharged from medium secure services have also previously been demonstrated to increase risks of reoffending. However, lifetime history of comorbid substance misuse or dependence, being born outside the UK and never having married had little effect in this study. The most powerful protective factors were Home Office restrictions requiring patients to accept supervision and treatment following discharge, and longer detention in medium secure services.

Different categories of offending behaviour had different profiles of risk factors. Our findings can be used to identify patients who are at especially high risk of reoffending. Those with two or more previous violent convictions, a primary diagnosis of personality disorder, or a comorbid diagnosis of antisocial personality disorder are at increased risk of future violent offending. These risks are further increased among Black and other minority ethnic groups, and those younger when first appearing in court. However, risk management to prevent future violence must include a long-term perspective, as the cumulative probability of future violence increases linearly in these patients. Risk of subsequent sex offending was especially high among those with primary diagnoses of affective disorder, those from minority ethnic groups and those previously convicted for sex offences. It was unsurprising that sexual deviation considerably increased risks. This indicates that interventions for, and subsequent monitoring of, deviant sexual propensities should be prioritised over perceived risks from symptomatic conditions such as schizophrenia among sex offenders admitted to these services. Previous studies of sex offenders suggest that their risks are long-term, with some support from our findings on the probability of offending for all patients. However, this study has identified that patients with previous sex offending behaviour require special vigilance in their after-care during the first 3–4 years post-discharge.

Convictions for arson were more common among women, among those with a history of alcohol dependence/misuse and those with previous convictions for arson. Risk management is long-term in this subgroup. Cumulative probability of recidivism increases linearly, with emergence of increased risk at 6 years post-discharge. Similar previous convictions were also the strongest predictor of future acquittal offending, with increased risks among younger patients, male patients, patients with personality disorder and those who had started their criminal careers earlier.

The variables included in our analysis represent largely historical or ‘static’ risk factors for different categories of offending. It has been argued that long-term
reoffending is best predicted by static factors (Hanson & Bussiere, 1998) as these indicate established characteristics of the individual that are readily incorporated into an actuarial measure. However, future reoffending can only be prevented by addressing problems that present in the community following discharge, including criminogenic needs and dynamic risk factors, which can be changed and are amenable to intervention (Bonta, 1996; Andrews & Bonta, 1998). This would include adherence to prescribed medication and after-care. Future research should concentrate on examining the effectiveness of interventions after discharge that are designed to influence changeable factors encountered outside a secure setting. The question also remains whether more prolonged application of restrictions, including enhanced supervision and surveillance, and with compulsion to accept treatment, instead of reliance on the care programme approach, will result in more effective reduction of reoffending in patients who are identified as high risk. Longer periods in security and restrictions on patients’ behaviour and lifestyles following discharge were associated with significant reductions in risk of serious reoffending in this sample.

ACKNOWLEDGEMENTS

This study was funded by a grant from the Wellcome Trust. It would not have been possible without the cooperation of clinical and administrative staff in the medium secure units of North West Thames, North East Thames, West Midlands, Merseyside, North Western, South Western and East Anglian regions, and Keevesworth House, Stockton Hall and St. Andrew’s, Northampton.

REFERENCES


Butler Committee (1975) Report of the Committee on Mentally Abnormal Offenders TSO (The Stationery Office).


National Institute for Mental Health in England (2003) Personality Disorder: No Longer a Diagnosis of Exclusion. NIMHE.


Anterior cingulate grey-matter deficits and cannabis use in first-episode schizophrenia

PHILIP R. SZESZKO, DELBERT G. ROBINSON, SERGE SEVY, SANJIV KUMRA, CLAUDIA I. RUPP, JULIA D. BETENSKY, TODD LENCZ, MANZAR ASHTARI, JOHN M. KANE, ANIL K. MALHOTRA, HANDAN GUNDUZ-BRUCE, BARBARA NAPOLITANO and ROBERT M. BILDER

Background  Despite the high prevalence of cannabis use in schizophrenia, few studies have examined the potential relationship between cannabis exposure and brain structural abnormalities in schizophrenia.

Aims  To investigate prefrontal grey and white matter regions in patients experiencing a first episode of schizophrenia with an additional diagnosis of cannabis use or dependence (n = 20) compared with similar patients with no cannabis use (n = 31) and healthy volunteers (n = 56).

Method  Volumes of the superior frontal gyrus, anterior cingulate gyrus and orbital frontal lobe were outlined manually from contiguous magnetic resonance images and automatically segmented into grey and white matter.

Results  Patients who used cannabis had less anterior cingulate grey matter compared with both patients who did not use cannabis and healthy volunteers.

Conclusions  A defect in the anterior cingulate is associated with a history of cannabis use among patients experiencing a first episode of schizophrenia and could have a role in poor decision-making and in choosing more risky outcomes.

Declaration of interest  None. Funding detailed in Acknowledgements.

The prevalence of cannabis use in schizophrenia can range up to 43% (Bersani et al, 2002) and poses unique treatment challenges (Green et al, 2004). The use of cannabis in schizophrenia tends to be associated with earlier age at first psychotic episode, especially among male patients (Van Mastrigt et al, 2004), and is associated with an unfavourable outcome (Henquet et al, 2005). Although numerous studies have identified frontal brain structural abnormalities in schizophrenia (Goldstein et al, 1999; Gur et al, 2000), little research has been directed at understanding the potential association between these abnormalities and cannabis use. Animal studies suggest that using Δ⁹-tetrahydrocannabinol, the main psychoactive component of cannabis, may be neurotoxic to the frontal lobes (Verrico et al, 2003), which are believed to have a key role in the neurobiology of schizophrenia (Goldman-Rakic & Selemon, 1997). In this study we investigated three prefrontal cortical regions (the superior frontal gyrus, anterior cingulate gyrus and orbital frontal lobe) implicated in drug addiction (Goldstein & Volkow, 2002; Tucker et al, 2004), in a sample of patients with a first episode of schizophrenia with or without a history of cannabis use compared with healthy volunteers. We tested the hypothesis that patients with the dual diagnosis would have greater prefrontal structural abnormalities compared with patients who did not use cannabis and with healthy volunteers.

METHOD

The 51 patients included in this study were recruited from admissions to the in-patient service at the Zucker Hillside Hospital in Glen Oaks, New York, and were participating in clinical trials comparing the efficacy of atypical antipsychotic drugs. All patients were interviewed using the Structured Clinical Interview for Axis I DSM–IV Disorders (SCID; First et al, 1994) and met DSM–IV criteria (American Psychiatric Association, 1994) for schizophrenia (n = 36), schizoaffective disorder (n = 8) or schizophrenia-form disorder (n = 7). Of the 51 patients with schizophrenia included in this study, 8 had a diagnosis of cannabis abuse and 12 had a diagnosis of cannabis dependence. Of the 20 patients with a diagnosis of either cannabis abuse or dependence, 6 had a diagnosis of alcohol abuse (n = 5) or dependence (n = 1). Among the 6 patients who had a diagnosis of alcohol abuse or dependence, other substance use diagnoses included cocaine abuse (n = 1), hallucinogen abuse (n = 1) and opioid dependence (n = 1). None of the 14 patients with a diagnosis of cannabis abuse or dependence had any other substance use diagnosis. Only 2 of the 31 patients without a diagnosis of cannabis abuse or dependence had any other substance use diagnosis (1 with alcohol abuse and 1 with alcohol dependence). Twenty-five patients were antipsychotic drug-naïve at the time of the scan, including 6 patients from the group with cannabis use and 19 patients from the group without cannabis use. The median duration of antipsychotic drug exposure from entry into the clinical trial and the magnetic resonance imaging examination was 0 weeks (range 0–34 weeks).

Fifty-six healthy volunteers were recruited from local newspaper advertisements and through word of mouth in the community and denied any history of psychiatric or medical illness as determined by clinical interview and the non-patient version of the SCID (SCID–NP, Spitzer & Williams, 1988). Thus, no one in the healthy comparison group had a substance use diagnosis. Exclusion criteria for all study participants included serious neurological or endocrine disorder, any medical condition or treatment known to affect the brain, and meeting DSM–IV criteria for mental retardation. All procedures were approved by the local institutional review board and written informed consent was obtained from all participants.

Classification of handedness was based on a modified version of the Edinburgh Inventory consisting of 20 items (Oldfield, 1971). Participants with a laterality quotient greater than 0.70 were classified as dextral and the rest as non-dextral (Schachter et al, 1987). Handedness for 3 patients without cannabis use and 9 healthy volunteers was assessed solely on the basis of handwriting preference.
Imaging procedures
Magnetic resonance imaging (MRI) scans were conducted at Long Island Jewish Medical Center and were acquired in the coronal plane using three-dimensional fast spoiled gradient recalled acquisition with inversion recovery (time to repetition) 12.7 ms or 14.7 ms, echo time 4.5 ms or 5.5 ms, field of view 22 cm) on a 1.5 T whole-body superconducting imaging system (General Electric, Milwaukee, Wisconsin, USA). This sequence produced 124 contiguous images (slice thickness 1.5 mm) through the whole head with in-plane resolution of 0.86 mm × 0.86 mm in a 256 × 256 matrix.

Measurement procedures
All measurements were completed in MEDx (Sensor Systems, Sterling, Virginia, USA). The images were aligned along the anterior and posterior commissures for standardisation across individuals and flipped randomly along the right-left axis. Scans were mixed together randomly and no identifying information was available to the operator from the scan. All measurements were thus completed by an operator who was masked to group membership and hemisphere.

Total intracranial contents
Measurement of total intracranial contents was completed in MEDx by computing the volume of the total cerebrum, cerebrospinal fluid, cerebellum and brainstem. Interrater reliability between two raters as assessed by intraclass correlations in nine cases was 0.99.

Frontal lobe subregions
Measurement of the frontal lobe subregions was completed using methods described previously (Szeszko et al., 1999), which were adapted from Rademacher et al. (1992) for use in our magnetic resonance images. This method has been used in our previous work (Szeszko et al., 2000, 2004) and utilises the cerebral sulci in combination with a set of coronal planes that close the selected regions of interest. The boundaries of the superior frontal gyrus were the tip of the cingulate sulcus (anterior), the connection of the superior and precentral sulci (posterior), the precentral sulcus (lateral) and the cingulate sulcus (medial). The boundaries of the anterior cingulate gyrus were the tip of the cingulate sulcus (anterior), the connection of the superior and precentral sulci (posterior), the callosal sulcus (ventral) and the cingulate sulcus (dorsal). The boundaries of the orbital frontal region were the last appearance of the anterior horizontal ramus (anterior), the last appearance of the olfactory sulcus (posterior), the anterior horizontal ramus/circular sulcus of insula (lateral) and the olfactory sulcus (medial) (Fig. 1). Because one of the limiting sulci required for measurement of the orbital frontal region (the anterior horizontal ramus) was not present in every hemisphere (Ono et al., 1990; Szeszko et al., 1999), orbital frontal volumes could not be computed for some individuals (see Table 2).

All regions were outlined manually in the coronal plane on a slice-by-slice basis and included both grey and white matter (Fig. 2). After outlining the frontal region of interest, the operator automatically segmented it into grey and white matter using a thresholding method generated from histograms (Otsu, 1979), as described previously (Lim et al., 1992; Szeszko et al., 2004). Intraclass correlations between two or three operators for these brain structures (number of cases 8–10) were as follows: anterior cingulate gyrus grey matter, right hemisphere, ICC = 0.90, left hemisphere, ICC = 0.94; anterior cingulate gyrus white matter, right, ICC = 0.94, left, ICC = 0.94; superior frontal gyrus grey matter, right, ICC = 0.92, left, ICC = 0.97; superior frontal gyrus white matter, right, ICC = 0.95, left, ICC = 0.95; orbital frontal lobe grey matter, right, ICC = 0.92, left, ICC = 0.99; orbital frontal lobe white matter, right, ICC = 0.94, 0.94, left, ICC = 0.90.

Statistical procedures
The mixed-models approach (SAS version 8.2 for Windows) for repeated-measures analysis of variance was used to compare brain structure volumes between patients and healthy volunteers. Analyses were conducted separately for the anterior cingulate, superior frontal gyrus and orbital frontal lobe because of their neuroanatomical heterogeneity. For each of the three frontal regions the statistical model included group (healthy volunteers v. patients with cannabis use v. patients without cannabis use) and gender as between-subject factors. Tissue type (grey v. white) and hemisphere.
were repeated measures. Age was included as a covariate because the groups differed in age and because age correlated with the brain structure volumes. Intracranial volume was included as a covariate to control for non-specific differences in brain size among individuals. Subsequent analyses excluded any patient with a substance use diagnosis other than cannabis abuse or dependence. Group differences in demographic variables were examined using independent group t-tests. Chi-square tests were used to examine differences in categorical variables. Analyses of frontal brain structure volumes were conducted using two-tailed tests with α = 0.05 (0.05 divided by number of brain regions). Results of analyses for individual brain structure volumes are presented for descriptive purposes only (see Table 2).

RESULTS

Sample characteristics for the two patient groups and the healthy comparison group are given in Table 1. The three groups did not differ significantly from each other in distributions of age, gender, handedness or the absence of the right or left anterior horizontal ramus (P > 0.05). Also, as expected, the groups differed in education (F = 6.64, d.f. = 2.97, P = 0.002), with healthy volunteers having more education than either patient group. Patients with cannabis use did not differ significantly from patients without cannabis use in distributions of age at first psychotic symptoms, years of education, duration of antipsychotic treatment, duration of untreated psychosis or global assessment of functioning (all P > 0.05; Table 1). There were, however, significantly more patients who were antipsychotic drug-naïve among the cannabis user group compared with the non-cannabis user group ($\chi^2 = 4.76, \text{d.f.} = 1, P = 0.029$).

Mean brain structure volumes for the three study groups are given in Table 2, along with the adjusted 95% confidence intervals for the difference between group means. The main finding that distinguished the groups was a significant group-by-tissue type interaction for the anterior cingulate ($F_{2,108} = 6.39, P = 0.002$). Follow-up tests revealed that patients who used cannabis had significantly less anterior cingulate grey matter compared with patients who did not ($t_{1,108} = -2.41, P = 0.018$) and with healthy volunteers ($t_{1,108} = -2.19, P = 0.031$). Repeating the analysis with antipsychotic drug-naïve status as a covariate revealed that patients who used cannabis had significantly less anterior cingulate grey matter compared with patients who did not ($t_{1,48} = -2.40, P = 0.020$). Individual data points illustrating total anterior cingulate grey-matter volumes for the three groups are provided in Fig. 3. None of the interactions involving gender was statistically significant for the anterior cingulate. Neither the main effect of group nor group-by-tissue type interaction was statistically significant for the orbital frontal lobe (all P > 0.05). In addition, the main effect of group was not statistically significant for the superior frontal gyrus.

There also was a significant main effect of hemisphere for the anterior cingulate ($F_{1,110} = 19.3, P < 0.001$) and orbital frontal lobe ($F_{1,84} = 16.1, P < 0.001$). Overall, participants had more grey matter ($t_{1,110} = -4.78, P < 0.001$) and white matter ($t_{1,110} = -3.33, P = 0.001$) in the right anterior cingulate compared with the left anterior cingulate. In addition, participants had more grey matter ($t_{1,94} = -3.80, P < 0.001$) and white matter ($t_{1,94} = 4.07, P < 0.001$) in the left compared with the right orbital frontal lobe. No significant hemispheric asymmetry was evident in the sample for either superior frontal gyrus grey-matter or white-matter volumes ($P > 0.05$).

Subsequent analyses investigated whether having other substance use diagnoses influenced the observed findings. The group-by-tissue type interaction remained statistically significant ($F_{2,100} = 6.07, P = 0.003$) for the anterior cingulate when we excluded patients from analysis who had any substance use diagnosis other than cannabis abuse or dependence. Specifically, patients with either cannabis abuse or dependence as their sole substance use diagnosis had significantly less anterior cingulate grey matter than patients without any substance use diagnosis ($t_{1,100} = -2.45, P < 0.016$) and healthy volunteers ($t_{1,100} = -2.34, P = 0.021$).

<table>
<thead>
<tr>
<th>Table I Sample characteristics</th>
<th>Healthy comparison group (n=56)</th>
<th>Schizophrenia group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No cannabis use (n=31)</td>
<td>Cannabis use (n=20)</td>
</tr>
<tr>
<td>Gender, n</td>
<td>Male</td>
<td>36</td>
</tr>
<tr>
<td>Gender, n</td>
<td>Female</td>
<td>20</td>
</tr>
<tr>
<td>Age, years: mean (s.d.)</td>
<td>25.7 (6.7)</td>
<td>24.8 (4.9)</td>
</tr>
<tr>
<td>Handedness, n&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Dextral</td>
<td>42</td>
</tr>
<tr>
<td>Handedness, n&lt;sup&gt;1&lt;/sup&gt;</td>
<td>Non-dextral</td>
<td>12</td>
</tr>
<tr>
<td>Education, years: mean (s.d)&lt;sup&gt;2&lt;/sup&gt;</td>
<td>14.6 (2.6)</td>
<td>13.2 (1.8)</td>
</tr>
<tr>
<td>Age at first psychotic symptoms: mean (s.d)</td>
<td>–</td>
<td>21.4 (4.6)</td>
</tr>
<tr>
<td>Antipsychotic treatment, weeks: mean (range)</td>
<td>–</td>
<td>3.1 (0-26.7)</td>
</tr>
<tr>
<td>Antipsychotic drug-naïve, n</td>
<td>–</td>
<td>19</td>
</tr>
<tr>
<td>Duration of untreated psychosis, weeks: mean (range)</td>
<td>–</td>
<td>169 (1-828)</td>
</tr>
<tr>
<td>Global Assessment of Functioning score: mean (s.d.)</td>
<td>–</td>
<td>33.1 (13.4)</td>
</tr>
</tbody>
</table>

1. Missing data from healthy group (n=2).
2. Missing data from healthy group (n=7).
Table 2
Unadjusted frontal lobe volumes and adjusted confidence intervals for group differences

<table>
<thead>
<tr>
<th>Frontal lobe volume, cm³</th>
<th>Healthy control group (n = 31)</th>
<th>FESZ without cannabis use (n = 20)</th>
<th>FESZ with cannabis use (n = 20)</th>
<th>Adjusted confidence intervals of difference between groups</th>
<th>ES</th>
<th>95% CI</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grey matter total</td>
<td>33.3 (5.51)</td>
<td>32.9 (5.10)</td>
<td>33.3 (5.51)</td>
<td>-0.25 to 4.35</td>
<td>0.04</td>
<td>[0.31, 0.07]</td>
<td>[0.31, 0.07]</td>
</tr>
<tr>
<td>White matter total</td>
<td>25.4 (4.81)</td>
<td>25.4 (5.52)</td>
<td>25.4 (5.52)</td>
<td>0.16 to 1.96</td>
<td>0.01</td>
<td>[0.00, 0.22]</td>
<td>[0.00, 0.22]</td>
</tr>
<tr>
<td>Right hemisphere</td>
<td>16.8 (2.89)</td>
<td>16.0 (3.22)</td>
<td>16.5 (2.96)</td>
<td>0.39 to 2.21</td>
<td>0.02</td>
<td>[0.01, 0.03]</td>
<td>[0.01, 0.03]</td>
</tr>
<tr>
<td>Grey matter total</td>
<td>16.7 (2.61)</td>
<td>17.0 (3.51)</td>
<td>16.5 (2.95)</td>
<td>0.39 to 2.21</td>
<td>0.02</td>
<td>[0.01, 0.03]</td>
<td>[0.01, 0.03]</td>
</tr>
<tr>
<td>White matter total</td>
<td>12.8 (2.36)</td>
<td>12.4 (2.81)</td>
<td>13.0 (2.34)</td>
<td>-0.46 to 1.39</td>
<td>0.01</td>
<td>[0.00, 0.02]</td>
<td>[0.00, 0.02]</td>
</tr>
<tr>
<td>Left hemisphere</td>
<td>12.7 (3.10)</td>
<td>12.1 (1.05)</td>
<td>12.7 (1.05)</td>
<td>0.16 to 1.10</td>
<td>0.01</td>
<td>[0.00, 0.02]</td>
<td>[0.00, 0.02]</td>
</tr>
<tr>
<td>Grey matter total</td>
<td>4.39 (1.11)</td>
<td>4.46 (1.54)</td>
<td>4.39 (1.11)</td>
<td>0.075 to 0.53</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>White matter total</td>
<td>2.28 (0.58)</td>
<td>2.46 (0.79)</td>
<td>2.24 (0.78)</td>
<td>0.39 to 0.38</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>Anterior cingulate gyrus</td>
<td>10.1 (2.22)</td>
<td>10.4 (2.59)</td>
<td>10.5 (2.69)</td>
<td>-0.50 to 0.36</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>Grey matter total</td>
<td>3.89 (1.00)</td>
<td>4.06 (1.02)</td>
<td>4.05 (1.06)</td>
<td>0.30 to 0.29</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>White matter total</td>
<td>2.38 (0.63)</td>
<td>2.46 (0.79)</td>
<td>2.44 (0.78)</td>
<td>0.038 to 0.48</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>Right hemisphere</td>
<td>5.89 (1.19)</td>
<td>5.54 (1.54)</td>
<td>5.89 (1.19)</td>
<td>0.41 to 1.19</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>Grey matter total</td>
<td>3.96 (1.00)</td>
<td>3.94 (1.02)</td>
<td>3.96 (1.02)</td>
<td>-0.17 to 0.89</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>White matter total</td>
<td>1.26 (0.63)</td>
<td>1.36 (0.65)</td>
<td>1.36 (0.65)</td>
<td>0.038 to 0.48</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>Orbital frontal</td>
<td>10.1 (2.22)</td>
<td>10.4 (2.59)</td>
<td>10.5 (2.69)</td>
<td>-0.50 to 0.36</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>Grey matter total</td>
<td>3.89 (1.00)</td>
<td>4.06 (1.02)</td>
<td>4.05 (1.06)</td>
<td>0.30 to 0.29</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>White matter total</td>
<td>2.38 (0.63)</td>
<td>2.46 (0.79)</td>
<td>2.44 (0.78)</td>
<td>0.038 to 0.48</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>Right hemisphere</td>
<td>5.89 (1.19)</td>
<td>5.54 (1.54)</td>
<td>5.89 (1.19)</td>
<td>0.41 to 1.19</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>Grey matter total</td>
<td>3.96 (1.00)</td>
<td>3.94 (1.02)</td>
<td>3.96 (1.02)</td>
<td>-0.17 to 0.89</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>White matter total</td>
<td>1.26 (0.63)</td>
<td>1.36 (0.65)</td>
<td>1.36 (0.65)</td>
<td>0.038 to 0.48</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>Left hemisphere</td>
<td>5.89 (1.19)</td>
<td>5.54 (1.54)</td>
<td>5.89 (1.19)</td>
<td>-0.50 to 0.36</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>Grey matter total</td>
<td>3.89 (1.00)</td>
<td>4.06 (1.02)</td>
<td>4.05 (1.06)</td>
<td>0.30 to 0.29</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
<tr>
<td>White matter total</td>
<td>2.38 (0.63)</td>
<td>2.46 (0.79)</td>
<td>2.44 (0.78)</td>
<td>0.038 to 0.48</td>
<td>0.00</td>
<td>[0.00, 0.00]</td>
<td>[0.00, 0.00]</td>
</tr>
</tbody>
</table>

ES, effect size; FESZ, first-episode schizophrenia.

1. Volumes adjusted for age and total intracranial contents.
2. Owing to the absence of the anterior horizontal ramus volumes could not be computed for 5 healthy volunteers, 4 patients without cannabis use and 2 patients with cannabis use.

233
DISCUSSION

Understanding the relationship between prefrontal grey matter and cannabis use in schizophrenia may have important implications for improving our understanding of the potentially deleterious effects of these substances on brain structure in this disorder. Using methods for cortical parcellation of the prefrontal cortex based on the sulcal anatomy, we report that patients experiencing a first episode of schizophrenia who have a history of cannabis use have less anterior cingulate grey matter compared with similar patients who do not use cannabis and with healthy volunteers. We obtained similar findings when we excluded patients with substance use diagnoses other than cannabis abuse or dependence from analysis.

Other studies

Little research has been directed at understanding the relationship between cannabis use and brain structure, especially in schizophrenia, and thus it is difficult to compare our findings with prior work. In a structural neuroimaging study Cahn et al (2004) did not identify differences in total grey- and white-matter volumes between patients with recent-onset schizophrenia comorbid with cannabis abuse or dependence and patients with no cannabis use, but did not examine discrete frontal cortical regions. Several studies reported grey-matter structural alterations in cannabis users, however, and this may have relevance for the findings reported here. For example, Matochik et al (2005) reported that individuals who used cannabis had lower grey-matter density in the right parahippocampal gyrus and greater density bilaterally near the precentral gyrus and right thalamus compared with those who did not. In addition, Wilson et al (2000) reported lower whole-brain grey-matter volume among individuals who started using cannabis before age 17 years compared with individuals who started using cannabis later. Moreover, the use of other illicit substances such as cocaine has been linked to anterior cingulate grey-matter structural alterations (Franklin et al, 2002; Matochik et al, 2003).

The anterior cingulate is believed to play an important part in mediating executive functions, including set-shifting and response inhibition, which have been reported to be abnormal among individuals who use cannabis (Gruber & Yurgelun-Todd, 2005). Several studies reported aberrant anterior cingulate activity among cannabis users while performing the Stroop task, which requires the ability to inhibit prepotent tendencies to respond (Eldreth et al, 2004; Gruber & Yurgelun-Todd, 2005). It is also noteworthy that abnormal anterior cingulate activity was also reported in cannabis users while performing a motor sequencing task (Pillay et al, 2004) and in individuals exposed to marijuana prenatally (Smith et al, 2004). People who are substance users also find it difficult to inhibit their own actions as working memory demands increase (Hester & Garavan, 2004), and individuals who use cannabis may need the anterior cingulate to ‘work harder’ to complete task demands (Kanayama et al, 2004). Drug craving has also been linked with anterior cingulate activity (Kilts et al, 2001) and more specifically with attentional biases for cannabis-related stimuli (Field et al, 2004).

Implications

Risky decision-making is considered integral to the phenomenology of drug use (Fishbein et al, 2005) and such decisions are intimately linked with reward and punishment, which is mediated by neural systems involving the anterior cingulate (Shidara & Richmond, 2002). Patients with schizophrenia who use cannabis may have deficits in the ability to balance rewards and punishments, which could contribute to drug-taking behaviour. Specifically, people who use cannabis tend to make decisions based on large immediate gains in spite of more costly losses (Whitlow et al, 2004). Among drug users, risky choices during a decision-making test were associated with abnormal metabolic activity in the anterior cingulate (Tucker et al, 2004), which may partly form the neuroanatomical substrate of choosing risky outcomes. Our results are compatible with the hypothesis that grey-matter structural alterations involving the anterior cingulate in patients with schizophrenia using cannabis could be associated with poor decision-making and partly mediate the compulsive drive towards drug use (Adinoff, 2004).

We did not observe significant differences between the group with first-episode schizophrenia with no history of cannabis use and the healthy volunteer group in any of the prefrontal grey-matter or white-matter volumes. One potentially important consideration in the assessment of brain structure in schizophrenia, however, is illness duration, especially given some evidence that grey-matter deficits in schizophrenia are progressive (Mathalon et al, 2001; Cahn et al, 2002; Pantelis et al, 2003) and that such deficits occur only after the first few years following illness onset (Molina et al, 2004). It is therefore conceivable that prefrontal grey-matter structural alterations might become apparent later in the course of schizophrenia, at least in our cohort of patients without cannabis use. In addition, it would be helpful to elucidate the potential effects of cannabis on the anterior cingulate in longitudinal studies, especially given that this region has been implicated in the transition to psychosis (Pantelis et al, 2003).
GREY-MATTER DEFICITS AND CANNABIS USE IN SCHIZOPHRENIA

Limitations
There were several limitations to our study that should be acknowledged. One potential limitation is the extensive amount of time required to outline the frontal lobe regions of interest. An alternative approach might be the use of voxel-based morphometry, although this method requires brain normalisation and smoothing, which could result in the loss of information if abnormalities are subtle and localised to small regions. Also, there are inherent challenges in using sulcal anatomical features as the boundaries of regions of interest, given their heterogeneous presentation; however, we believe that this approach provides greater cytoarchitectonic validity compared with methods based on invariant landmarks not appearing on the cortical surface. In addition, qualitative methods for mapping cingulate and paracortex regions identified (Yucel et al., 2002) may be useful in complementing the volumetric approach described here. Another potential limitation is the lack of a study group who used cannabis but did not have schizophrenia, to test whether our findings were more generally associated with cannabis use. Moreover, given the cross-sectional nature of this study, we could not determine whether anterior cingulate grey-matter volume deficits predispose patients to use cannabis or whether these deficits are a consequence of cannabis use.

ACKNOWLEDGEMENTS
This work was supported in part by grants from NARSAD (PR5) and the National Institute of Mental Health to P.R.S. (MH01990), R.M.B. (MH60374), S.S. (DA015541), J.M.K. (MH60575), D.G.R. (MH66004) and the North Shore—Long Island Jewish Research Institute General Clinical Research Center (MO1 RR018535).

REFERENCES


Mortality in people with schizophrenia in rural China

10-year cohort study

MAO-SHENG RAN, ERIC YU-HAI CHEN, YEATES CONWELL, CECILIA LAI-WAN CHAN, PAUL S. F. YIP, MENG-ZE XIANG and ERIC D. CAINE

Background Long-term mortality and the risk factors for premature death among patients with schizophrenia living in rural communities are unknown.

Aims To explore the 10-year mortality and its risk factors among patients with schizophrenia.

Method We used data from a 10-year prospective follow-up study (1994–2004) of mortality among people with schizophrenia, and death registration data for Xinjin County, Chengdu, China.

Results The mortality rate was 2228 per 100 000 person-years during follow-up. Both all-cause mortality and suicide rates were significantly greater in male than in female patients. Age at illness onset (>45 years), duration of illness (>10 years), age greater than 50 years, physical illness, inability to work, male gender, and never having received treatment were identified as independent predictors of increased mortality.

Conclusions Higher mortality rates in male patients may contribute to the higher prevalence of schizophrenia in women compared with men in China. The findings of risk factors for mortality should be taken into account when developing interventions to improve outcomes among people with schizophrenia.

Declaration of interest None. Funding detailed in Acknowledgements.

Although the mortality rate is higher among people with schizophrenia than among the general population (Mortensen & Juel, 1993; Harris & Barraclough, 1997; Harrison et al., 2001), little is known about the factors predicting death in these individuals (Salokangas et al., 2002). Most previous studies of mortality in schizophrenia were conducted in industrialised locales (Simpson & Tsuang, 1996; Brown, 1997). Mortality in a broader range of cultures should be explored (Harrison et al., 2001; Simpson & Tsuang, 1996). Mental health services are less available in rural China than in urban areas of the country (Xiang et al., 1994), possibly contributing to increased mortality rates there. However, studies of causes of death among people with schizophrenia in rural communities are sparse.

Although there are higher rates of schizophrenia in men than in women worldwide (Murray & Lopez, 1996; Aleman et al., 2003), evidence indicates that there are substantially higher rates of schizophrenia in women than in men in China (Cooper & Sartorius, 1996; Zhang et al., 1998; Liu et al., 2000; Ran et al., 2003; Phillips et al., 2004). The reasons why the pattern of schizophrenia in China differs from that in other parts of the world are unknown and largely unstudied.

Suicide is one of the most common causes of premature death in people with schizophrenia (Mortensen & Juel, 1993; De Hert & Peuskens, 2000). Given the unique pattern of suicide in China – rural rates are three times greater than urban rates, and rates in women are 25% higher than those for men in the general population (Phillips et al., 2002) – it is crucial to explore the characteristics of suicide in people with schizophrenia in rural China (Ran & Chen, 2004). Previous studies have been limited by use of retrospective or cross-sectional designs, small numbers of patients, general problems of diagnostic standardisation, restricting the focus to in-patients only, and the relatively low rates of follow-up. Therefore, long-term prospective follow-up study of mortality in people with schizophrenia living in the community should be worthwhile (Harrison et al., 2001; Ran & Chen, 2004).

Our study objectives were to explore the rates of all-cause mortality and suicide among people with schizophrenia, and to identify major causes of death and factors increasing the risk of death.

METHOD

Study population

The study sample consisted of people with schizophrenia (n=510) identified in an epidemiological investigation of 123 572 persons aged 15 years and older in six townships of Xinjin County in October 1994. Cases were identified through screening procedures for psychosis (face-to-face interviews with the head of each household, together with the key informant method) and general psychiatric interview. The details of this investigation have been described in previous papers (Ran et al., 2001, 2003). All participants lived in rural communities and met ICD–10 criteria (World Health Organization, 1992) for a diagnosis of schizophrenia based on standardised administration of the Present State Examination (PSE–9; Wing et al., 1974) by trained research interviewers. Using the baseline data acquired in 1994, we followed up and interviewed all the individuals with schizophrenia and their informants in May 2004. All respondents gave informed consent.

Measurement

The principal assessment tools included the PSE and Social Disability Screening Schedule (SDSS; Shen et al., 1986) used in the baseline investigation in 1994 (Ran et al., 2001, 2003). For cohort members who were alive when followed up in 2004, at least one informant familiar with the person’s life and circumstances and/or the cohort members themselves were interviewed. For those who had died, the next of kin or at least one informant familiar with the dead person was interviewed. All the interviews were conducted by trained psychiatrists using the Patients Follow-up Schedule (PFS) for about 30 min; this questionnaire was used to collect information concerning demographic characteristics, cause and time of death, treatment and...
social support. For all cases, medical and psychiatric treatment records were obtained from hospital, village doctors’ clinics and traditional healers. Information from the death certification and suicide note (where applicable) was also obtained. The classification of each death as due to suicide, accident or natural causes represented the consensus opinion of interviewers and independent researchers after reviewing all information obtained during the interviews. Participants were defined as homeless and lost to follow-up if an informant reported that they had wandered and slept in public places and that their whereabouts were now unknown. Participants’ physical illnesses (e.g. heart disease, respiratory disease, cancer) at baseline and follow-up were defined according to the baseline data, informants’ reports and doctors’ diagnoses. Family economic status was defined according to the family mean income. Marked symptoms were defined according to the assessment of the PSE.

**Statistical analysis**

The follow-up period for every participant started at recruitment and ended at interview, death or the point when the individual was lost to follow-up. Mortality rates were calculated overall and by subgroups defined according to various characteristics. Mortality rates were estimated using the person–time method (number of deaths divided by person-years of follow-up). The effects of gender on all-cause mortality and suicide rates were tested using univariate Cox hazard regression analyses. Survival analyses were also used to explore gender differences in survival rates. Standardised mortality ratios were calculated by dividing observed deaths by expected deaths, with the general population of Xinjin County as the standard population. Death registration records for Xinjin County were used to derive data for the general population.

Predictors of mortality (crude mortality rates) were identified using univariate and multivariate Cox hazard regression analyses. Hazard ratios for potential predictors and the corresponding 95% confidence intervals were determined using univariate Cox regression. All variables with \( P \leq 0.10 \) in univariate analyses were included in a multivariate Cox model using a backwards procedure. Those with \( P \leq 0.05 \) were retained in the final model. The backwards procedure was used to exclude the potential confounding effect of variables by adding them, one at a time, in the final model. In addition, \( 2 \times 2 \) interactions between independent predictors were tested.

In the Cox analyses and for estimation of the mortality and suicide rates, all independent variables other than gender were treated as time-dependent. All these variables were based on the measures at baseline or follow-up.

**RESULTS**

**Characteristics of the cohort**

Of the 510 people identified as having schizophrenia, 10 were excluded because they were lost to follow-up; therefore 500 people (98.0%) were followed up from 1994 to 2004. Their characteristics are described in Table 1. Half of this group were male, two-thirds were married and in 55.8% the family’s economic status was less than the mean. One-fifth had been admitted to hospital and a third had never

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Gender, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>233 (46.6)</td>
</tr>
<tr>
<td>Married</td>
<td>321 (64.2)</td>
</tr>
<tr>
<td>Single</td>
<td>104 (20.8)</td>
</tr>
<tr>
<td>Divorced</td>
<td>31 (6.2)</td>
</tr>
<tr>
<td>Family economic status below mean, n (%)</td>
<td>279 (55.8)</td>
</tr>
<tr>
<td>Family history of mental illness, n (%)</td>
<td>141 (28.2)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Marital status, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous hospitalisation</td>
<td>105 (21.0)</td>
</tr>
<tr>
<td>Medication</td>
<td>240 (48.0)</td>
</tr>
<tr>
<td>Never received treatment</td>
<td>153 (30.6)</td>
</tr>
<tr>
<td>Age, years: mean (s.d.)</td>
<td>44.7 (15.5)</td>
</tr>
<tr>
<td>Duration of illness, years: mean (s.d.)</td>
<td>12.5 (11.3)</td>
</tr>
<tr>
<td>Age at onset, years: mean (s.d.)</td>
<td>31.1 (13.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total (n=500)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deaths Survivors</td>
<td>372 (74.4)</td>
</tr>
<tr>
<td>Deaths Suicide</td>
<td>21 (4.2)</td>
</tr>
<tr>
<td>Death due to accident</td>
<td>13 (2.6)</td>
</tr>
<tr>
<td>Death due to natural causes</td>
<td>64 (12.8)</td>
</tr>
<tr>
<td>Homeless and lost to follow-up</td>
<td>30 (6.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Women (n=267)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survivors</td>
<td>213 (79.7)</td>
</tr>
<tr>
<td>Deaths Suicide</td>
<td>6 (2.3)</td>
</tr>
<tr>
<td>Death due to accident</td>
<td>6 (2.3)</td>
</tr>
<tr>
<td>Death due to natural causes</td>
<td>28 (10.5)</td>
</tr>
<tr>
<td>Homeless and lost to follow-up</td>
<td>14 (5.2)</td>
</tr>
</tbody>
</table>

Table 1 Characteristics of the cohort in 1994 (n=500)

Table 2 Status of the cohort patients in 2004
to be alive (Table 2). The status of 30 people (6.0%) who had been homeless was unknown to their family and friends. Twenty-one people (4.2%) had died by suicide, 13 (2.6%) had died due to accident and 64 (12.8%) had died from natural causes during the follow-up period. Among all the 500 participants, 134 (26.8%) had had various physical illnesses at baseline or at some time during the follow-up period, 11 (2.2%) had migrated to cities for temporary work, and 6 women (1.2%) had married and moved to other counties during the 10 years of follow-up.

The 98 deaths observed during follow-up represented a mortality rate of 2228 per 100,000 person-years (95% CI 1792–2664). Among men the mortality rate was 2913 (95% CI 2174–3652) and among women it was 1661 (95% CI 1150–2172). The rate was significantly higher in male than in female cohort members (hazard ratio 2.0, 95% CI 1.3–3.2, P < 0.005). The standardised mortality ratio for the whole cohort was 4.0 (95% CI 2.4–5.8); for men it was 4.9 (95% CI 2.8–8.1) and for women it was 3.3 (95% CI 1.9–6.1).

Among those who died in the follow-up period, 21 (21.4%) took their own lives, representing a suicide rate of 477 per 100,000 person-years (95% CI 273–681). Among men, the suicide rate was 753 per 100,000 person-years (95% CI 373–1133) and among women it was 249 per 100,000 person-years (95% CI 50–448). The rate was significantly higher in male than in female cohort members (HR = 3.1, 95% CI 1.2–8.0, P < 0.005). The standardised mortality rate for the cohort members who died by suicide was 32.0 (95% CI 18.5–52.5), for men it was 63.5 (95% CI 43.6–94.5) and for women it was 13.4 (95% CI 6.2–32.8). The 13 people who died as a result of accidents during the follow-up period represented a mortality rate due to accident of 296 per 100,000 person-years (95% CI 135–457). The standardised mortality ratio for the people who died by accidents was 6.6 (95% CI 4.3–10.2). The 64 natural deaths during follow-up period represented a mortality rate due to natural causes of 1455 per 100,000 person-years (95% CI 1101–1809). The standardised mortality ratio for all the cohort members who died from natural causes was 2.6 (95% CI 1.7–4.1). Among these 64 deaths, the specific causes of death in 42 cases (66%) were known (various cancers, n = 9; heart disease, n = 7; respiratory disease, n = 7; other disease, n = 19).

The survival probability for the whole cohort in 2004 was 0.80 (95% CI 0.76–0.84). Compared with women, the survival rate of men during the 10-year follow-up was significantly lower (survival probability in 2004 for women, 0.84, 95% CI 0.80–0.88; survival probability in 2004 for men, 0.74, 95% CI 0.68–0.80; log-rank test χ² = 7.85, P < 0.01).

Mortality rates by clinical and demographic characteristics and results of the univariate and multivariate Cox regression analyses are presented in Table 3. The independent predictors of mortality identified in the final model were age at onset of schizophrenia (> 45 years), duration of illness (> 10 years), age greater than 30 years, physical illness, inability to work, male gender, and never having received treatment. None of the variables excluded by the backwards procedure had a confounding effect.

**DISCUSSION**

To our knowledge this is the first 10-year prospective cohort study on mortality among people with schizophrenia in rural China. It includes longitudinal follow-up and analyses based on time-dependent factors. The strengths of our study include the use of a large representative community sample in rural China, its longitudinal 10-year follow-up design, and high rates of follow-up. Although studies of new-onset illness may accurately estimate suicide risk during the initial years of the illness (Palmer et al., 2005), study of a representative sample that includes both new-onset and chronic cases may better capture the death risk for the population of people with schizophrenia living in the community. As in most previous studies in China (Zhang et al., 1998; Phillips et al., 2004), all participants were 15 years or older at baseline in this study. It would not have substantially affected our results if children under 15 years old (a small group, representing less than 1% of all persons with schizophrenia) had been included.

**Mortality and suicide**

Compared with previous studies (Harris & Barraclough, 1997; Osby et al., 2000; Harrison et al., 2001), our results showed that mortality and suicide rates are relatively high among people with schizophrenia, both men and women, in rural China. The overall mortality rate of 2228 per 100,000 person-years observed in our study is extremely high, exceeding by 4 times the rate observed among people over 15 years old in the general population. This result is consistent with other results from low- and middle-income countries (Mojtabai et al., 2001). The suicide rate that we observed (477 per 100,000 person-years) is close to the estimated rate in a previous study of people with schizophrenia in China (0.68% per year; Phillips et al., 2004). Standardised mortality ratios were 32.0 for all suicide and 63.5 for males, which is much higher than in other countries (Osby et al., 2000). After excluding homeless individuals whose data were unavailable, we calculated the risk of suicide during the follow-up period as 4.5%, which is close to the 4.9% risk found in a meta-analysis by Palmer et al. (2003). Our study also indicated that the direct use of proportionate mortality rates in previous studies assuming a constant rate of suicide over a lifetime may overestimate suicide risk (Caldwell & Gottesman, 1990).

**Mortality and prognosis**

Although some authors have suggested that there is a better prognosis for individuals with schizophrenia in low- and middle-income countries (Left et al., 1992), our study showed a higher rate of mortality and missing due to homelessness among people with this disorder. It may be premature to suggest that there is a better prognosis for schizophrenia in these countries if withdrawals or attrition due to death and homelessness are not included in follow-up analyses. If the deceased and homeless cases were included in such analyses, the picture might change significantly. Deaths and homelessness among people with schizophrenia should be explored more in future natural history studies of this illness.

**Mortality and prevalence**

Different mortality rates in men and women in our study may explain the unique phenomenon of schizophrenia being more prevalent in women than in men in
Table 3  Mortality rates and Cox regression analyses of mortality (n=500)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Deaths n</th>
<th>Mortality rate per 100 000 person-years</th>
<th>Hazard ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at onset &gt; 45 years</td>
<td>34</td>
<td>6415</td>
<td>9.2 (4.9–17.1)</td>
</tr>
<tr>
<td>No</td>
<td>64</td>
<td>1654</td>
<td></td>
</tr>
<tr>
<td>Duration of illness ≥ 10 years</td>
<td>55</td>
<td>2517</td>
<td>5.1 (2.5–10.4)</td>
</tr>
<tr>
<td>Yes</td>
<td>43</td>
<td>1942</td>
<td></td>
</tr>
<tr>
<td>Age &gt; 50 years</td>
<td>64</td>
<td>2232</td>
<td>4.8 (2.5–9.1)</td>
</tr>
<tr>
<td>No</td>
<td>34</td>
<td>2062</td>
<td></td>
</tr>
<tr>
<td>Physical illness</td>
<td>44</td>
<td>4015</td>
<td>2.7 (1.8–4.2)</td>
</tr>
<tr>
<td>Yes</td>
<td>54</td>
<td>1635</td>
<td>2.1 (1.4–3.3)</td>
</tr>
<tr>
<td>Inability to work</td>
<td>39</td>
<td>5571</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>59</td>
<td>1595</td>
<td></td>
</tr>
<tr>
<td>Male gender</td>
<td>58</td>
<td>2913</td>
<td>2.0 (1.3–3.2)</td>
</tr>
<tr>
<td>No</td>
<td>40</td>
<td>1661</td>
<td></td>
</tr>
<tr>
<td>Never received treatment</td>
<td>40</td>
<td>3028</td>
<td>1.5 (1.0–2.4)</td>
</tr>
<tr>
<td>No</td>
<td>58</td>
<td>1884</td>
<td></td>
</tr>
<tr>
<td>Without income¹</td>
<td>55</td>
<td>4475</td>
<td>1.5 (0.9–2.6)</td>
</tr>
<tr>
<td>Yes</td>
<td>43</td>
<td>1357</td>
<td></td>
</tr>
<tr>
<td>Family history of mental disorder²</td>
<td>19</td>
<td>1501</td>
<td>1.5 (0.9–2.5)</td>
</tr>
<tr>
<td>Yes</td>
<td>79</td>
<td>2522</td>
<td></td>
</tr>
<tr>
<td>Married³</td>
<td>70</td>
<td>2169</td>
<td>1.3 (0.9–2.1)</td>
</tr>
<tr>
<td>No</td>
<td>28</td>
<td>2389</td>
<td></td>
</tr>
<tr>
<td>Living in shabby or unstable house¹</td>
<td>10</td>
<td>2674</td>
<td>1.3 (0.7–2.6)</td>
</tr>
<tr>
<td>Yes</td>
<td>88</td>
<td>2186</td>
<td></td>
</tr>
<tr>
<td>Number of family members ≤ 3²</td>
<td>48</td>
<td>2081</td>
<td>1.2 (0.8–1.9)</td>
</tr>
<tr>
<td>Yes</td>
<td>50</td>
<td>2390</td>
<td></td>
</tr>
<tr>
<td>Education ≤ 7 years</td>
<td>81</td>
<td>2342</td>
<td>1.2 (0.7–2.1)</td>
</tr>
<tr>
<td>No</td>
<td>17</td>
<td>1809</td>
<td></td>
</tr>
<tr>
<td>Family economic status below mean²</td>
<td>58</td>
<td>2344</td>
<td>1.2 (0.8–1.7)</td>
</tr>
<tr>
<td>Yes</td>
<td>40</td>
<td>2078</td>
<td></td>
</tr>
<tr>
<td>With marked symptoms³</td>
<td>67</td>
<td>2545</td>
<td>1.2 (0.8–1.8)</td>
</tr>
<tr>
<td>Yes</td>
<td>67</td>
<td>2545</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>31</td>
<td>1755</td>
<td></td>
</tr>
<tr>
<td>Previous hospitalisation²</td>
<td>20</td>
<td>1923</td>
<td>1.2 (0.7–1.9)</td>
</tr>
<tr>
<td>Yes</td>
<td>78</td>
<td>2322</td>
<td></td>
</tr>
<tr>
<td>Lives alone³</td>
<td>23</td>
<td>2975</td>
<td>1.1 (0.7–1.8)</td>
</tr>
<tr>
<td>No</td>
<td>75</td>
<td>2068</td>
<td></td>
</tr>
<tr>
<td>Previous suicide attempts²</td>
<td>12</td>
<td>2013</td>
<td>1.1 (0.6–2.0)</td>
</tr>
<tr>
<td>Yes</td>
<td>86</td>
<td>2261</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Considered for the multivariate model (P ≤0.10 in univariate analyses) but not retained.
2. Not considered for the multivariate model (P >0.10 in univariate analysis).

China. First, given that men have an earlier age at onset than women (Ran et al, 2003), the results of our study indicate that the higher mortality rate in men might be the major reason for the higher prevalence of schizophrenia in women than in men in China. Higher survival rates for female patients compared with male patients also support this opinion. This result may also explain in part why one finds more women among geriatric and late-onset patients with schizophrenia (Ran et al, 2004). Second, the results also indicated that women with schizophrenia were more likely to have married and moved away from the area. Even though the number is small, it is possible that female patients move to cities, which may increase the prevalence rates for women in urban areas. Third, our previous study did not support differences in the full remission rates between men and women (Ran et al, 2003).

All the risk factors identified in this study reflect the influence of both the socio-economic characteristics of rural China and the clinical characteristics of the cohort. Gender was an independent predictor of death during the follow-up period. Male gender was associated with an increased risk of death in this group, which is consistent with previous studies (Salokangas et al, 2002); the survival advantage of women held also among these participants. Why did the men have a significantly higher mortality rate than the women? Possible reasons may be that oestrogen has a protective effect in women (Seeman & Lang, 1990), or that men with schizophrenia may accept less support and treatment than female patients in rural China (Ran et al, 2003). These possibilities warrant further study.

Mortality and age

Age (> 50 years) was an independent predictor of mortality during follow-up, as one would expect – older people with schizophrenia are much more likely to die (Palmer et al, 2005). Although young patients, early in the course of the illness, are more likely to attempt suicide than older patients, more patients may die from other causes with increasing age; the proportion who died by suicide is relatively small among all age groups. Our study showed that the mortality rate was significantly higher among people with later onset of schizophrenia (> 45 years) than among those with onset before 45 years of age. Although evidence indicates that individuals
with a later onset may have a more benign illness course, symptom severity and cognitive deficits may be similar in both early-onset and late-onset cases (Jeste et al., 1995). Although the finding that older patients and those with later-onset disorder have higher mortality rates may not relate to lack of treatment, our study still suggests that higher mortality rates may be associated with the poor treatment received by older patients in rural China (Ran et al., 2004).

Mortality and treatment
Evidence indicates that a significant proportion of treated incident cases of schizophrenia achieve favourable long-term outcome (Harrison et al., 2001). Suicide risk among patients with schizophrenia-spectrum disorders declines quickly after treatment and recovery (Qin & Nordentoft, 2005). The results of our study indicated that never receiving treatment might increase the mortality rate among people with schizophrenia and, conversely, that treatment reduced mortality risk. In a 17-year follow-up study, the number of antipsychotic drugs taken by patients with schizophrenia showed a graded relation to mortality (Joukamaa et al., 2006). Given that certain classes of antipsychotic have been associated with death (Montout et al., 2002), we suggest that basic medication is important in decreasing the mortality and that excessive antipsychotic administration (e.g. overdose, multiple medication) may increase the mortality. The relationship between antipsychotics and mortality needs further study.

The results of our study indicate that long duration of illness and inability to work, which reflect poor social functioning, may increase the risk of mortality. The finding that previous hospitalisation is not a risk factor for mortality may be related to the lower rate of hospitalisation in rural China. Evidence indicates that people with schizophrenia have high rates of potentially reversible medical morbidity that additionally increase mortality (Green et al., 2003; Goff et al., 2005). In our study, over a quarter of the sample had physical illness, which predicted increased mortality. We suggest that treating medical comorbidity might reduce premature mortality among these patients (Goff et al., 2005).

Implications for services
Our results have implications for reducing mortality and suicide rates among people with schizophrenia in China and elsewhere. The risk factors of mortality should be taken into account when developing interventions to prevent premature death among these patients. Suicide prevention strategies should also be developed. Given the limited resources in contemporary China, prevention programmes should emphasise community-based mental healthcare to provide earlier diagnosis, antipsychotic treatment, treatment of comorbid medical conditions, function rehabilitation and family support. Given the severe stigma associated with psychiatric illness (Xiang et al., 1994), efforts to reduce stigma in the community will be necessary to enable individuals with schizophrenia to rejoin their community and allow interventions to be made to decrease their mortality rate.

Our findings indicate that homelessness among people with schizophrenia is a common phenomenon in rural China. Given that homeless individuals might experience mortality and suicide rates much higher than those of their counterparts in the general population (Roy et al., 2004), we suggest that suicide rates among people with schizophrenia might have been underestimated in previous studies in China (Phillips et al., 2004) because homeless people with the disorder were not included. Community-based mental health services, especially family and housing services, should be developed to prevent patients becoming homeless.

Given the representative sample used in this study, we are confident that our findings are generalisable to the population of people with schizophrenia in rural areas in China, and even to other countries that have a similar social environment. Premature death, suicide and homelessness are serious problems in people with schizophrenia in rural China. Supplying community mental health services and medication to these people should be a mental healthcare priority to prevent these early deaths.

ACKNOWLEDGEMENTS

The 1994 Chengdu study was supported by a grant from the China Medical Board in New York (CMB 92-537). This work was supported in part by the Hong Kong Jockey Club Centre for Suicide Research and Prevention, University of Hong Kong, and by International Clinical Operational and Health Services Research Training Award D43 TW05814 (E.D.C.) and Global Research Initiative Program 1 R01 TW007260-01 (M.S.R.) from the Fogarty International Center, National Institutes of Health, USA.

The authors thank the Chengdu Mental Health Center (Wen-Jun Mao, MD, and Shi-Hui Hu, MD) and the Xinjin Mental Hospital (Cui-Ping Tang, RN, Fu-Rong Lin, MD, Li Li, MD and Li Si-Gan, MD) for their collaboration.

REFERENCES


SCHIZOPHRENIA MORTALITY IN CHINA

The 1994 Chengdu study was supported by a grant from the China Medical Board in New York (CMB 92-537). This work was supported in part by the Hong Kong Jockey Club Centre for Suicide Research and Prevention, University of Hong Kong, and by International Clinical Operational and Health Services Research Training Award D43 TW05814 (E.D.C.) and Global Research Initiative Program 1 R01 TW007260-01 (M.S.R.) from the Fogarty International Center, National Institutes of Health, USA.

241


RAN ET AL

Correspondence: Dr Mao-Sheng Ran, College of Professional Studies, University of Guam, Mangilao, Guam 96923, USA. Tel: +1 671 735 2635; fax: +1 671 734 1203; email: ranmaosh@yahoo.com

(First received 8 April 2006, final revision 12 September 2006, accepted 27 October 2006)
Assessing insight in schizophrenia: East meets West

BALASUBRAMANIAN SARAVANAN, K. S. JACOB, SHANTHI JOHNSON, MARTIN PRINCE, DINESH BHUGRA and ANTHONY S. DAVID

Background Lack of insight has been observed in people with schizophrenia across cultures but assessment of insight must take into account prevailing illness models.

Aims To determine whether culturally specific and Western biomedical interpretations of insight and psychosis can be reconciled.

Method Patients with schizophrenia (n=131) were assessed during their first contact with psychiatric services in Vellore, South India. Patients’ explanatory models, psychopathology and insight were investigated using a standard schedule translated into Tamil.

Results Supernatural explanations of symptoms were frequent. Some insight dimensions were weakly associated (inversely) with severity of symptoms whereas preserved insight was associated with anxiety, help-seeking and perception of change. Willingness to attribute symptoms to disease, in others and in one’s self, but not to supernatural forces was strongly associated with insight.

Conclusions The relationship between insight, awareness of illness and other clinical variables is similar in South India to elsewhere. However, the assessment of insight might have failed to capture locally accepted explanatory frameworks. An inclusive conceptual model which emphasises help-seeking is recommended.

Declaration of interest None. Funding detailed in Acknowledgements.

Insight is a complex multidimensional construct which is shaped by individual psychology (i.e., motivation and denial) and the constraints of biology (as in cognitive impairment and anosognosia) and is influenced by social constructions of illness and culturally specific explanatory models (Amador & David, 2004). Lack of insight was found to be almost invariably associated with a diagnosis of acute schizophrenia across all countries and cultures surveyed within the World Health Organization International Pilot Study of Schizophrenia (Amador et al, 1991). Standardised tools for the assessment and quantification of insight have been developed over the past 15 years (e.g., the Schedule for the Assessment of Insight (SAI; David, 1990; Sanz et al, 1998) and the Scale to Assess Unawareness of Mental Disorder (SUMD; Amador et al, 1993)). These have been found to have clinical utility for diverse populations and patient groups worldwide, with little modification besides translation (Saravanan et al, 2003).

There is a consensus about the nature of insight emerging from systematic reviews and meta-analyses. For example, there is a weak but consistent inverse relationship between psychopathology and insight, with the exception of anxiety and low mood, which are positively associated with insight (Mintz et al, 2003; David, 2004). Intellectual ability, particularly executive functioning, seems to be related to insight (Keshavan et al, 2004; Morgan & David, 2004; Aleman et al, 2006). Another frequent observation is that patients are more able to recognise and label certain behaviours as the consequence of mental illness in others than they are the same behaviours in themselves (McEvoy et al, 1993; Swanson et al, 1995; Chung et al, 1997; Startup, 1997). However, some authors (Moodley & Perkins, 1993; Johnson & Orrell, 1995) have questioned what they regard as the Western conceptualisation of insight, arguing that it is overly biomedical and fails to allow for social constructions and culturally appropriate explanatory models of mental illness (for a discussion see Saravanan et al, 2004).

In this study we investigated the effect of culture, psychopathology and other clinical variables on insight of patients with schizophrenia in South India to determine which aspects are common across cultures and which are culture-specific. We tested the following hypotheses: (a) the relationship between insight and psychotic and depressive symptoms is similar to that in Western populations assessed in the same way; (b) a tendency to ascribe illness to another rather than oneself is inversely related to insight; (c) explanatory models, elicited in a standardised manner, are independent of clinician-rated insight.

METHOD

Study site

This study was carried out in the Department of Psychiatry, Christian Medical College, Vellore, which is in the north central part of Tamil Nadu. The total area of the Vellore district is 4314.29 km² and is divided into 12 blocks with a total population of 3,026,432. The Department of Psychiatry and Community Health has worked within the Kaniyambadi Block for the past 40 years. The town of Vellore (10.54 km²) has a population of 175,061. The 100-bed hospital provides short-term care for patients with all types of psychiatric diagnoses from the town of Vellore and a much wider rural area beyond. The emphasis is on a multidisciplinary approach and eclectic care using a wide variety of pharmacological and psychological therapies. The hospital has a daily out-patient clinic in which 200–250 patients are seen. The study was approved by the ethics committees of the Christian Medical College, Vellore, and Institute of Psychiatry, London.

Sample

The study group consisted of patients with schizophrenia having their first contact with mental health services and living within a 100 km radius of the study site. Patients were carefully screened for a DSM-IV diagnosis of schizophrenia (American Psychiatric Association, 1994) and then interviewed at intake using the Structured Clinical Interview for DSM-III-R—Patient
Version (SCID–P; Spitzer et al, 1990) to confirm the diagnosis. Patients with a primary diagnosis of substance use disorder, mood disorder or organic mental disorder were excluded. Patients meeting inclusion criteria and providing written consent were interviewed as soon as possible after the start of treatment for psychosis, with all patients assessed within a week of the onset of treatment. Patients were informed that the purpose of the study was to assess their level of awareness about their illness, but were unaware of the specific hypotheses. A semi-structured interview was used to elicit data regarding demographic characteristics (age, gender, marital status, religion, education, employment, duration of illness and economic data).

Assessments

All assessment instruments were independently translated into the local language (Tamil) by two health professionals. The vernacular version thus obtained was then back-translated into English by two different bilingual professionals. The four translators then arrived at a consensus on the final vernacular version. Content, semantic, technical and conceptual equivalence of the Tamil version of the instruments was examined regularly during the process of translation. Assessments were administered by a local research psychiatrist trained in their use.

Insight

The expanded version of the Schedule of Assessment of Insight (SAI–E; Kemp & David, 1997; Sanz et al, 1998) was used for assessment of insight. This has been applied widely in Western and non-Western countries (Kulhara et al, 1992; Aga et al, 1995) and comprises questions to assess three dimensions of insight: awareness, re-labelling of symptoms and adherence, plus a ‘hypothetical contradiction’ item added to evaluate the person’s capacity to consider another’s perspective. Each dimension comprises two or three questions which are scored on a 3-point scale from 0 (no insight) to 2 (good insight), with a maximum total score of 24. The supplementary question is scored from 0 to 4 and this is added to the total score. This expanded version also includes items on awareness of change, difficulties resulting from the mental condition and insight into key symptoms.

Explanatory model interview

The Tamil version of the Short Explanatory Model Interview (SEMI; Lloyd et al, 1998; Joel et al, 2003) was used to elicit patients’ attributions of their presenting complaints; their previous help-seeking behaviour (including visiting a temple, a shaman/mantrawadi, a traditional healer, or a doctor); their causal models (e.g. previous deeds/karma, evil spirits, punishment by god, black magic, or disease); perceived consequences (change in the body or mind); and their expectations regarding the index consultation. The SEMI, which combines open-ended questions and a case vignette with a structured coding frame, has been used successfully in a variety of countries and cultures, including India (Manoharam et al, 2001). The vignette described a young man with florid psychotic symptoms (see Appendix). Participants were asked the same range of questions regarding illness attribution for the vignette and the help-seeking they would advise for such a person.

Psychopathology

The Brief Psychiatric Rating Scale (BPRS; Ventura et al, 1993) and the Global Assessment of Functioning (GAF; Endicott et al, 1976) were used to assess psychopathology. The BPRS total score and a depression subscale (investigating guilt, low mood and suicidality) were utilised.

RESULTS

Over a 1-year recruitment period, 196 patients with schizophrenia attended the Department of Psychiatry, Christian Medical College, Vellore, and 188 met the entry criteria. Of these, 37 were excluded because the severity of psychopathology precluded an interview, 14 did not attend and 6 refused consent, yielding a final sample of 131 participants. Demographic and clinical characteristics of the sample are given in Table 1. The vast majority were Hindu, with a predominance of young men living in rural areas, who had received formal education but were unemployed. Data were analysed with the Statistical Package for the Social Sciences, version 10.0. Age, gender, years of education and length of illness were not significantly correlated with BPRS and SAI–E total scores. There was no significant difference in insight between patients from urban and rural areas (mean total SAI–E score 5.5, s.d. 4.5); however, patients who came voluntarily had significantly higher scores (mean 7.0, s.d.–6.6) than those who attended involuntarily (mean 4.3, s.d. = 4.1; P = 0.02).

Table 1  Characteristics of 131 patients with schizophrenia

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>mean (s.d.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>29.5 (7.2)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>72 (55)</td>
</tr>
<tr>
<td>Female</td>
<td>49 (45)</td>
</tr>
<tr>
<td>Religion, n (%)</td>
<td></td>
</tr>
<tr>
<td>Hindu</td>
<td>115 (87.8)</td>
</tr>
<tr>
<td>Muslim</td>
<td>4 (3.1)</td>
</tr>
<tr>
<td>Christian</td>
<td>11 (8.4)</td>
</tr>
<tr>
<td>Residence, n (%)</td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>105 (80.2)</td>
</tr>
<tr>
<td>Urban</td>
<td>26 (19.8)</td>
</tr>
<tr>
<td>Literacy, n (%)</td>
<td></td>
</tr>
<tr>
<td>Illiterate</td>
<td>22 (16.8)</td>
</tr>
<tr>
<td>Read only</td>
<td>9 (6.9)</td>
</tr>
<tr>
<td>Read and write</td>
<td>100 (76.3)</td>
</tr>
<tr>
<td>Age at onset of illness, years:</td>
<td></td>
</tr>
<tr>
<td>mean (s.d.)</td>
<td>27.8 (6.85)</td>
</tr>
<tr>
<td>Duration of illness, weeks:</td>
<td></td>
</tr>
<tr>
<td>mean (s.d.)</td>
<td>95.5 (134.2)</td>
</tr>
<tr>
<td>Status, n (%)</td>
<td></td>
</tr>
<tr>
<td>Involuntary patient</td>
<td>115 (87.8)</td>
</tr>
<tr>
<td>Voluntary patient</td>
<td>16 (12.2)</td>
</tr>
<tr>
<td>GAF score: mean (s.d.)</td>
<td>28.7 (8.19)</td>
</tr>
<tr>
<td>BPRS total score: mean (s.d.)</td>
<td>56.7 (5.2)</td>
</tr>
<tr>
<td>SAI–E total score: mean (s.d.)</td>
<td>4.7 (4.57)</td>
</tr>
</tbody>
</table>

GAF, Global Assessment of Functioning; BPRS, Brief Psychiatric Rating Scale; SAI–E, Schedule for the Assessment of Insight – Expanded Version.

Insight and psychopathology

There was no significant correlation between BPRS total score and SAI–E total score (Pearson’s r = –0.13, P = 0.1), general awareness (r = –0.05) or treatment adherence sub-scales (r = –0.15). However, a significant negative correlation was observed between the re-labelling dimension of the SAI–E and BPRS total score (r = –0.2, P = 0.04). Of the individual items on the BPRS, anxiety showed the strongest (positive) correlation with SAI–E total score and with the illness awareness items (r = 0.25 and 0.28 respectively, P < 0.01).
Help-seeking

Adjusted using conditional logistic regression with gender, age, years of schooling, location of residence (rural vs. urban) and marital status (currently married v. not currently married).

Table 2 Relationship between insight and explanatory models variables

<table>
<thead>
<tr>
<th>SEMI questions</th>
<th>Total number of responses</th>
<th>SAI–E total score Mean (s.d.)</th>
<th>F (P)¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black magic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>94</td>
<td>4.1 (4.0)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>26</td>
<td>7.7 (5.6)</td>
<td>9.0 (.0001)</td>
</tr>
<tr>
<td>Not sure</td>
<td>11</td>
<td>2.5 (2.6)</td>
<td></td>
</tr>
<tr>
<td>Previous deeds</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>12</td>
<td>5.7 (4.8)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>78</td>
<td>5.8 (5.0)</td>
<td>9.8 (.001)</td>
</tr>
<tr>
<td>Not sure</td>
<td>41</td>
<td>2.2 (1.9)</td>
<td></td>
</tr>
<tr>
<td>Evil spirits</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>22</td>
<td>3.6 (4.3)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>80</td>
<td>5.7 (4.9)</td>
<td>6.7 (.002)</td>
</tr>
<tr>
<td>Not sure</td>
<td>29</td>
<td>2.5 (2.0)</td>
<td></td>
</tr>
<tr>
<td>Punishment by god</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>13</td>
<td>4.0 (2.7)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>71</td>
<td>5.8 (4.9)</td>
<td>6.3 (.003)</td>
</tr>
<tr>
<td>Not sure</td>
<td>47</td>
<td>3.0 (3.8)</td>
<td></td>
</tr>
<tr>
<td>Disease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>18</td>
<td>10.4 (6.1)</td>
<td>21.4 (.001)</td>
</tr>
<tr>
<td>No</td>
<td>64</td>
<td>3.7 (3.4)</td>
<td></td>
</tr>
<tr>
<td>Not sure</td>
<td>49</td>
<td>3.8 (3.8)</td>
<td></td>
</tr>
<tr>
<td>Social dysfunction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>110</td>
<td>5.1 (4.8)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>13</td>
<td>2.9 (2.5)</td>
<td>3.1 (.05)</td>
</tr>
<tr>
<td>Not sure</td>
<td>8</td>
<td>1.7 (1.4)</td>
<td></td>
</tr>
<tr>
<td>Occupational dysfunction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>124</td>
<td>4.8 (4.6)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>5</td>
<td>2.1 (2.5)</td>
<td>1.2 (.3)</td>
</tr>
<tr>
<td>Not sure</td>
<td>2</td>
<td>2.0 (2.8)</td>
<td></td>
</tr>
<tr>
<td>Well-being affected</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>103</td>
<td>5.4 (4.7)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>8</td>
<td>2.1 (3.0)</td>
<td>6.1 (.003)</td>
</tr>
<tr>
<td>Not sure</td>
<td>20</td>
<td>2.1 (2.3)</td>
<td></td>
</tr>
</tbody>
</table>

SEMI, Short Explanatory Model Interview; SAI–E, Schedule for Assessment of Insight — Expanded version.
1. One-way analysis of variance (ANOVA).

Explanatory models

Table 2 shows the relationship between insight and responses to the SEMI summarised according to major theme. The 18 patients who endorsed a disease explanation had significantly higher scores on the SAI–E, indicating greater insight. Most patients acknowledged social (but not occupational) dysfunction and that their well-being had been affected; and these had higher insight scores than those who did not. There was a significant inverse correlation between non-biomedical beliefs and insight scores. Patients who held such beliefs, most commonly regarding black magic ($n=94$), but also those who attributed their problems to evil spirits, previous deeds and punishment by god, tended to have lower insight scores. Patients were frequently unsure or unable to give a definite attribution of their illness and this tended to be associated with low insight scores.

Awareness of illness — case vignette

Patients were more likely to judge the person described in the vignette as having an illness or disease ($n=63$) than they were to make this attribution about themselves ($n=18$; $\chi^2=13.9$, d.f.=1, $P<.001$), excluding those making attributions to neither ($n=66$). Slightly fewer patients judged the person described in the vignette to be suffering from the effects of black magic ($n=81$, 61.8%) than they themselves ($n=94$, 71.8%; 95% CI — 21.3 to 1.4, NS).

Factors associated with insight

A stepwise conditional logistic regression analysis was performed with total insight score as the dependent variable, entering psychopathology, endorsed explanatory models, pathways to care and other clinical variables, controlling for age, gender and education and other demographic factors. This identified causal attribution (disease as a positive predictor and black magic as a negative predictor), the perceived consequences item of the SEMI and presence of help-seeking as factors significantly associated with insight (Table 3).

DISCUSSION

This study examined the relationship between insight, psychopathology and explanatory models of illness and showed that insight in patients with schizophrenia from
South India shows many of the same clinical, psychopathological and cognitive associations as in Western populations.

**Psychopathology**

There was some evidence to support the hypothesis that insight is related to the psychopathology of schizophrenia. We found a weak but significant inverse correlation between total BPRS score and the relabelling dimension of insight, and a positive relationship between insight total score and indices of anxiety and worry. Given that correlations were in opposite directions, it is not surprising that the overall correlation was non-significant, although the weight of psychosis items has usually outweighed the mood/anxiety items in previous studies (for review see David, 2004). Nevertheless this pattern is in line with studies carried out in both high- and low- and middle-income countries (Kulhara et al., 1992; Aga et al., 1995; Mintz et al., 2003; David, 2004). The strength of the correlations was modest, bolstering the argument that insight and psychopathology are not merely two sides of the same coin. Lower insight scores were seen in patients brought to the clinic involuntarily, as expected from previous studies (McEvoy et al., 1989; David et al., 1992). The lack of correlation with scores of social functioning is again in line with some (Barthó et al., 1988; Schwartz, 1998; Simon et al., 2004) but not all previous studies (Amador et al., 1994; Trauer & Sacks, 2002), and might reflect the particularly low levels of social functioning (and small variance) in this patient sample. As in previous studies, duration of illness and gender were not determinants of insight.

The association between insight and symptoms of anxiety – assessed somewhat crudely using the BPRS – may be predicted on the basis of previous work (Amador et al., 1996; Mintz et al., 2003; Rathod et al., 2005), but the direction of causality is by no means clear. The received wisdom is that anxiety and lowered mood are a reaction to illness awareness, but it is equally plausible that anxiety leads to a more self-critical attitude, akin to ‘depressive realism’. Whatever the explanation, this pattern of associations appears not to be culturally specific.

**Explanatory models**

Insight was also evaluated indirectly using a vignette that forms part of the SEMI. In line with our second hypothesis, a pattern was seen wherein the person in the vignette was readily labeled as having a mental illness whereas patients rarely used this attribution for themselves. This pattern has been observed in other vignette studies in the UK, USA and Asia (McEvoy et al., 1993; Swanson et al., 1995; Chung et al., 1997; Startup, 1999). The self-other comparison overcomes the criticism levelled at cultural comparisons in psychiatry that the ‘content’ of attributions, delusions, hallucinations etc. may differ but that this is a relatively trivial consideration. Previous vignette studies show that the ‘medical model’ is an available heuristic to people with major mental illness, including those in South India. This is not surprising given the high levels of literacy and rapid cultural transition – first through colonisation and subsequently through globalisation – affecting that part of the world. It is of course highly likely that complex illness attributions of the kind elicited have been around in Asia for centuries, but here we are talking about a narrower Western biomedical model. The pattern of responses in relation to self and others might be described as an example of self-serving bias, a common, perhaps universal, defensive cognitive style (Pronin et al., 2004). It is understandable that a person might be uncomfortable in ascribing to themself a diagnosis of mental illness, given the stigma that this carries and, perhaps, implications of blame. However, the alternative explanations of behaviour such as black magic or past misdeeds are not value free. The meaning and implications of mental illness within cultures are clearly varied, and we are undertaking further detailed qualitative work to explore this.

Finally, we tested the hypothesis that indigenous explanatory models or ‘emic perspectives’ would be independent of insight. The study showed that the expression of culturally appropriate illness attributions (e.g. evil spirits, black magic, etc.) was common. Adherence to such ideas was deemed, at best, not consistent with high levels of insight or, at worst, indicative of poor insight as judged by psychiatrists from the same culture. In the case of black magic, this illness attribution predicted significantly lower levels of insight in the multiple regression analysis after controlling for other factors. However, the same analysis revealed a strong association between awareness of the consequences of illness, early help-seeking and level of insight, elicited separately through the open-ended SEMI in a way that did not rely upon the conventions of a psychiatric interview. This supports the notion that insight or an awareness of illness that is adaptive in terms of appropriate help-seeking might easily be missed if the assessment of insight is undertaken in a rigid culturally insensitive way (Saravanan et al., 2004).

Our results compliment a recent study of second-generation minority ethnic groups in the UK of West African, Caribbean and Bangladeshi origin. This study found that there was an association between attribution of illness to supernatural phenomena – not invoked by the White UK comparison group – and poorer insight as assessed using the SEMI (McCabe & Priebe, 2004). We also noted a tendency for those who were unsure whether disease or supernatural forces accounted for their state to have low insight scores. This might imply that any attribution is better than none.

**Conclusions**

We have shown that insight is amenable to study in a non-Western setting using standard assessment instruments. Patients with schizophrenia, wherever they are, appear to vary in their willingness to accept an illness attribution for their state, but are more likely to do so if they are anxious and if considering someone else’s predicament rather than their own. However, the study supports a more inclusive conceptual model for assessing insight which emphasises help-seeking. This will be of value not just for clinical and research work in non-Western countries but for a richer understanding of the experience of the individual in their cultural context (White et al., 2000; McCabe & Priebe, 2004). Understanding the relationship between cultural variations of insight and prevailing belief systems will be a step forward in designing interventions to enhance engagement of patients with mental health services in Western and non-Western countries.

**ACKNOWLEDGEMENTS**

We thank the participants and staff of the Christian Medical College, Vellore, for their cooperation and commitment to the study. B.S. was supported by a grant from the Welcome Trust.

**APPENDIX**

**Case vignette**

I'd like to ask your opinion about some other people's visit to the doctor. I'd like to read a short
account of their problems and then ask you a few questions about them.'

Mr M.E. was hospitalised at the age of 18. He was single and lived with his parents. His relatives described that he functioned quite well as a child and that he was well adjusted at school until the age of 11–12. At that age he became preoccupied with strange ideas and for this reason he saw a psychiatrist weekly for about 1 month. This treatment, which did not include any medication, had a positive effect on his worries, but he became more withdrawn and participated in fewer social activities in the months to come. Before hospitalisation he felt that others could hear his thoughts and he also felt that a ‘satanic’ group living in his native place were persecuting him. In a mysterious way he felt that this group did ‘black magic’ against him and that they could influence his body from a long distance. He could feel this as a pain in his stomach. These symptoms lasted for several months. At the time of hospitalisation he felt that his brain was damaged and ‘empty’, and that ‘someone’ was inserting thoughts into his head. He was withdrawn and preoccupied with the idea that the ‘black magic’ that they had done might have destroyed his brain tissue.

REFERENCES


BALSUBRAMANIAN SARAVANAN, MD, Institute of Psychiatry, London and Christian Medical College, Vellore, India; K. S. JACOB, MD, PhD, SHANITI JOHNSON, Christian Medical College, Vellore, India; MARTIN PRINCE, MD, FRCPych, DINESH BHUGRA, PhD, FRCPych, ANTHONY S. DAVID, MD, FRCP, FRCPych, Institute of Psychiatry, King’s College London, London, UK

Correspondence: Professor Anthony David, Section of Cognitive Neuropsychiatry, PO Box 68, Institute of Psychiatry, King’s College London, Denmark Hill, London SE5 8AF, UK. Email: a.david@iop.kcl.ac.uk

(First received 1 August 2006, accepted 8 November 2006)
The Stigma Scale: development of a standardised measure of the stigma of mental illness

MICHAEL KING, SOKRATIS DINOS, JENIFER SHAW, ROBERT WATSON, SCOTT STEVENS, FILIPPO PASSETTI, SCOTT WEICH and MARC SERFATY

Background  There is concern about the stigma of mental illness, but it is difficult to measure stigma consistently.

Aims  To develop a standardised instrument to measure the stigma of mental illness.

Method  We used qualitative data from interviews with mental health service users to develop a pilot scale with 42 items. We recruited 193 service users in order to standardise the scale. Of these, 93 were asked to complete the questionnaire twice, 2 weeks apart, of whom 60 (65%) did so. Items with a test–retest reliability kappa coefficient of 0.4 or greater were retained and subjected to common factor analysis.

Results  The final 28-item stigma scale has a three-factor structure: the first concerns discrimination, the second disclosure and the third potential positive aspects of mental illness. Stigma scale scores were negatively correlated with global self-esteem.

Conclusions  This self-report questionnaire, which can be completed in 5–10 min, may help us understand more about the role of stigma of psychiatric illness in research and clinical settings.

Declaration of interest  None.

Stigma is the negative evaluation of a person as tainted or discredited on the basis of attributes such as mental disorder, ethnicity, drug misuse or physical disability (Goffman, 1963). There is no doubt that such prejudice has substantial negative social, political, economic and psychological consequences for stigmatised people (Dovidio et al., 2000). They may feel unsure of how ‘normal’ people will identify or receive them (Goffman, 1963) and become constantly self-conscious and calculating about what impression they are making (Rush, 1998).

A number of attempts have been made to measure attitudes to mental illness and stigma, most of which have focused on attitudes towards mental illness held by people in the community (Bhugra, 1989; Link et al., 1991; Ritchie et al., 1994; Wolff et al., 1996; Byrne, 1997; Corrigan et al., 2000, 2001). Far fewer attempts have been made to measure stigma directly with service users themselves. One instrument developed in the USA focused on stigma associated with seeking psychotherapy (Judge, 1998), and a second concerned the shame and withdrawal felt by people with mental illness (Link et al., 2001). After our study was completed, a fourth measure has been published in which a more comprehensive attempt was made to evaluate stigma using thoughts and opinions from focus groups of mental health users in the USA (Ritsher et al., 2003). Corrigan and colleagues (Corrigan, 2000, 2004; Corrigan & Watson, 2002) have extended their research on public attitudes to mental illness to include conceptual and methodological work on what they called self-stigma (i.e. the reactions of stigmatised individuals towards themselves) and on the perception of discrimination by people with mental illness (Corrigan et al., 2003; Rusch et al., 2005).

We aimed to design a standardised measure of the stigma of mental illness that is firmly anchored in the experiences and views of mental health service users, and then to test its relationship to a measure of self-esteem. We predicted that stigma and self-esteem would be negatively correlated.

METHOD

Participants and procedure

The study was approved by the local research ethics committee. We recruited 193 people with a range of psychiatric diagnoses and of varying age, gender and ethnicity from mental health user groups, day centres, crisis centres, out-patient departments and hospitals in North London. Service users were approached either by the researchers or by members of staff and were informed about the study and its aims, and then asked to participate. No exclusion criteria were used. Our aim was to recruit as many participants as possible from diverse psychiatric and demographic backgrounds. The requirements of ethical approval constrained any collection of data about potential participants who refused. Two service users (J.S. and R.W.) who had already received training in research methods in earlier work on this theme (Dinos et al., 2004) underwent further training to contribute to the questionnaire content, and to conduct further data collection. A proportion of participants completed the questionnaire on two occasions approximately 2 weeks apart.

Measures

We asked participants standard demographic questions, followed by questions about when they first experienced mental health problems, whether or not they had received a diagnosis from a mental health professional, the nature of any diagnosis, the time that the diagnosis was given and whether they agreed with it, treatment received and whether they had ever been admitted to hospital compulsorily. Participants then completed the following two questionnaires.

Stigma Scale

Forty-two questions on the stigma of mental illness were developed from the detailed, qualitative accounts of 46 mental health service users recruited in an earlier study (Dinos et al., 2004). Stigma was a pervasive concern for almost all of these 46 participants. People with psychosis or drug dependence were most likely to report feelings and experiences of stigma and were...
most affected by them. Participants with depression, anxiety or personality disorders were more concerned about patronising attitudes and often perceived stigma even if they had not experienced any overt discrimination. However, experiences were not universally negative, and people employed various strategies to protect their self-esteem and maintain a positive self-concept. The content of statements used in this study arose directly from these findings. Themes that were more salient than others because they appeared in most of the qualitative interviews – such as how to manage telling others about the illness – were given priority. Thus, items that were based on each of several different disclosure types were included in the scale. The 42 items covered all of the themes and sub-themes from these interviews. The wording of each item was based on participants’ phrases in the qualitative interviews, adapted with minor modifications to fit most people’s experiences. Participants indicated whether they agreed or disagreed with each of these 42 statements on a five-point Likert scale ranging from ‘Strongly agree’ to ‘Strongly disagree’. Response set bias was addressed by alternating between negative and positive wording. We chose a five-point Likert scale as a straightforward, widely used response style that avoided more difficult formats such as visual analogue scales and yet accurately reflected participants’ experiences.

**Self-Esteem Scale**

The Self-Esteem Scale (Rosenberg, 1965, 1979) has been shown to have high test–retest reliability and concurrent validity with a number of measures of psychological well-being and self-efficacy. Participants indicate whether they agree or disagree with ten statements on a five-point Likert scale ranging from ‘Strongly agree’ to ‘Strongly disagree’. Examples of statements are ‘On the whole I am satisfied with myself’ and ‘I feel that I have a number of good qualities’. The aim of including this questionnaire was to explore the relationship between perceived stigma and self-esteem. Although we expected scores on the two scales to be negatively correlated, we did not regard this as a validation of our stigma scale.

**Analysis**

We first examined the pattern and distribution of responses in order to detect items that had little variation in response and would therefore not distinguish between people with differing experiences of stigma. We examined the test–retest reliability of responses to the statements using the weighted kappa statistic and items with a weighted kappa coefficient below 0.4 were removed. Remaining items were subjected to a common factor analysis and subsequent oblique (promax) rotation as we assumed at least two factor scores would be correlated. We found, however, that the factor scores derived were not correlated and thus, as a sensitivity check, we also performed an orthogonal rotation which assumes no correlation between any two factors. We chose common factor analysis (in contrast to principal components analysis) because our primary purpose was to understand the factor structure of the instrument, rather than summarise or reduce the data. Common factor analysis enables an examination of simple patterns in the relationships among the statements. The scree plot of successive eigenvalues was inspected to identify the point where the plot abruptly levelled out, indicating that adding further factors would not help describe the overall relationship between the statements. Internal consistency of the final scale (and sub-scales) was estimated using Cronbach’s α. We also explored the correlation of each item with the total score (item excluded), the average correlation with other items and Cronbach’s α with that item removed. Concurrent validity with the Self-Esteem Scale was assessed by comparing mean scores using Pearson’s correlation coefficient. Data were analysed using Stata version 7 for Windows.

**RESULTS**

**Participants**

Altogether 193 service users took part. The first 93 were asked to complete the stigma questionnaire on two occasions; 60 (65%) of them complied and 33 completed it only once. The 60 patients who completed the questionnaire twice did not differ from the 33 who refused, in terms of their diagnoses, mean number of years since diagnosis or whether they had ever been compulsorily admitted to hospital. A further 100 participants agreed to complete the questionnaire once in order to boost the sample size for factor analysis. A total of 109 men and 82 women (2 respondents did not state their gender), whose mean age was 42.9 years (s.d. = 12.4, range 19–76), took part; 159 (76.5%) were White, 11 (5.5%) were Black, 7 (3.5%) were of Indian or Bangladeshi origin, 18 (9%) were of other origin and 11 did not state their ethnic background. Regarding occupation, 34 (17%) were employed, 68 (34%) were on sick leave from work, 40 (20%) were unemployed seeking work, 12 (6%) were students, 24 (12%) were retired, two (1%) were home managers and 20 were unable to answer the question. Most participants had received a diagnosis of schizophrenia, bipolar affective disorder, depression and/or mixed anxiety and depression (Table 1) and most had received more than one diagnosis; 135 patients (67.5%) agreed with their diagnoses, 36 did not, I was unsure and 21 did not answer the question. A third of participants (n=63) reported that they had been admitted to a psychiatric unit compulsorily (8 did not answer the question) and 26 (16%) reported having received electroconvulsive therapy.

**Distribution of responses**

Responses to all items were reasonably evenly distributed, in that each response

<table>
<thead>
<tr>
<th>Table I</th>
<th>Diagnoses and treatments reported by the 193 participants. More than one diagnosis or form of treatment could be reported</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnosis</td>
<td>n</td>
</tr>
<tr>
<td>Schizophrenia/schizoaffective disorder</td>
<td>52</td>
</tr>
<tr>
<td>Bipolar affective disorder</td>
<td>37</td>
</tr>
<tr>
<td>Mixed anxiety and depression</td>
<td>77</td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td>54</td>
</tr>
<tr>
<td>Drug problems</td>
<td>27</td>
</tr>
<tr>
<td>Alcohol problems</td>
<td>29</td>
</tr>
<tr>
<td>Personality disorder</td>
<td>23</td>
</tr>
<tr>
<td>Depression</td>
<td>94</td>
</tr>
<tr>
<td>OCD</td>
<td>12</td>
</tr>
<tr>
<td>Eating disorder</td>
<td>24</td>
</tr>
<tr>
<td>PTSD</td>
<td>13</td>
</tr>
<tr>
<td>Treatment</td>
<td></td>
</tr>
<tr>
<td>ECT</td>
<td>26</td>
</tr>
<tr>
<td>Antidepressants</td>
<td>146</td>
</tr>
<tr>
<td>Sleeping tablets</td>
<td>101</td>
</tr>
<tr>
<td>Tranquillisers</td>
<td>78</td>
</tr>
<tr>
<td>Counselling/CBT</td>
<td>111</td>
</tr>
<tr>
<td>Antipsychotics</td>
<td>86</td>
</tr>
<tr>
<td>Mood stabilisers</td>
<td>47</td>
</tr>
<tr>
<td>None</td>
<td>2</td>
</tr>
</tbody>
</table>

CBT, cognitive–behavioural therapy; ECT, electroconvulsive therapy; OCD, obsessive–compulsive disorder; PTSD, post-traumatic stress disorder.
The general public is understanding of people with mental health problems (D) 0.41
Other people have made me feel ashamed of myself because of my mental health problems (A) 0.38
The way people have treated me upsets me (A) 0.34
I have been discriminated against by housing departments/landlords because of my mental health problems (A) 0.38
I have been discriminated against in education because of my mental health problems (A) 0.60
Sometimes I feel that I am being talked down to because of my mental health problems (A) 0.42
Having had mental health problems has made me a more understanding person (D) 0.51
I am to blame for my mental health problems (A) 0.50
I feel ashamed of myself that I have had mental health problems (A) 0.38
I do not feel bad about having had mental health problems (D) 0.45
Other people think less of me because I have had mental health problems (A) 0.52
Newspapers/television take a balanced view about mental health problems (D) 0.24
I am open to my family about my mental health problems (D) 0.50
I worry about telling people I receive psychological treatment (A) 0.43
Some people with mental health problems are dangerous (A) 0.67
Other people have never made me feel embarrassed because of my mental health problems (D) 0.33
People have been understanding of my mental health problems (D) 0.45
I have been discriminated against by police because of my mental health problems (A) 0.64
I have been discriminated against by employers because of my mental health problems (A) 0.53
I have been physically threatened or attacked because of my mental health problems (A) 0.28
My mental health problems have made me more accepting of other people (D) 0.44
Very often I feel alone because of my mental health problems (A) 0.48
I am scared of how other people will react if they find out about my mental health problems (A) 0.45
I would have had better chances in life if I had not had mental health problems (A) 0.55
I am as good as other people, even though I have had mental health problems (D) 0.57
I do not mind people in my neighbourhood knowing I have had mental health problems (D) 0.55
I would say I have had mental health problems if I was applying for a job (D) 0.71
I worry about telling people that I take medicines/tablets for mental health problems (A) 0.58
People's reactions to my mental health problems make me keep myself to myself (A) 0.50
I am angry with the way people have reacted to my mental health problems (A) 0.59
I have not had any trouble from people because of my mental health problems (D) 0.51
I have been discriminated against by health professionals because of my mental health problems (A) 0.51
People have avoided me because of my mental health problems (A) 0.53
People have insulted me because of my mental health problems (A) 0.49
Having had mental health problems has made me a stronger person (D) 0.48
I do not feel embarrassed because of my mental health problems (D) 0.57
I avoid telling people about my mental health problems (A) 0.52
Having had mental health problems makes me feel that life is unfair (A) 0.58
When I see or read something about mental health in the papers or television, it makes me feel bad about myself (A) 0.53
I feel the need to hide my mental health problems from my friends (A) 0.49
I find it hard telling people I have mental health problems (A) 0.44
I do not understand the diagnosis I have been given (A) 0.64

1. Each question scored 0–4 in the direction of greater stigma: A, scored 0–4 in direction of agreement; D, scored 0–4 in direction of disagreement.
The questionnaire) produced similar results. Scale scores based on the 0–4 scoring of generated in the analysis (rather than sub-generated in the analysis (rather than sub-

A sensitivity analysis using factor scores. TURING SEPARATE ASPECTS OF STIGMA (Table). Overall stigma score than with each other, scores had higher correlations with the Stigma Scale Mean scores were as follows: Stigma Scale item deletion improved the internal reliability above 0.88. Cronbach’s x for the first sub-scale (discrimination) was 0.87, for the second (disclosure) 0.85 and for the third (positive aspects) 0.64.

Sub-scale scores
Mean scores were as follows: Stigma Scale 62.6 (s.d.=15.4), discrimination sub-scale 29.1 (s.d.=9.5), disclosure sub-scale 24.7 (s.d.=8.0) and positive aspects sub-scale 8.8 (s.d.=2.8). As expected, mean sub-scale scores had higher correlations with the overall stigma score than with each other, supporting the notion that they were capturing separate aspects of stigma (Table 5). A sensitivity analysis using factor scores generated in the analysis (rather than sub-scale scores based on the 0–4 scoring of the questionnaire) produced similar results.

Concurrent validity
Scores on the Self-Esteem Scale (high score indicates high self-esteem) were negatively correlated with the overall Stigma Scale core and sub-scale scores (Table 5).

Discussion
We have developed a brief self-report scale to measure the stigma of mental illness based directly on service users’ detailed accounts of their feelings and experiences of prejudice and discrimination (Dinos et al., 2004). We constructed more items than we thought would be needed in a final version and used assessments of reliability and consistency, as well as common factor analysis, to examine its underlying dimensions. The first factor or sub-scale explained much more of the variance (44%) than the other two factors and it could be argued that this might form the full scale. However, the principal aim of the factor analysis was to understand the latent dimensions of the instrument rather than reduce it further and we believe the dimensions found in the other two sub-scales are important in our understanding the complexity of stigma. The questionnaire takes 5–10 min to complete. Our scale is similar in content to that the Internalised Stigma of Mental Illness scale developed by Ritsher et al (2003). However, test–retest reliability of this latter scale remains uncertain as it was based on only 16 respondents.

Strengths and limitations
A major strength of our study is that the content of this stigma scale arose directly from earlier qualitative research into...
Table 4  Descriptive statistics of final 28 item stigma scale

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree</th>
<th>Agree</th>
<th>Neither agree nor disagree</th>
<th>Disagree</th>
<th>Strongly disagree</th>
<th>Responses</th>
<th>Mean (s.d.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5  I have been discriminated against in education because of my mental health problems (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>188</td>
<td>1.59 (1.03) 1.5</td>
</tr>
<tr>
<td>6  Sometimes I feel that I am being talked down to because of my mental health problems (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>189</td>
<td>2.40 (1.24) 3.0</td>
</tr>
<tr>
<td>7  Having had mental health problems has made me a more understanding person (P)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>190</td>
<td>1.08 (0.89) 1.0</td>
</tr>
<tr>
<td>10 I do not feel bad about having had mental health problems (D)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>188</td>
<td>2.32 (1.26) 3.0</td>
</tr>
<tr>
<td>14 I worry about telling people I receive psychological treatment (D)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>189</td>
<td>2.71 (1.18) 3.0</td>
</tr>
<tr>
<td>15 Some people with mental health problems are dangerous (P)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>190</td>
<td>2.82 (0.95) 3.0</td>
</tr>
<tr>
<td>17 People have been understanding of my mental health problems (P)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>185</td>
<td>1.84 (1.06) 2.0</td>
</tr>
<tr>
<td>18 I have been discriminated against by police because of my mental health problems (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>188</td>
<td>1.72 (1.21) 2.0</td>
</tr>
<tr>
<td>19 I have been discriminated against by employers because of my mental health problems (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>187</td>
<td>2.08 (1.16) 2.0</td>
</tr>
<tr>
<td>21 My mental health problems have made me more accepting of other people (P)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>191</td>
<td>1.19 (1.01) 1.0</td>
</tr>
<tr>
<td>22 Very often I feel alone because of my mental health problems (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>190</td>
<td>2.85 (1.14) 3.0</td>
</tr>
<tr>
<td>23 I am scared of how other people will react if they find out about my mental health problems (D)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>192</td>
<td>2.65 (1.13) 3.0</td>
</tr>
<tr>
<td>24 I would have had better chances in life if I had not had mental health problems (D)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>191</td>
<td>2.89 (1.15) 3.0</td>
</tr>
<tr>
<td>26 I do not mind people in my neighbourhood knowing I have had mental health problems (D)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>192</td>
<td>2.58 (1.34) 3.0</td>
</tr>
<tr>
<td>27 I would say I have had mental health problems if I was applying for a job (D)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>189</td>
<td>2.16 (1.31) 2.0</td>
</tr>
<tr>
<td>28 I worry about telling people that I take medicines/tablets for mental health problems (D)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>191</td>
<td>2.58 (1.18) 3.0</td>
</tr>
<tr>
<td>29 People's reactions to my mental health problems make me keep myself to myself (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>188</td>
<td>2.40 (1.19) 3.0</td>
</tr>
<tr>
<td>30 I am angry with the way people have reacted to my mental health problems (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>190</td>
<td>2.23 (1.18) 2.0</td>
</tr>
<tr>
<td>31 I have not had any trouble from people because of my mental health problems (Dc)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>192</td>
<td>2.24 (1.14) 2.0</td>
</tr>
<tr>
<td>32 I have been discriminated against by health professionals because of my mental health problems (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>189</td>
<td>1.95 (1.28) 2.0</td>
</tr>
<tr>
<td>33 People have avoided me because of my mental health problems (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>189</td>
<td>2.30 (1.18) 3.0</td>
</tr>
<tr>
<td>34 People have insulted me because of my mental health problems (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>192</td>
<td>2.01 (1.20) 2.0</td>
</tr>
<tr>
<td>35 Having had mental health problems has made me a stronger person (P)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>188</td>
<td>1.78 (1.23) 2.0</td>
</tr>
<tr>
<td>36 I do not feel embarrassed because of my mental health problems (D)</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>190</td>
<td>2.16 (1.22) 2.0</td>
</tr>
<tr>
<td>37 I avoid telling people about my mental health problems (D)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>191</td>
<td>2.68 (1.10) 3.0</td>
</tr>
<tr>
<td>38 Having had mental health problems makes me feel that life is unfair (Dc)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>191</td>
<td>2.53 (1.16) 3.0</td>
</tr>
<tr>
<td>40 I feel the need to hide my mental health problems from my friends (D)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>190</td>
<td>2.12 (1.22) 2.0</td>
</tr>
<tr>
<td>41 I would have had better chances in life if I had not had mental health problems (D)</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>191</td>
<td>2.70 (1.16) 3.0</td>
</tr>
</tbody>
</table>

D, disclosure; Dc, discrimination; P, positive aspects.

patients’ experiences of mental illness (Dinos et al., 2004). We do not suggest that this approach is superior to, or distinct from, one based on theoretical conceptions of perceived stigma; the items derived resonate with current theory about stigma. However, our instrument directly reflects the lived experience of stigma and may help us to extend our current theoretical concepts. Furthermore, data collection in this study was carried out by mental health service users, an approach which we hoped
would allow respondents to express their feelings frankly. Patients recruited were unselected and came from a variety of clinical and community settings. We did not examine how stigma varied with the demographic and clinical characteristics of participants, as they might not have been representative of all people with mental health problems. Thus, the instrument needs further evaluation in larger groups of patients in distinct diagnostic groups or in particular settings (such as in-patients) to understand its applicability. Furthermore, diagnoses and treatments were ascertained exclusively by self-report. Although the range of age, gender and diagnoses included indicates that we recruited a broad spectrum of mental health service users, the majority were White and hence the instrument needs further evaluation in a larger population of people from Black and minority ethnic populations. Three factors and 35 items mean that our sample size of 193 was adequate for the factor analysis. There is an inevitable element of subjectivity in the interpretation of the results of factor analysis and there may be other ways of describing the three factors arising. Whether the factor structure is consistent awaits confirmatory factor analysis in other populations. We confirmed our hypothesis that perceived stigma and self-esteem are negatively correlated. However, we stress that this analysis is exploratory and does not validate the stigma scale.

Forms of stigma

The distinction between stigma in the form of actual and feared discrimination is not new. Jacoby (1994) drew a distinction between ‘felt’ and ‘enacted’ stigma. Both may occur, regardless of whether or not the person feels any sense of personal shame or inferiority. Enacted stigma can be described as episodes of discrimination against people with mental illness. It can involve loss of job opportunities and negative reactions of family or friends, and it can also take the form of subtle, patronising attitudes and behaviours towards people with mental illness. The discrimination sub-scale contains items that refer to the negative reactions of other people, including acts of discrimination by health professionals, employers and police. As Jacoby (1994) emphasised, stigma may be also felt in the absence of any direct discrimination and may critically affect disclosure. It may not be possible for some people to conceal that they have a mental illness, but the key issue for the many who can is how to manage information about their condition (disclosure). Although ‘felt stigma’ is often used to refer to an internalised negative view of being mentally ill that leads to behaviours to hide it, reluctance to disclose is common without any attendant feelings of shame or embarrassment. Lack of disclosure may simply be the result of fear of what others will think, avoidance of unpleasant situations and a reluctance to invoke prejudice. Similar caution about disclosure in the absence of any personal shame is seen in other contexts, for example sexual orientation (Day & Shonbrun, 2000). Thus, we would take issue with an assumption (e.g. Corrigan et al., 2003; Ritsher et al., 2003; Rusch et al., 2005) that fear of disclosure is always the result of internalised stigma. As can be seen from the statements in our disclosure sub-scale, only two questions refer to embarrassment or feeling bad about the illness (items 10 and 36, Table 3) whereas the remainder refer to managing disclosure to avoid discrimination. Although the third factor, positive aspects of mental illness, contributed to less of the overall variance of the questionnaire items, it taps into how people accept their illness, become more open and make positive changes as a result, and lifts the mainly negative tone of the instrument. It is important to note that (given the direction of scoring shown in Table 4) high scores on this sub-scale indicate that the respondent perceives few positive outcomes from the illness. Its lower correlation with other parts of the scale suggests that people who do believe they are more empathetic human beings because of their illness may be less affected by stigma.

### Association with self-esteem

The relationship between stigma and self-esteem has been the focus of theoretical and empirical debates for decades: see Crocker & Major (1989) and Crocker & Wolfe (2001) for reviews. Unfortunately, the concept of stigma of mental illness has tended to rule out potential positive constructions of identity (e.g. Finlay et al., 2001; Camp et al., 2002; Dinos et al., 2005; Rusch et al., 2005, 2006). However, the majority of past studies were speculative in nature because there has not been a straightforward way to test the relationship between the two constructs (mainly because of lack of robust stigma scales). Scores on the Stigma Scale and its sub-scales were negatively correlated with global self-esteem, confirming our hypothesis that a negative relationship would be found between high self-esteem and high levels of perceived stigma. Ritsher et al. (2003) also reported that their new stigma scale and the Rosenberg Self-Esteem Scale were measuring distinct constructs. However, they did not report any direct correlation between their new scale and self-esteem. Our study is the only one, to our knowledge, that has developed a stigma scale and subsequently explored the relationship between self-esteem and stigma.

### Use of the Stigma Scale in clinical care and research

Stigma about mental illness may determine how and even whether people seek help for mental health problems, their level of engagement with treatment and the outcome of their problems (Hayward & Bright, 1997). This instrument now requires further assessment in clinical and research populations. We believe that it may contribute usefully to our understanding of processes that affect help-seeking, treatment uptake and outcome of mental illness.
ACKNOWLEDGEMENTS

We thank all the participants, staff in the services where the study was conducted and Mr Bob Blizard for his statistical advice. We also thank the reviewers for their helpful insights into the results and their meaning. This study was carried out with the support and collaboration of the Camden and Islington Mental Health and Social Care Trust.

REFERENCES


(First received 9 April 2006, final revision 25 September 2006, accepted 27 October 2006)
Risk of major adverse perinatal outcomes in women with eating disorders

NADIA MICALI, EMILY SIMONOFF and JANET TREASURE

Background Low birth weight, prematurity and higher miscarriage rates have previously been reported in women with eating disorders.

Aims To determine whether women with a history of eating disorders are at higher risk of major adverse perinatal outcomes.

Methods Adjusted birth weight, preterm delivery and miscarriage history were compared in those with a history of eating disorders (anorexia nervosa (n=171), bulimia nervosa (n=199) and both (n=82)) and those with other (n=1166) and no psychiatric disorders (n=10 636) in a longitudinal cohort study.

Results The group with bulimia nervosa had significantly higher rates of past miscarriages (relative risk ratio 2.0, P=0.01) and the group with anorexia nervosa delivered babies of significantly lower birth weight than the general population (P=0.01), which was mainly explained by lower pre-pregnancy body mass index. Preterm delivery rates were comparable across groups.

Conclusions Women with a history of eating disorders are at higher risk of major adverse obstetric outcomes. Antenatal services should be aware of this higher risk.

Declaration of interest None. Funding detailed in Acknowledgements.

Eating disorders are a common source of psychiatric morbidity in women of childbearing age (Van Hoeken et al, 2003). Previous studies on clinical samples have reported that women with anorexia nervosa, bulimia nervosa and women hospitalised for an eating disorder deliver lower birth weight and more preterm babies (Stewart et al, 1987; Brinich et al, 1988; Bulik et al, 1999; Waugh & Bulik, 1999; Soldid et al, 2004). Higher miscarriage rates have also been reported in women with eating disorders, especially those with bulimia nervosa (Mitchell et al, 1991; Abraham, 1998; Bulik et al, 1999; Blais et al, 2000).

No studies to date have determined, in an epidemiologically representative sample, whether the effect on adverse pregnancy outcomes is specific to the eating disorders and their symptoms, rather than to any severe psychiatric disorder. Moreover, most studies on women with eating disorders have not taken into account the effect of other mediating factors that may affect perinatal outcomes. In this study we investigated the effect of a history of eating disorders on the outcome of pregnancy in a representative sample of the British population.

METHOD

Sample

The Avon longitudinal Study of Parents and Children (ALSPAC) is a longitudinal, prospective study of women and pregnancy (Golding et al, 2001). All pregnant women living in the geographical area of Avon, UK, who were expected to deliver their baby between 1 April 1991 and 31 December 1992 were recruited. All women gave informed and written consent. It was estimated that 85–90% of those eligible took part. The sample has been shown to be representative of the British population.

There were 14 663 women enrolled at the 9th week of pregnancy. Data were obtained on 14 472 women via postal questionnaires. Women were excluded from the current study if they had not answered the questionnaire sent at approximately 12 weeks (2019). We only included singleton births in the study (12 234), as babies from multiple pregnancies have different patterns of foetal growth and gestational length. At 12 weeks women were also asked whether they had any recent or past history of psychiatric problems, including depression, schizophrenia, alcoholism, anorexia nervosa, bulimia nervosa or any other psychiatric disorder. Their pre-pregnancy weight and height were also obtained. Socio-demographic data were obtained during pregnancy. At 18 weeks of gestation information was obtained on vomiting and the use of laxatives for weight loss prior to and during pregnancy. Data on smoking and alcohol intake before and during the first and second trimesters of pregnancy were obtained at two time-points during pregnancy.

Body mass index (BMI) was calculated as pre-pregnancy weight/height squared.

Outcomes

Birth weight, outcome of pregnancy (live or stillbirth), gender of the baby and gestational age at birth were obtained from obstetric records. Birth weights were corrected for gestational age and gender. Preterm delivery was defined as birth before 37 weeks of gestation. Only pregnancies where clinical estimates of length of gestation based on ultrasonography agreed with mothers’ dates (plus or minus 2 weeks) were included. Women were asked at 18 weeks about any previous miscarriages. The data were then categorised as none, one and two or more.

Data analysis

Parametric (one-way analysis of variance) and non-parametric tests were used as appropriate for group comparisons, after testing for normality. Bivariate linear regression models were used to test for predictors of continuous outcomes. Multinomial and binary logistic regression models examined predictors of categorical and binary outcomes respectively.

Potential covariates likely to influence outcomes were first tested in bivariate models and included in multivariate models when significant. The final model accounted for the main effects of each covariate. Factors considered to be possible mediators (Kraemer et al, 2001) of main effects were included in the multivariate
model at a second stage. All analyses were performed using Stata version 8 for Windows. All statistical tests presented are two-tailed. Statistical significance was defined as \( P < 0.05 \).

Although our sample was relatively big, the sizes of groups with eating disorders were variable and some groups were small (anorexia nervosa plus bulimia nervosa in particular) in relation to the ‘general population’ control sample. We were therefore concerned that differences in rarer outcomes might not be detectable when comparing groups with eating disorders and the reference group. Hence we carried out a power calculation and found that effect sizes of 0.3 in continuous outcomes could be detected with a power of 75–93% at the 5% significance level. Group differences in proportions for common outcomes could be detected with 92–99% power and group differences in proportions for uncommon outcomes (such as prematurity) could be detected with 63–99% power at the 5% significance level.

### Ethical approval

The study was approved by the ethics committees of the Institute of Psychiatry and ALSPAC.

## RESULTS

Women who were included in the current study (n=12,254) were divided into five groups: (a) 171 (1.4%) who only endorsed the question ‘Have you ever had anorexia nervosa?’ (7 of these reported a recent episode); (b) 199 (1.6%) who only endorsed the question ‘Have you ever had bulimia nervosa?’ (51 of these reported a recent episode); (c) 82 (0.7%) who endorsed both questions; (d) 1166 (9.5%) who reported having had schizophrenia, severe depression or other psychiatric disorders (including drug addiction and alcoholism) and formed the ‘other psychiatric disorders’ group; (e) 10,636 who formed the ‘general population’ comparison group.

### Socio-demographic data

Maternal age at delivery and ethnicity did not differ across the five groups (see Table 2). Women with other psychiatric disorders were less likely to be in full-time or part-time employment, or full-time education or training and were more likely to be multiparous than the general population sample. Women in the three eating disorder groups did not differ from the general population sample on parity or employment status. Women with a history of anorexia nervosa, anorexia nervosa plus bulimia nervosa and other psychiatric disorders were significantly more likely to have smoked during the first trimester of pregnancy. Women with other psychiatric disorders were significantly more likely to have smoked during the second trimester of pregnancy and drunk alcohol during the first trimester. All four clinical groups were less likely to be living with a partner than the ‘general population’ group.

### Eating disorders and related symptoms

We compared BMI across the five groups and the proportions of women reporting past vomiting and laxative use for weight loss (Table 2). Women in the three eating disorder groups were significantly more likely to have used laxatives and self-induced vomiting. Women with a history of anorexia nervosa and anorexia nervosa plus bulimia nervosa had a significantly lower mean BMI than the other groups (Table 2).

### Pregnancy outcomes

Foetal deaths (n=66) were excluded from these analyses. Women with a history of anorexia nervosa had 2 foetal deaths (1.2%), those with bulimia nervosa and those with anorexia nervosa plus bulimia nervosa had none, those with other psychiatric disorders had 7 (0.6%) and general population controls had 57 (0.7%). Differences were not statistically significant.

### Birth weight

We excluded 67 women who developed gestational diabetes because of high rates of macrosomia in this group. Rates of gestational diabetes were significantly higher in the group with anorexia nervosa plus bulimia nervosa (2 positive, 2.4%, Fisher’s exact=17.9, \( P=0.01 \)) and that with other psychiatric disorders (16 positive, 1.4%) compared with the general population (48 positive, 0.5%). Data were missing on birth weight for 148 babies.

Mean birth weights corrected for gender and gestational age were calculated for 11,973 babies. The mean birth weight for babies born to women with a history of anorexia nervosa was 3340 g (95% CI 3220–3460 g).

### Table 1 Socio-demographic data

<table>
<thead>
<tr>
<th></th>
<th>Anorexia nervosa</th>
<th>Bulimia nervosa</th>
<th>Anorexia nervosa plus bulimia nervosa</th>
<th>Other psychiatric disorders</th>
<th>General population</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n=171</td>
<td>n=199</td>
<td>n=82</td>
<td>n=1166</td>
<td>n=10,636</td>
</tr>
<tr>
<td><strong>Age at delivery, years (s.d.)</strong></td>
<td>28.9 (5.2)</td>
<td>28.2 (4.6)</td>
<td>29.2 (4.6)</td>
<td>28.5 (5.5)</td>
<td>28.2 (4.8)</td>
</tr>
<tr>
<td><strong>Multiparity, % (OR, 95% CI)</strong></td>
<td>52.5 (0.9, 0.6–1.2)</td>
<td>51.6 (0.9, 0.7–1.2)</td>
<td>53.3 (0.9, 0.6–1.5)</td>
<td>59.1 (1.2, 1.0–1.3)**</td>
<td>54.9</td>
</tr>
<tr>
<td><strong>White ethnicity, % (OR, 95% CI)</strong></td>
<td>96.2 (0.6, 0.3–1.4)</td>
<td>97.4 (0.9, 0.4–2.2)</td>
<td>98.8 (1.0, 0.3–13.8)</td>
<td>98 (1.2, 0.8–1.9)***</td>
<td>97.6</td>
</tr>
<tr>
<td><strong>Employment, % (OR, 95% CI)</strong></td>
<td>490 (1.0, 0.7–1.4)</td>
<td>48.9 (1.0, 0.7–1.3)</td>
<td>42.7 (0.8, 0.5–1.2)</td>
<td>32.8 (0.6, 0.6–0.7)**</td>
<td>49.1</td>
</tr>
<tr>
<td><strong>Any smoking in first trimester, % (OR, 95% CI)</strong></td>
<td>27.8 (1.4, 1.0–2.0)*</td>
<td>26.2 (1.3, 0.9–1.8)</td>
<td>39.5 (2.4, 1.5–3.8)**</td>
<td>40.2 (2.5, 2.2–2.8)*****</td>
<td>21.4</td>
</tr>
<tr>
<td><strong>Any smoking in second trimester, % (OR, 95% CI)</strong></td>
<td>20.2 (1.3, 0.9–1.9)</td>
<td>20.6 (1.3, 0.9–1.8)</td>
<td>23.5 (1.6, 0.9–2.6)</td>
<td>32.6 (2.5, 2.2–2.8)*****</td>
<td>16.3</td>
</tr>
<tr>
<td><strong>Has a partner, % (OR, 95% CI)</strong></td>
<td>95.2 (0.4, 0.2–0.8)*</td>
<td>95.6 (0.4, 0.2–0.9)*</td>
<td>90.5 (0.2, 0.1–0.3)**</td>
<td>94.6 (0.4, 0.3–0.5)***</td>
<td>98.1</td>
</tr>
<tr>
<td><strong>Any alcohol use in first trimester, % (OR, 95% CI)</strong></td>
<td>11.6 (0.7, 0.4–1.2)</td>
<td>18.8 (1.3, 0.9–1.9)</td>
<td>24.7 (1.8, 1.1–3.0)*</td>
<td>19.3 (1.3, 1.1–1.6)*</td>
<td>15.2</td>
</tr>
</tbody>
</table>

1. Percentage in full-time/part-time employment or full-time education/training vs unemployed, housewives or retired.

\*P < 0.05, **P < 0.001 vs. general population.
Table 2  Lifetime weight control behaviours and pre-pregnancy body mass index

<table>
<thead>
<tr>
<th></th>
<th>Anorexia nervosa</th>
<th>Bulimia nervosa</th>
<th>Anorexia nervosa plus bulimia nervosa</th>
<th>Other psychiatric disorders</th>
<th>General population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-induced vomiting for weight control, %</td>
<td>23.4***</td>
<td>56.3***</td>
<td>62.2***</td>
<td>4.0</td>
<td>3.6</td>
</tr>
<tr>
<td>Laxative use for weight control, %</td>
<td>25***</td>
<td>29.1***</td>
<td>55***</td>
<td>4.5</td>
<td>3.2</td>
</tr>
<tr>
<td>Body mass index pre-pregnancy: mean (s.d.)</td>
<td>21.5 (3.2)**</td>
<td>23.1 (4.3)</td>
<td>21.5 (3.0)*</td>
<td>23.1 (4.2)</td>
<td>22.9 (3.8)</td>
</tr>
</tbody>
</table>

*P < 0.05, **P < 0.01, ***P < 0.001, v general population.

Table 3  Linear regression analysis of birth weight after stepwise adjustment for relevant covariates

<table>
<thead>
<tr>
<th></th>
<th>Anorexia nervosa</th>
<th>Bulimia nervosa</th>
<th>Anorexia nervosa plus bulimia nervosa</th>
<th>Other psychiatric disorders</th>
<th>General population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Birth-weight'</td>
<td>−83.9 (−151.9 to −15.9)</td>
<td>14.3 (−48.7 to 77.4)</td>
<td>−2.6 (−101.7 to 96.5)</td>
<td>−33.5 (−60.9/−6.1)</td>
<td>−0.018*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight' adjusted for maternal factors</td>
<td>−75.1 (−143.6 to −6.5)</td>
<td>20.5 (−42.4/83.5)</td>
<td>−2.6 (−101.7 to 96.5)</td>
<td>−36.4 (−64.0 to −8.8)</td>
<td>−0.018*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birth weight' adjusted for maternal factors and smoking in the second trimester</td>
<td>−64.5 (−132.3/3.3)</td>
<td>27.6 (−34.5 to 89.7)</td>
<td>8.8 (−90.2 to 107.9)</td>
<td>−4.4 (−31.9 to 23.2)</td>
<td>−0.016*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Birthweight' adjusted for maternal factors, smoking in the second trimester and BMI</td>
<td>−40.7 (−109.1 to 27.8)</td>
<td>11.5 (−52.6 to 75.5)</td>
<td>23.9 (−78.8 to 126.6)</td>
<td>−2.7 (−31.0 to 25.5)</td>
<td>−0.009</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Anorexia nervosa</th>
<th>Bulimia nervosa</th>
<th>Anorexia nervosa plus bulimia nervosa</th>
<th>Other psychiatric disorders</th>
<th>General population</th>
</tr>
</thead>
<tbody>
<tr>
<td>B (95% CI) /</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**P < 0.05, ***P < 0.01.
1. Adjusted for gestational age and gender of the baby.
2. P = 0.06 v general population.

3272–3407); to women with bulimia nervosa 3439 g (3377–3502); to women with anorexia nervosa plus bulimia nervosa 3422 g (3323–3521); to women with other psychiatric disorders 3392 g (3366–3413); and to the general population sample 3425 g (3416–3433). Babies of women with anorexia nervosa were significantly lighter than babies of control women, as were babies of women with other psychiatric disorders (overall F (6, 11966)=918.8, P < 0.05) (Table 3).

We studied the role of covariates known to influence birth weight, including maternal factors such as parity, maternal age, employment status, whether women had a partner and alcohol intake (a factor relating to the studied pregnancy). Alcohol intake, relationship status and employment status were not significantly related to the outcome and were not included in the final model.

Smoking in the first and second trimester, pre-pregnancy BMI, laxative use and self-induced vomiting in pregnancy were investigated as possible mediators of effect. Laxative use and self-induced vomiting in pregnancy were not significantly related to birth weight in bivariate analyses. When maternal covariates (parity, maternal age) were included in the model, babies born to women with a lifetime history of anorexia nervosa were still significantly lighter than babies of control women (B = −75.1, β = −0.016, P = 0.03) (Table 3). When smoking in the second trimester was included in the model, a marginal difference remained for babies of women with anorexia nervosa compared with general population controls (B = −63.5, β = −0.013, P = 0.06). When BMI pre-pregnancy was included in the model, the effect of maternal history of anorexia nervosa on birth weight disappeared.

Preterm delivery

Data for evaluation of preterm delivery were available in 12 188 births. The rates of preterm delivery were: anorexia nervosa 6.5%; bulimia nervosa 5.0%; anorexia nervosa plus bulimia nervosa 4.9%; other psychiatric disorders 5.8%; general population 4.8%; with no group differences on logistic regression analysis. After controlling for ethnicity, maternal age, and parity, the group with other psychiatric disorders had significantly higher rates of preterm delivery compared with the general population (odds ratio 1.3, 95% CI 1.0–1.8, P = 0.03).

Previous miscarriages

Data on previous miscarriages were analysed in 11 700 women. An initial multinomial logistic regression showed that women with bulimia nervosa, those with anorexia nervosa plus bulimia nervosa and those with other psychiatric disorders were significantly more likely to report previous miscarriages (Table 4). When adjusted for relevant covariates (lifetime smoking and alcohol use, age, parity), only women with a history of bulimia nervosa and of other psychiatric disorders remained significantly more likely to have a history of previous miscarriages than the general population. A trend remained for women with anorexia nervosa plus bulimia nervosa.
The same three groups of women were significantly more likely to have had two or more miscarriages compared with the general population. The difference remained after controlling for relevant covariates (Table 4).

**DISCUSSION**

**Main findings**

Women with a history of bulimia nervosa (with or without a history of anorexia nervosa) had an increased rate of lifetime miscarriages, as did women with a history of other psychiatric disorders. This persisted after controlling for potential covariates. Women with a history of anorexia nervosa were more likely to deliver babies of lower birth weight than control women, although weights were comparable to babies of women with other psychiatric disorders.

**Miscarriages**

Higher rates of miscarriage in women with bulimia nervosa have been reported previously (Mitchell et al., 1991; Morgan et al., 2006). A higher risk of miscarriage for women with current and past bulimia nervosa was reported in two studies (Abraham, 1998; Blais et al., 2000). Our results confirm these findings. Possible hypotheses include polycystic ovary syndrome and leptin abnormalities (Morgan et al., 2006). Future research will need to address the issue of direct cause of miscarriages in women with bulimia nervosa and the exact physiology.

**Birth weight**

Previous studies have shown that women with current or past eating disorders have a higher risk of delivering lower birth weight babies (Stewart et al., 1987; Bulik et al., 1999; Sollid et al., 2004) and our study confirms this finding. However, we found that the lower birth weight of babies born to women with anorexia nervosa may be mediated by lower pre-pregnancy BMI and to a lesser extent by smoking in the second trimester of pregnancy. None of the previous studies has investigated the effect of either variable in a population with eating disorders. However, the effect of maternal weight pre-pregnancy on birth weight of offspring has been documented in population studies; low maternal weight at conception or delivery has been found to have a significant impact on perinatal outcomes, mainly birth weight and preterm delivery (Kaminsky et al., 1973; Wolfe et al., 1991; Cnattingius et al., 1998; Ehrenberg et al., 2003). It is likely that a low pre-pregnancy BMI is an indicator of poor maternal nutritional status during pregnancy, but we were not able to evaluate this in this study.

Previous studies have highlighted an increased risk for adverse perinatal outcomes in women with severe mental illness (Jablensky et al., 2005), but no previous study has compared women with eating disorders with women with other severe psychiatric disorders. In our study, smoking during the second trimester seemed to be mainly responsible for the low birth weight in women with other psychiatric disorders. This suggests that the mechanism for low birth weight might be different in women with other severe psychiatric disorders compared with women with anorexia nervosa.

**Preterm delivery**

Two previous studies of clinical samples have shown higher rates of prematurity in babies of women with eating disorders (Bulik et al., 1999; Sollid et al., 2004). Bulik et al. (1999) relied on a small sample and self-report of premature birth. The study of Sollid et al. (2004), although larger, was register-based and included only women who had been hospitalised for an eating disorder, which was likely to be severe. Recall and sampling differences might therefore partly explain the disparity of these findings with those of our study. Our study is in line with that of Franko et al. (2001) who found no difference in rates of prematurity when comparing women with anorexia and bulimia nervosa. There is the possibility that this finding might be a result of a low power to detect differences in our sample. This finding needs replication.

**Strengths and limitations**

The strengths of the study include the use of data from a large longitudinal prospective community cohort. We were able to include a comparison group of women with psychiatric disorders other than eating disorders in addition to a general population control group. We were also able to take into account the role of several covariates relevant to the outcomes.

The main weakness of this study is that women were classified according to self-report of lifetime anorexia nervosa or bulimia nervosa or both. It is uncertain how accurate this classification is in terms of psychiatric classificatory systems. However, the availability of rates of lifetime eating disorder behaviours and BMI pre-pregnancy lends weight to self-reported diagnoses. The prevalence of eating disorders in this sample was 3.7%. According to estimates of the prevalence of eating disorders in women of child-bearing age...
et al (2006), the prevalence of anorexia nervosa is between 0 and 1.5% and that of full-syndrome bulimia nervosa between is 0.4 and 0.8%. When partial syndromes are included the prevalence rate of eating disorders reaches about 5%. The prevalence of anorexia nervosa in our sample is 1.4% and that of bulimia nervosa 1.6%. It is therefore likely that a proportion in these two groups might have had an eating disorder not otherwise specified or a milder eating disorder compared with clinical samples. The current study is therefore likely to have underestimated rather than overestimated the rates of adverse perinatal outcomes in women with eating disorders.

Another limitation of the study is that weights and heights pre-pregnancy were also obtained by self-report. Moreover, we were not able to determine the temporal relationship between previous miscarriages and the course of bulimia nervosa. The sample did not have sufficient power to determine whether rare complications such as foetal deaths were more common in women with anorexia nervosa, although there was a trend in this direction.

Implications

Our results, together with previous reports in the literature, suggest that maternal eating disorders are associated with higher risk of some obstetric complications. This is extremely relevant to the prevention of adverse foetal outcomes. Moreover, the extent to which perinatal complications are predictors of later psychiatric disorders is still unclear. We found that women with eating disorders have similar rates of major adverse perinatal outcomes to women with other psychiatric disorders, although some of the causal factors implicated might differ. Women with bulimia nervosa are at higher risk of miscarriage. Future research will need to clarify the exact mechanism.

Women with a history of anorexia nervosa should be informed when planning a pregnancy that good general health includes having a healthy BMI as well as smoking cessation. Previous studies suggest that the association of smoking with high levels of body image distortion, and the role of smoking in weight control are relevant to women with and without eating disorders (George & Waller, 2005; John et al., 2006). If this is so, the link with body image and weight control may need to be considered when counselling women about smoking cessation in pregnancy.

NADIA MICALI, MD, MRCPsych, Department of Child and Adolescent Psychiatry and Eating Disorders Research Unit, Institute of Psychiatry, King’s College London; EMILY SIMONOFF, MD, FRCPsych, Department of Child and Adolescent Psychiatry, Institute of Psychiatry, King’s College London; JANET TREASURE, PhD, FRCPsych, Eating Disorders Research Unit, Department of Academic Psychiatry (Guy’s Hospital), King’s College London and ALSPEC, Department of Pediatric and Perinatal Epidemiology, University of Bristol, Bristol, UK

Correspondence: Dr Nadia Micali, Department of Child and Adolescent Psychiatry, Institute of Psychiatry, King’s College London, Box 085, De Crespigny Park, London SE5 8AF, UK. Email: N.Micali@op.kcl.ac.uk

(First received 14 December 2005, final revision 10 October 2006, accepted 8 November 2006)

Experts agree that women should be counselled to delay pregnancy until the eating disorder is in complete remission (Sollid et al., 2004). Advising women with eating disorders on possible effects of the disorder on fertility and the possibility of adverse outcomes in their offspring could be important for motivating women to implement changes in their behaviour.

ACKNOWLEDGEMENTS

We thank Dr Sam Leary, ALSPEC, for her help with the data, the women and children who took part in this study, the midwives and the whole ALSPEC team. The study was funded by the National Alliance for Schizophrenia and Depression and the Psychiatry Research Trust. The UK Medical Research Council, the Welcome Trust and the University of Bristol provided core support for ALSPEC.

REFERENCES


Effect of antidepressant therapy on executive function after stroke*

KENJI NARUSHIMA, SERGIO PARADISO, DAVID J. MOSER, RICARDO JORGE and ROBERT G. ROBINSON

**Background** Executive dysfunction is common after stroke and may impair long-term outcome. Remedies for this condition are limited.

**Aims** To examine the effect of antidepressants on executive function after stroke.

**Method** Forty-seven patients who had had a stroke during the prior 6 months received 12 weeks of antidepressant treatment in double-blind placebo-controlled fashion, followed by assessment of executive function at the end of treatment and after 2 years.

**Results** No significant group effect was found at the end of treatment. However, 21 months after the end of treatment the placebo group showed deterioration of executive function, whereas the active treatment group showed clear and significant improvement independent of depressive symptoms (F = 2.1, d.f. = 1.45, P = 0.001).

**Conclusions** Antidepressant treatment fosters long-term improvement of executive function following stroke. This phenomenon is consistent with a reorganisation of neuronal networks associated with prefrontal functions based on modulation of monoaminergic neurotransmission and the activity of neurotrophins.

**Declaration of interest** None.

Funding detailed in Acknowledgements.

*Presented in part at the annual meeting of the American Neuropsychiatry Association, La Jolla, California, 18–21 February 2006.

Executive functions are mental abilities that allow the individual to respond appropriately to unfamiliar and complex situations and support several cognitive, emotional and social capacities. Impairment of executive functions occurs in the majority of stroke patients (Lesser et al, 1996), and may slow rehabilitation and return to premorbid levels of psychosocial adjustment.

Antidepressant treatment improves outcome following stroke, independently of depression. Nortriptyline and fluoxetine improve activities of daily living (Narushima & Robinson, 2003) and mortality (Jorge et al, 2003), and sertraline improves morbidity (Rasmussen et al, 2003). Antidepressants may exert outcome-improving effects on the brain in at least two ways: one is through modulation of cortical striato-pallido-thalamo-cortical pathways subsequent to action on the raphe nuclei, locus ceruleus and ventral tegmental area (Alexopoulos et al, 2000); another is through reorganisation of neural circuitry favoured by their activity on brain-derived neurotrophic factor (BDNF) (Saarelaïnen et al, 2003).

This evidence prompted our study hypothesis that antidepressants improve frontal executive function following stroke. This hypothesis has not been tested before. The positive effect of treatment with antidepressants on the cognitive abilities of older adults (Allard et al, 2003) and the identification of similar mechanisms of executive deficit in late life and stroke (Coffey et al, 1988a,b) further supported the rationale behind this study. We predicted that antidepressant treatment would improve executive dysfunction independently of depression.

**METHOD**

**Patient selection**

Data allowing examination of the study hypothesis were collected during a double-blind placebo-controlled treatment study of post-stroke depression (Robinson et al, 2000). Participants were patients consecutively admitted to the Younkers Rehabilitation Center at the Iowa Methodist Medical Center in Des Moines, Iowa. Patients were included if they had experienced a stroke within the preceding 6 months. Exclusion criteria were:

(a) medical illness threatening the patient’s life or recovery;

(b) severe comprehension deficit, i.e. failing part 1 of the Token Test (De Renzi et al, 1978);

(c) a history of head injury or other brain disease.

Patients taking antidepressants discontinued them for a 2-week wash-out period before the study. Written consent was obtained in accordance with institutional review board requirements. Ninety-two patients entered the double-blind placebo-controlled phase and 69 completed it (Fig. 1).

After completing the treatment phase, seven patients refused to repeat the assessment. Fifteen patients developed complications and were excluded. In the active treatment group, one patient developed Parkinson’s disease, one had a transient cerebrovascular ischaemic event just before the evaluation and five experienced seizures. In the placebo group, one patient developed psychosis, two had a stroke, and five experienced seizures. All patients who developed seizures took anticonvulsants.

There was no significant group difference in the frequency of complications (x^2 = 3.04, P = 0.08). Clinical and background variables of patients not included in the analysis were not significantly different from those of participants included in the analyses. Data from 47 patients were analysed, 30 of these patients were treated with either nortriptyline (n = 11) or fluoxetine (n = 19), and 17 received placebo.

**Treatment protocol**

Patients were randomly assigned to 12 weeks of either fluoxetine, nortriptyline or placebo, unless there was a definite contraindication. Nortriptyline was chosen because it is reported to be effective in post-stroke depression (Lipsey et al, 1984); fluoxetine was chosen because it is a commonly used selective serotonin reuptake inhibitor with the advantage of a long half-life. Nortriptyline was not given to patients with cardiac abnormalities, and fluoxetine

---

*BRITISH JOURNAL OF PSYCHIATRY (2007), 190, 260–265. doi: 10.1192/bjp.bp.106.025064*
was contraindicated in patients with intracerebral haemorrhage. Eight patients had a contraindication to nortriptyline and nine to fluoxetine: these patients were randomly reassigned to the alternative active medication or to placebo. Thus, 85% of the patients were randomly assigned to nortriptyline, fluoxetine or placebo. All patients were randomly assigned to either active or placebo medication.

The dosages of nortriptyline were 25 mg per day for the first week, 50 mg per day for weeks 2 and 3, then 75 mg per day for weeks 4–6 and 100 mg per day for the final 6 weeks. Dosages of fluoxetine were 10 mg per day for the first 3 weeks, 20 mg per day for weeks 4–6, then 30 mg per day for weeks 7–9 and 40 mg per day for the final 3 weeks. In nine patients the dosages had to be decreased owing to severe side-effects; five were being treated with nortriptyline and four with fluoxetine. Reduction of dosage was achieved in double-blind fashion. All the patients were fully compliant with their medication regimen throughout the treatment period. After completion of the treatment phase of the study, all the medications were discontinued.

Assessment

The initial neuropsychological examination took place after the completion of the 12-week double-blind phase. This was followed by a second evaluation 21 months after the end of active treatment.

The neurological examination was conducted by a neurologist using the National Institutes of Health (NIH) Stroke Scale (Brott et al., 1989). Severity of depression was assessed using the 17-item version of the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960). Overall cognitive functioning was assessed with the Mini-Mental State Examination (MMSE; Folstein et al., 1975). All assessments were conducted by examiners masked to the participants’ background conditions.

The following tests of executive function constituted the primary outcome variables: the Controlled Oral Word Association (COWA; Benton, 1967) test measured initiation and psychomotor speed; the Wisconsin Card Sorting Test Perseverative Errors (WCST–PE; Grant & Berg, 1948) and the Wechsler Adult Intelligence Scale–Revised (WAIS–R; Wechsler, 1981) Similarities sub-tests measured conceptualisation and problem-solving; and the WAIS–R Arithmetic and Digit Span sub-tests measured attention and working memory.

Neuroimaging

Computerised tomography or magnetic resonance scans were obtained and evaluated for anatomical location and lesion volume by a neuroradiologist or a neurologist unaware of the psychiatric findings. Lesion volume was estimated using the ratio of the largest cross-sectional area of the lesion to the area of the brain slice that included the body of the lateral ventricles (Robinson et al., 1986).

Statistical analysis

Between-group comparisons were made using means, standard deviations and repeated-measures analysis of variance (ANOVA). Frequency distributions were compared using the chi-squared test. Individual executive function scores were transformed into standardised z scores which were combined into an ‘executive index’. The follow-up executive index scores were based on z scores calculated using the baseline means and standard deviations. Changes between the initial and follow-up measures were evaluated by change scores, where positive scores showed improvement and negative scores showed deterioration. Multiple linear regression was used to examine the potential contribution of clinical factors to executive improvement. All tests were two-tailed and significance was set at $P < 0.05$.

RESULTS

Patient characteristics, neurological and radiological findings

There was no statistically significant group difference in background characteristics and demographic data (Table 1). During the 21-month post-treatment period, one patient in the active treatment group refused to participate, two complained of side-effects, and four died; in the placebo
group three patients refused follow-up evaluations and one died. Thus, a total of 36 patients completed all evaluations. There was no significant group difference in rates of withdrawal. The background variables of patients who completed the follow-up neuropsychological evaluations were not significantly different from those of patients who did not complete them.

In the period between the initial and follow-up evaluations, some patients were given antidepressants by their treating physicians. Type, dosage and duration of all prescribed medications were recorded. The frequency and the duration of antidepressant treatment in the two groups were not significantly different (Table 1). Additionally, there was no significant group difference in MMSE scores at the initial evaluation (active treatment, mean=27.5, s.d.=0.71; placebo group, mean=27.0, s.d.=0.95; P=0.66) and at follow-up (active treatment, mean=27.8, s.d.=0.85; placebo group, mean=25.6, s.d.=0.95; P=0.12). There was no significant group difference in neurological and radiological variables (Table 2).

Executive index
At the completion of the treatment phase, executive performance showed no statistically significant between-group effect (F=0.01, d.f.=1,45, P=0.91).

A significant time-by-group interaction was found between the initial and follow-up evaluation. This was true for both efficacy analysis (i.e. analysing only patients who completed the 24-month follow-up: F=13.1, d.f.=1,34, P=0.001) and intention-to-treat (ITT) analysis (i.e. analysing all patients who enrolled in the treatment study, with missing data interpolated based on actual observations using the last observation carried forward method: F=12.1, d.f.=1,45, P=0.001). Patients given active treatment showed an improvement in executive performance after 21 months, whereas patients given placebo showed a decline (Fig. 2).

There was no significant difference between patients treated with nortriptyline and fluoxetine. Paired t-test analysis among actively treated patients showed significant improvement (efficacy analysis: t=−2.39, d.f.=22, P=0.03; ITT analysis: t=−2.33, d.f.=29, P=0.03), whereas placebo-treated patients showed significant deterioration (efficacy analysis: t=−2.75, d.f.=12, P=0.02, ITT analysis: t=−2.59, d.f.=16, P=0.02). All but one patient who received placebo showed deterioration of executive function 2 years after the initiation of the study.

Individual tests of executive function
Participants’ performance on the individual tests composing the executive index was also examined (Fig. 3). The COWA and WCST–PE tests showed significant treatment effects compared with placebo in change scores (COWA test, efficacy analysis: F=6.46, d.f.=1,34, P=0.02; COWA test, ITT analysis: F=6.22, d.f.=1,45, P=0.02; WCST–PE, efficacy analysis: F=7.19, d.f.=1,34, P=0.01, WCST–PE ITT analysis: F=6.87, d.f.=1,45, P=0.01). The WAIS–R Similarities, Digit Span and Arithmetic sub-tests showed the same direction of change, but failed to reach statistical significance (Fig. 3).

Multiple linear regression analysis
Many factors may influence executive function recovery in addition to antidepressant treatment. In order to examine the contribution of these potentially confounding factors, an exploratory stepwise regression analysis was conducted including the
Following variables: active treatment or placebo; severity of depressive symptoms at baseline, on completion of the 3-month placebo-controlled treatment phase and after 2 years; age; gender; past and family psychiatric history; antidepressant treatment during the 21-month follow-up period; presence of neurological impairment (motor, sensory or visual impairments, or aphasia); lesion location (left hemisphere, right hemisphere or brain-stem); lesion’s proximity to the frontal pole (i.e. relative distance between the frontal pole and the most rostral margin of the stroke lesion compared with the distance between the frontal and occipital poles); type (infarction or haemorrhage) and volume of lesion.

The final model predicting executive function consisted of active treatment or placebo, age, past psychiatric history, neurological impairment and total lesion volume ($F(5,25)=3.91, P<0.05$). The only factor that showed a significant independent effect was active treatment or placebo ($F(1,25)=14.9, P<0.01$) (Table 3).

**Table 2** Stroke characteristics and neurological findings

<table>
<thead>
<tr>
<th>Location of lesion</th>
<th>Nortriptyline or fluoxetine group</th>
<th>Placebo (n=17)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=30) n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroke type</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infarction</td>
<td>27 (90)</td>
<td>16 (94)</td>
<td>0.62</td>
</tr>
<tr>
<td>Haemorrhage</td>
<td>3 (10)</td>
<td>1 (6)</td>
<td></td>
</tr>
<tr>
<td>Location of lesion</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left hemisphere</td>
<td>11 (37)</td>
<td>6 (35.3)</td>
<td></td>
</tr>
<tr>
<td>Right hemisphere</td>
<td>17 (57)</td>
<td>10 (59)</td>
<td>0.99</td>
</tr>
<tr>
<td>Brain-stem/other</td>
<td>2 (7)</td>
<td>1 (6)</td>
<td></td>
</tr>
<tr>
<td>Lesion volume (% of brain volume)</td>
<td>5</td>
<td>4</td>
<td>0.64</td>
</tr>
</tbody>
</table>

**Table 3** Multiple linear regression analysis of change in executive function during 21 months of follow-up

<table>
<thead>
<tr>
<th>Factor</th>
<th>d.f.</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment¹</td>
<td>1</td>
<td>14.9</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Age</td>
<td>1</td>
<td>0.73</td>
<td>0.40</td>
</tr>
<tr>
<td>Past psychiatric history</td>
<td>1</td>
<td>0.97</td>
<td>0.33</td>
</tr>
<tr>
<td>Lesion volume</td>
<td>1</td>
<td>3.02</td>
<td>0.09</td>
</tr>
<tr>
<td>Impairments²</td>
<td>1</td>
<td>1.37</td>
<td>0.26</td>
</tr>
</tbody>
</table>

1. Active vs. placebo treatment.  
2. Motor/sensory/visual deficits or aphasia.

**DISCUSSION**

To our knowledge this is the first study to examine the effect of antidepressant treatment on executive function following stroke. We found that 12 weeks of treatment, started within 6 months of the stroke, exerted a positive effect on executive function 21 months after treatment. The effect on cognitive function was independent of improvement in depression, was delayed and was independent of intervening antidepressant treatment. The effect of antidepressants on cognition did not extend to overall cognitive status as measured by the MMSE.

**Caveats**

Before discussing these findings, some factors that might have influenced our results should be acknowledged. The majority of patients were high-school or college-educated, White, married and belonged to Hollingshead social class I to III (I is the highest, V is the lowest). Therefore, the results may not be applicable to all stroke patients. As expected in studies examining elderly people, follow-up evaluations could not be obtained for all patients. This may limit generalisation of our findings to people willing or able to receive follow-up evaluations. Some patients were prescribed antidepressants during the naturalistic follow-up, and a small number of patients were reassigned to the alternative active treatment or placebo for medical reasons; these factors might have introduced bias (although naturalistic antidepressant treatment was not an independent predictor of recovery). Finally, there was no neuropsychological assessment to ensure the two groups had equivalent levels of executive function prior to treatment; hence, despite
randomisation, this might have influenced the results.

Clinical implications
The findings in this study have important implications for the neuropsychiatry of stroke and rehabilitation medicine. Antidepressants administered within 6 months of stroke appear to improve long-term executive function outcome. This effect is twofold. As can be observed in Fig. 1, antidepressants both improve and prevent decline of executive function. How can this interesting and clinically important phenomenon be explained?

Mechanisms
Frontal cortical–subcortical circuits
Five frontal cortical–subcortical circuits—motor, oculomotor, dorsolateral prefrontal, lateral orbital frontal, and anterior cingulate—subserve distinct motor and cognitive abilities (Alexander et al., 1986). All circuits originate in the prefrontal cortex, project to the striatum, synapse at the level of the globus pallidus, substantia nigra and thalamus, and finally return to the prefrontal cortex, forming closed loops—the cortical–striato-pallido-thalamo-cortical (CSPTC) pathway. The dorsolateral prefrontal, orbitofrontal, and anterior cingulate cortical circuits subserve both executive (Drewe, 1974; Baker et al., 1996) and affective functions (Baxter et al., 1985, 1989; Drevets et al., 1992; Baker et al., 1996; Elliott et al., 1997). They can be modulated by the activity of monoaminergic nuclei, including the raphe nuclei, the locus ceruleus and the ventral tegmental area (Alexopoulos et al., 2000), which are sites of action of antidepressant medications. Thus, the mechanism of executive function recovery may be the modulation of monoaminergic nuclei exerting effects on CSPTC circuits.

Neurogenesis
Another possible mechanism stems from the association of chronic antidepressant administration and neurogenesis. Neurogenesis in the adult brain is generally thought to be restricted to germinonal centres in the subventricular zone and the hippocampal/dentate gyrus (Petersen, 2002). Chronic administration of antidepressants enhances the development of immature neurons and promotes the survival and function of adult neurons by enhancing BDNF and its receptor trkB, resulting in functional and anatomical changes. Activation of BDNF and trkB receptor has been shown to be required for antidepressants to induce behavioural effects (Saarelainen et al., 2003), and has been posited as an explanation for the delayed treatment effect of antidepressants (Nibuya et al., 1995). Although there has been no demonstration of neurogenesis in the prefrontal cortex, chronic antidepressant treatment induces activation of trkB receptor in the prefrontal cortex and is responsible for the sensitisation to the effects of BDNF (Saarelainen et al., 2003). Neurotrophins, particularly BDNF, have been shown to regulate neurite outgrowth (membrane-enclosed protrusions of neuronal cell cytoplasm), synaptic plasticity and the selection of functional connections in the central nervous system in general (Katz & Shatz, 1996; McAllister et al., 1999). Consistent with these findings is the recent notion that chronic administration of antidepressants prevents stress-induced reduction of BDNF (Manji & Duman, 2001; McEwen & Lasley, 2003; Brown et al., 2004). The finding in this study may make the study of neurotrophin-mediated mechanisms of improved executive function worth pursuing in greater detail.

Future studies
Our study has shown that early treatment with antidepressants following stroke has a remote positive effect on recovery and prevention of decline in executive function. Monoaminergic modulation of frontal executive functions and/or enhancement of neuronal plasticity and reorganisation of limbic and frontal structures may underlie this phenomenon. Our findings require confirmation in further studies, which might also explore whether any particular antidepressant is to be preferred, and the optimal time, duration and dosage of treatment.

ACKNOWLEDGMENTS
We thank Stephen Arndt, PhD, Teresa Kopel, MA, Stephanie Rosazza, BS, and J. Todd Kosier, BS, for their help in this research project. Thanks to Lauren Wagner, MD, for help in editing the manuscript. We are indebted to the patients of the Toukiers’ Rehabilitation Center of the Iowa Methodist Medical Center in Des Moines, Iowa. The study was supported in part by National Institute of Mental Health ROI grants MH-40355, MH-52879, MH-53592, MH63405 (R.G.R.), MH-52879 (S.P.) and Research Scientist Award MH-00163 (R.G.R.). S.P. is supported by the Edward J. Malinckrodt foundation and by a National Institutes of Health sponsored Institutional Career Development Award (K22).

REFERENCES


Effect of barriers on the Clifton suspension bridge, England, on local patterns of suicide: implications for prevention

OLIVE BENNEWITH, MIKE NOWERS and DAVID GUNNELL

Summary  We assessed the effect of the installation of barriers on the Clifton suspension bridge, Bristol, England, in 1998 on local suicides by jumping. Deaths from this bridge halved from 8.2 per year (1994–1998) to 4.0 per year (1999–2003; P=0.008). Although 90% of the suicides from the bridge were by males, there was no evidence of an increase in male suicide by jumping from other sites in the Bristol area after the erection of the barriers. This study provides evidence for the effectiveness of barriers on bridges in preventing site-specific suicides and suicides by jumping overall in the surrounding area.

Declaration of interest  None.

A number of sites around the world, particularly bridges, have gained notoriety as places from which suicide by jumping is popular (Gunnell & Nowers, 1997). As many acts of self-harm are impulsive in nature (Mann, 2003), restricting access to commonly used methods can result in reductions in both method-specific and overall suicide rates.

While two studies have found barriers to be effective in the prevention of suicide by jumping from particular bridges (O’Carroll et al, 1994; Beautrais, 2001) neither study investigated thoroughly the effects on suicide by jumping from other sites nearby and overall suicides. In December 1998, two metre-high wire barriers were installed on the main span of the Clifton suspension bridge in Bristol. For architectural reasons similar protective measures were not placed on the buttress walls at either end of the bridge (a photograph of the bridge is available as a data supplement to the online version of this paper). We used local and national suicide data to assess the effectiveness of these barriers in suicide prevention.

METHOD

The Clifton suspension bridge is located at the centre of the geographic area served by the Bristol coroner (Nowers & Gunnell, 1996). The bridge is over 6 km from the nearest psychiatric hospital; it is 75 m above the river and the case fatality of jumps from the bridge is over 95%.

Coroners’ inquest files were examined to obtain information on all suicides occurring in the Bristol area, 5 years before (1994–1998) and 5 years after (1999–2003) the installation of the barriers. All deaths with an inquest verdict of suicide were included in the study. Records of deaths given an open, accidental or misadventure verdict by the coroner were also examined, as previous research suggests that some deaths that are likely to be suicide are given such verdicts for legal reasons (O’Donnell & Farmer, 1995). For cases given these verdicts, vignettes describing the events leading up to the death were written (O.B.). The likelihood (high, medium, low or unclear) that these deaths were suicide was rated independently by D.G. and M.N., masked to the year of death. Only cases rated as medium or high likelihood were included in the study. Where the raters disagreed in their initial coding, consensus was reached through discussion. Of the 451 cases given a verdict other than suicide (open, n=189; accident or misadventure, n=260; no verdict, n=2), independent ratings by D.G. and M.N. resulted in agreement on inclusion or exclusion in 383 (84.9%) cases. After discussion a consensus on inclusion or exclusion was reached in the remaining 68 cases. We did not examine the coroner’s files for accidental acute alcohol poisonings or deaths from illegal drug use or methadone poisoning, as determining the possibility of suicide in such deaths is particularly problematic.

For all cases of suicide information was obtained on the person’s date of death, age and gender. To compute local and national rates of suicide, relevant population and mortality data were obtained from the Office for National Statistics on: (a) the number of suicides by jumping in England and Wales: ICD–10 codes X80 and Y30 (World Health Organization, 1992); (b) the overall number of suicides in England and Wales: ICD–10 codes X60–X84, Y10–Y34 excluding Y33.9 (where verdict pending); (c) population figures for the years 1994 to 2003.

Statistical analyses were carried out using Stata version 8.2 for Windows. Poisson regression was used to compare the number of deaths by jumping in the years before and after the construction of the barriers.

RESULTS

There were 987 suicides in the Bristol area over the 10-year study period. Of these deaths, 134 (13.6%) were suicides by jumping, 61 from the Clifton suspension bridge. There were a further 4 deaths where both the location of the body or skeletal remains and indications of trauma suggested that the person might have fallen from the bridge (n=3) or from nearby cliffs (n=1). All these deaths occurred before the barriers were erected, were given open verdicts and the remains were never identified; none of these deaths was included in subsequent analyses.

The number of deaths by jumping from the Clifton suspension bridge halfed (from 41 to 20; P=0.008) in the 5 years after the construction of the barriers compared with the previous 5 years (Table 1). Ninety per cent (55 of 61) of the people who died in this way were male, and the decline in deaths was seen in men only.

Before the barriers were erected (1994–1998) 30 of the 31 suicides (97%) for which the site of the jump was recorded were from the span of the bridge and only one (3%) from the buttresses. In the subsequent 5 years nearly half (8/17) of the jumps for which the site was recorded were from the buttresses where no fencing was in place. In the 5 years after the construction of the barriers there was a non-significant increase compared with the previous 5 years in the number of deaths by jumping from sites other than the suspension bridge: from 6.2 deaths per year to 8.4 deaths per year (P=0.2). This increase was entirely due to a rise in female deaths by jumping – in keeping with national trends in female suicide by jumping (see Table 1).
There was a non-significant fall in the mean number of deaths per year (14.4 to 12.4; \( P = 0.4 \)) by jumping from all sites in the area across the two study periods. This fall was due to a reduction in male (\( P = 0.017 \)) suicides by jumping. There was an increase in suicides by jumping among women (\( P = 0.001 \)). There was no change in the overall rate of suicide among those resident in the area during the periods before and after the placement of the barriers on the bridge: mean annual rate 11.2 per 100 000 vs. 10.5 per 100 000, difference \(-0.7\) (95% CI \(-1.9\) to \(-0.9\)), \( P = 0.39 \). This was the case for both men (difference \(-1.8\) per 100 000, 95% CI \(-1.7\) to \(-0.9\)) and women (difference 0.4 per 100 000, 95% CI \(-0.9\) to 2.1).

**DISCUSSION**

The number of deaths by jumping from the Clifton suspension bridge halved following the installation of the preventive barriers.

**Table I** Suicides by jumping before (1994–98) and after (1999–2003) the installation of preventive barriers on the Clifton suspension bridge

<table>
<thead>
<tr>
<th>Site of suicide by jumping</th>
<th>1994–1998</th>
<th>1999–2003</th>
<th>Difference in means (95% CI) ( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clifton suspension bridge</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All suicides</td>
<td>8.2</td>
<td>4.0</td>
<td>(-4.2) ((-5.9) to (-1.4)) 0.008</td>
</tr>
<tr>
<td>Total deaths</td>
<td>41</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8.0</td>
<td>3.0</td>
<td>(-5.0) ((-2.6) to (-6.3)) 0.001</td>
</tr>
<tr>
<td>Total deaths</td>
<td>40</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.2</td>
<td>1.0</td>
<td>(0.8) ((-0.08) to (8.4)) 0.1</td>
</tr>
<tr>
<td>Total deaths</td>
<td>1</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td><strong>Sites in Bristol other than the suspension bridge</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All suicides</td>
<td>6.2</td>
<td>8.4</td>
<td>(2.2) ((-0.9) to (7.2)) 0.2</td>
</tr>
<tr>
<td>Total deaths</td>
<td>31</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5.2</td>
<td>5.2</td>
<td>(0.2) ((-2.2) to (-3.8)) 1.0</td>
</tr>
<tr>
<td>Total deaths</td>
<td>26</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1.0</td>
<td>3.2</td>
<td>(2.2) ((0.2) to (7.7)) 0.023</td>
</tr>
<tr>
<td>Total deaths</td>
<td>5</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td><strong>All sites in England and Wales (rates per 100 000)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All suicides</td>
<td>0.34</td>
<td>0.36</td>
<td>(0.02) ((0.01) to (0.06)) 0.2</td>
</tr>
<tr>
<td>Male</td>
<td>0.54</td>
<td>0.53</td>
<td>(-0.01) ((-0.07) to (0.06)) 0.8</td>
</tr>
<tr>
<td>Female</td>
<td>0.15</td>
<td>0.20</td>
<td>(0.05) ((0.01) to (0.10)) 0.005</td>
</tr>
</tbody>
</table>

1. Poisson regression analyses.

Although there was a decrease overall in the number of deaths by jumping in the area among men, this was not the case for women. However, any impact on female suicide rates would be expected to be minimal, as only one woman jumped from the bridge in the 5 years prior to the installation of the barriers and national data suggest that suicide by jumping among females is increasing, although the proportional increase across the two study periods was higher in the Bristol area.

This study provides evidence for the preventive role of barriers on bridges. There was some evidence that the presence of the barriers did not lead to an increase in deaths by jumping from other sites. The case-fatality rate among those jumping from the Clifton bridge is greater than 95%. Therefore, any displacement of people deterred from jumping to other methods of suicidal behaviour is likely to have a beneficial effect on levels of suicide, because no other method is associated with a high case fatality. In view of continued suicides from some parts of the Clifton suspension bridge structure, further work to improve the safety of the site is warranted.

**ACKNOWLEDGEMENTS**

We thank Mr Paul Forrest, HM Coroner for Avon, staff employed at the coroner’s office, and Ms Alison Brown and search room staff at the Bristol Records Office, for their help in accessing suicide data. Local and national population data and national suicide data were provided by the Office for National Statistics for England and Wales. The study was funded by the American Foundation for Suicide Prevention.

**REFERENCES**


Impact of hospital consultants’ poor mental health on patient care

CATH TAYLOR, JILL GRAHAM, HENRY POTTS, JENNY CANDY, MICHAEL RICHARDS and AMANDA RAMIREZ

Summary In a survey of 1794 UK NHS hospital consultants 1308 (73%) responded. Psychiatric morbidity (General Health Questionnaire–12 score \( \geq 4 \)) was present in 32% of responders, who were twice as likely to report drinking hazardous levels of alcohol, being irritable with patients and colleagues, reducing their standards of care and intending to retire early (all \( \pi < 0.001 \)). Male and mid-aged consultants were also particularly at risk. Approaches that support consultants to practice medicine safely throughout their careers are required.

Declaration of interest None.

Funding detailed in Acknowledgements.

METHOD

A confidential postal survey was sent to 1794 UK NHS hospital consultants in late 2002. The sample included all surgical oncologists, medical oncologists and clinical oncologists, a random sample of gastroenterologists (two in three) and a random sample of radiologists (one in five). Consultants were ascertained through the UK medical Royal Colleges and professional bodies to which they were affiliated. Sampling and ascertainment are described in detail elsewhere (Taylor et al, 2005). Psychiatric morbidity was estimated using the General Health Questionnaire–12 (GHQ–12; Goldberg & Williams, 1988). Harmful alcohol consumption was screened using the World Health Organization’s Alcohol Use Disorders Identification Test (AUDIT; Saunders et al, 1993). Impaired clinical performance was measured using a scale adapted from Firth-Cozens et al (1997), whereby the frequency that stress at work had caused irritability with colleagues, irritability with patients and reduction in standards of care (such as taking short cuts) was rated on four-point scale from ‘never to my knowledge’ to ‘at least weekly’. Early retirement was defined as intending to retire aged \( \leq 55 \) years. Demographic measures included gender, age and marital status. Hierarchical logistic regression models were developed. Each univariately significant demographic variable was entered into a multivariate model (model 1) followed by the additional impact, if any, of poor mental health (GHQ score \( \geq 4 \); model 2). Analysis of relationships with planned early retirement necessarily excluded consultants aged over 55. Individuals with missing data were excluded on a test-by-test basis. Missing data constituted less than 5% except AUDIT scores (9%) and intended retirement age (7%). All tests were two-tailed, using a 5% significance level, and all analysis was conducted using SPSS v.12.0.1 for Windows.

RESULTS

Questionnaires were returned by 1308 consultants (73%), of whom 19% (251) were women, 89% (1151) were married or cohabiting, 4% (52) were aged \(< 35 \) years, 41% (534) were aged 36–45, 36% (473) were aged 46–55 and 19% (242) aged \( \geq 55 \) years (Table DS1 in the data supplement; 17% of consultants (207) reported consuming hazardous quantities of alcohol; 33% (432) reported that, at least monthly in the last 6 months, stress at work had caused them to be irritable with colleagues; 16% (212) reported being irritable with patients; 17% (221) had reduced their standards of care through, for example, taking short cuts or not following procedures; and 18% (176) planned to retire early.

Male consultants were more likely to report harmful consumption of alcohol; consultants aged between 36 and 45 years (mid-aged) and those who were unmarried were more likely to report being irritable with colleagues; younger consultants were more likely to report being irritable with patients; male consultants and mid-aged consultants were more likely to report reducing their standards of care; and female consultants and mid-aged consultants were more likely to intend to retire early (Table 1: model 1). Poor mental health independently increased consultants’ likelihood of reporting all of these behaviours (Table 1: model 2).

DISCUSSION

These findings suggest that hospital consultants with poor mental health are substantially more likely to report harmful consumption of alcohol, being irritable with patients, being irritable with colleagues, reducing their standards of care at work and/or planning to retire early. In addition, we have shown that male and mid-aged consultants are particularly at risk.

Our study included a large national cohort of consultants from five specialties. Despite the sensitive nature of the survey questions, the response rate was high, giving us confidence that the sample is representative. The cross-sectional design limits interpretation of causality, and assessments of mental health and consultant behaviours relied upon self-report measures. However, both the GHQ–12 and the AUDIT have been shown to be reliable and valid screening tools.

To our knowledge, this is the first study to examine the relationship between the mental health of hospital consultants and behaviours that detract from patient care.

Our finding that male consultants were more at risk of harmful alcohol consumption is consistent with findings from general population surveys. Perhaps more surprising is the increased risk of impaired clinical performance and planned early retirement
at mid-age. Consultants appear to be particularly vulnerable at this phase in their career (e.g. Spickard et al., 2002), which may be due to the fact that the honeymoon period of achieving consultant status is over but retirement is far from sight. They will have accumulated all of the responsibility and associated pressures, and are also more likely to have high demands on their time from home, with many having young families.

This study adds to the growing literature which highlights the critical importance to patients, as well as to doctors, of identifying approaches that protect consultants' mental health and support them to practice medicine safely throughout their careers.

ACKNOWLEDGEMENTS

We thank the consultants who took part in the survey, Royal Colleges of Physicians and Radiologists, British Society for Gastroenterology and British Association of Surgical Oncologists for their support. This study was jointly funded by the Charitable Foundation of Guy's and St Thomas' Hospitals (London, UK) and Cancer Research UK.

REFERENCES


Table 1: Predictors of harmful alcohol use (score ≥ 8 for men, ≥ 7 for women on the Alcohol Use Disorders Identification Test), impaired performance at work (being irritable with colleagues or with patients, or reducing standards of care, at least monthly in the past 6 months) and intention to retire early (aged ≤ 55 years): logistic regression

<table>
<thead>
<tr>
<th>Behaviour by variable</th>
<th>Model 1 OR (95% CI)</th>
<th>Model 2 OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harmful use of alcohol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (male)</td>
<td>1.6 (1.0–2.4)*</td>
<td>1.7 (1.1–2.6)*</td>
</tr>
<tr>
<td>GHQ–12 score ≥ 4</td>
<td>NA</td>
<td>2.0 (1.4–2.7)***</td>
</tr>
<tr>
<td>Irritable with colleagues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear term</td>
<td>3.1 (1.2–7.9)***</td>
<td>2.8 (1.1–7.6)***</td>
</tr>
<tr>
<td>Quadratic term</td>
<td>0.8 (0.7–0.9)</td>
<td>0.8 (0.7–0.9)</td>
</tr>
<tr>
<td>≤ 35</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>36–45</td>
<td>1.5</td>
<td>1.4</td>
</tr>
<tr>
<td>46–55</td>
<td>1.3</td>
<td>1.3</td>
</tr>
<tr>
<td>&gt; 55</td>
<td>0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>1.5 (0.9–2.3)**</td>
<td>1.4 (0.9–2.2)**</td>
</tr>
<tr>
<td>Separated</td>
<td>2.2 (1.3–3.8)</td>
<td>1.9 (1.1–3.4)</td>
</tr>
<tr>
<td>GHQ–12 score ≥ 4</td>
<td>NA</td>
<td>3.4 (2.6–4.3)***</td>
</tr>
<tr>
<td>Irritable with patients</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear term</td>
<td>0.8 (0.7–1.0)*</td>
<td>0.8 (0.7–1.0)*</td>
</tr>
<tr>
<td>Reduced standard of care</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear term</td>
<td>2.2 (0.7–7.4)*</td>
<td>1.9 (0.6–6.6)*</td>
</tr>
<tr>
<td>Quadratic term</td>
<td>0.8 (0.7–1.0)</td>
<td>0.9 (0.7–1.1)</td>
</tr>
<tr>
<td>≤ 35</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>36–45</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>46–55</td>
<td>1.1</td>
<td>1.0</td>
</tr>
<tr>
<td>&gt; 55</td>
<td>0.7</td>
<td>0.6</td>
</tr>
<tr>
<td>Gender (male)</td>
<td>2.2 (1.4–3.4)***</td>
<td>2.4 (1.5–3.9)***</td>
</tr>
<tr>
<td>GHQ–12 score ≥ 4</td>
<td>NA</td>
<td>2.8 (2.0–3.8)***</td>
</tr>
<tr>
<td>Early retirement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender (male)</td>
<td>0.3 (0.2–0.5)***</td>
<td>0.3 (0.2–0.5)***</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear term</td>
<td>5.0 (0.6–38.0)**</td>
<td>5.9 (0.7–47.5)***</td>
</tr>
<tr>
<td>Quadratic term</td>
<td>0.6 (0.4–1.0)</td>
<td>0.6 (0.4–1.0)</td>
</tr>
<tr>
<td>≤ 35</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>36–45</td>
<td>1.2</td>
<td>1.3</td>
</tr>
<tr>
<td>46–55</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>GHQ–12 score ≥ 4</td>
<td>NA</td>
<td>2.3 (1.6–3.2)***</td>
</tr>
</tbody>
</table>

*P < 0.05. **P < 0.01. ***P < 0.001. NA, not applicable.

1. Model 1: demographic predictors only.
3. Female consultants were the reference category for gender.
4. GHQ–12 score < 4 was the reference category for GHQ–12.
5. Effect of age modelled using a quadratic function to describe an inverse-U-shaped relationship. Resultant odds ratios for each age category are also given, using < 35 as the reference.
6. Married/cohabiting consultants were the reference category for marital status.
Correspondence

EDITED BY KIRIAKOS XENITIDIS and COLIN CAMPBELL

Contents  
- Is DISC1 really a gene predisposing to psychosis?  
- Hippocampal and amygdala volume reductions in first-episode schizophrenia  
- Effectiveness of cognitive–behavioural intervention by mental health nurses in schizophrenia  
- Contingency management for substance misuse  
- Depression and anxiety after myocardial infarction  
- High female suicide rates: ecological fallacy or sad reality?  
- Self-poisoning with pesticides in India

Is DISC1 really a gene predisposing to psychosis?

In their editorial on chromosomal abnormalities and psychosis Muir et al (2006) concluded that DISC1 ‘is an important modulator of risk for schizophrenia and severe affective disorder in people without cytogenetic abnormalities and may also influence cognition and brain structure in the general population’. They base their conclusions on work that originated in the finding of a rearrangement between chromosomes 1 and 11 in a single large family with polymorphic psychiatric syndromes (Millar et al, 2001). The two genes (DISC1 and DISC2) that they are concerned with were identified at the breakpoint and by linkage analysis were postulated to be relevant to psychiatric disease within that family.

Muir et al argue that these findings are relevant to schizophrenia in general. However, the evidence is less compelling than they suggest. Figure 1 presents the findings of the three largest linkage studies to date in relation to the location of DISC1 on chromosome 1 (the location of another ‘candidate gene’ RG34 is also shown). Each study included over 300 sibling pairs with schizophrenia or schizoaffective disorder and each included markers spaced at 10 cM intervals across the genome. The Lod (log of the odds) score is a measure of linkage – transmission of a disease state with particular genetic markers within families – and values above 3 are generally taken as significant evidence for linkage. In these three studies there is no evidence of linkage at the DISC1 locus or elsewhere on chromosome 1. The two claims of linkage made in Table 1 of Muir et al’s editorial relate to post hoc subdivision of one of these populations by diagnosis and to a finding in a separate smaller Finnish study. Given the ubiquity of psychosis across populations, and the relative uniformity of incidence of the core syndrome, and in the face of lack of evidence of linkage in populations of over 1000 sibling pairs (CROW, 2007), it is difficult to see that DISC1 can have an ‘important role in the development of psychosis’ as Muir et al argue. The evidence has been overinterpreted.

**Fig. 1** Linkage studies of DISC1 in sibling pairs with schizophrenia or schizoaffective disorder. —, Delisi et al, 2002 (382 sibling pairs); —, Williams et al, 2003 (353 sibling pairs); —, Suarez et al, 2006 (409 sibling pairs).

Authors’ reply: Professor Crow takes issue with our view that DISC1 is important to schizophrenia in general and is not restricted to the initial family in which disruption of this gene was reported. His argument is based on a selected set of sibling studies whose results do not support linkage anywhere on chromosome 1. This finding was unsurprising in view of the lack of power of such studies in the presence of genetic heterogeneity in schizophrenia susceptibility, which was not mentioned by Professor Crow. We and a large number of other workers in the field consider that such locus heterogeneity is highly likely and have shown that the sib-pair strategy has limited power to detect a locus that contributes less than 20% of the variance (Macgregor et al, 2002). Where heterogeneity is expected then linkage analysis, especially of extended multiplex pedigrees, and gene candidacy identified through the investigation of psychosis-associated karyotype anomalies are appropriate research strategies. Where there is a priori evidence from cytogenetic and linkage studies (such as the Finnish studies mentioned in the editorial) then the case–control association approach provides a useful resource to delineate potential population haplotype distortions that may indicate underlying functional mutations.

We would therefore disagree strongly with Crow in his statement that we have ‘overinterpreted’ the importance of DISC1 and commend an excellent review of schizophrenia neurobiology which emphasises heterogeneity (Ross et al, 2006). Although our theoretical framework differs from that of Bleuler (1950), we feel that the recent genetics and neurological discoveries are in agreement with his position that there is indeed a ‘group of schizophrenias’.


Hippocampal and amygdala volume reductions in first-episode schizophrenia

Steen et al (2006) performed a systematic review and meta-analysis of cross-sectional and longitudinal magnetic resonance imaging (MRI) studies of brain volumes in patients with first-episode psychosis and healthy controls. Despite some methodological differences, the findings were in line with a recent meta-analysis performed by our group (Vita et al, 2006).

A significant decrease in hippocampal but not amygdala volumes was found in patients at illness onset in both reviews. Another relevant paper reporting amygdala and hippocampal volumes in a large sample of patients with first-episode schizophrenia was published after these two meta-analyses (Velakoulis et al, 2006). Thus we considered it worthwhile to conduct a new set of meta-analyses including these MRI data.

The results of the new meta-analyses for hippocampus (7 studies, 290 patients, 355 controls) and amygdala (5 studies, 218 patients, 175 controls) confirmed our previous findings. Even with the inclusion of the study of Velakoulis et al (2006), the composite effect sizes for the hippocampus remained significant ($d = 0.357$, $95\%$ CI 0.208–0.541 for the right hippocampus and 0.574, $95\%$ CI 0.403–0.742 for the left hippocampus) whereas those for the amygdala were not ($d = -0.046$, $95\%$ CI −0.247 to 0.154 for the right amygdala and 0.025, $95\%$ CI −0.175 to 0.226 for the left amygdala).

These results, in line with those of Steen et al (2006), support the hypothesis of different patterns of involvement of temporal limbic structures over the course of schizophrenia, with the hippocampus affected earlier than the amygdala. In our opinion, these findings have important implications for future neurobiological studies of schizophrenia and emphasise the importance of longitudinal studies to address the issue of different times of occurrence and progression of brain abnormalities in people with first-episode schizophrenia.

Effectiveness of cognitive–behavioural intervention by mental health nurses in schizophrenia

Turkington et al (2006) report on outcomes of an effectiveness trial of brief cognitive–behavioural therapy (CBT) by mental health nurses in schizophrenia. Unfortunately there are flaws in the methodology, which casts major doubts on the validity of the study (Quintin et al, 2000). First, although the authors claim to have a control group, it seems that patients in the control group did not have a placebo-like intervention; for example, the nurses could have spent the same amount of time with the patients without providing the CBT intervention. What is more surprising is that the study was powered to give a 90% chance of detecting only a 25% level difference in overall symptoms at the 0.01 level of significance. A 25% difference between a treatment and non-intervention group can easily be accounted for by a placebo effect. It is well known that the placebo response rate is usually around 30% in psychiatric trials. For over 50 years the inclusion of a placebo control group has been the standard for determining the efficacy of an intervention. Without an adequate comparison group and without adequate comparison conditions, it is impossible to differentiate any specific effects from other ‘non-specific’ factors, including chance variation, regression to the mean, healthcare provider attention, treatment credibility and rationale, persuasion, patient expectation effects, researcher allegiance effects, effort justification, spontaneous remission, demand characteristics, etc. (Lohr et al, 1999).

Given the lack of a true control group this study would be called nothing but an open-label trial. Open-label trials require at least a 50% level difference in overall symptoms between baseline and post-intervention response; moreover they do not require huge numbers of patients to show a tendency towards improvement.

Authors’ reply: We believe that Dr Alam has misunderstood the difference between efficacy and effectiveness research. The national guidelines on the clinical management of schizophrenia (National Institute for Clinical Excellence, 2002) confirmed CBT to be an evidence-based treatment for persistent symptoms of schizophrenia. However, that decision was based almost entirely on efficacy trials where CBT was given by expert therapists to highly selected samples of people with schizophrenia without comorbidities and using an active comparator such as befriending or supportive counselling (e.g. Sensky et al, 2000). Expert therapists and uncomplicated patients are
rare in clinical psychiatric practice. Therefore, the next step was to design an effectiveness trial to see whether mental health nurses, without prior experience of CBT could be trained over a short period and then supervised to effectively and safely deliver brief CBT to large numbers of people with schizophrenia in the community. As this involved raters being masked to group allocation, this was therefore not an ‘open-label’ trial.

In relation to the effect size, it is certainly true that when an antipsychotic is compared with a placebo in drug-naive patients a much larger effect is demonstrable. The patients recruited to this trial were, however, almost entirely stabilised on antipsychotics and had already achieved such improvement from them. The effect size with any psychological treatment added to antipsychotics is always likely to be less than that initially achieved by the medication. We acknowledge that the effect size on symptoms at follow-up is modest but the impact on relapse is significant, clinically and in terms of resource savings, for such a brief intervention.

**Contingency management for substance misuse**

Petry (2006) provides a welcome review of contingency management in substance misuse settings and expresses surprise that it has not been employed more widely in Europe, particularly given the greater acceptance of ‘harm minimisation’ here than in the USA, where contingency management has been championed. This is broadly true but some UK drug services are experimenting with interventions informed by reinforcement principles.

The injectable opiate clinic at the Chelsea and Westminster Hospital in London has for some years used reinforcement principles to target illicit opiate and crack cocaine use. Urine samples are regularly tested and the results used alongside clinical judgement to determine the proportion of a client’s total daily opiate dose which may be administered intravenously as opposed to orally. In this way, access to injectable rather than oral opiate preparations is the ‘reward’ for positive behaviour. Staff increase or decrease the injectable proportion of the client’s prescription depending on the client’s stability.

As a first step towards developing an intervention study (Medical Research Council, 2000) we completed qualitative interviews with staff and clients to assess attitudes towards the further development of reinforcement methods. Staff and clients both cautiously supported reinforcement principles, and staff perceived clients to be more stable and less likely to use illicit substances under the present reinforcement scheme. Nevertheless, challenges were also highlighted. Most staff had reservations about developing voucher-based contingency management, citing possible increased workloads and a potential for damage to staff–client relationships. Despite a strong commitment to harm minimisation strategies at the clinic, some staff also had ethical objections to the development of voucher-based contingency management.

Our study was small and more research is required to explore the feasibility of voucher- or prize-based contingency management. However, as Petry emphasises, contingency management strategies have a good evidence base in a complex and challenging client group where positive outcomes are elusive. It is surely time to evaluate whether contingency management has a place in UK drug treatment services. Our work suggests that debate about the theoretical basis of contingency management and its ethical implications is needed to win support for experimentation among hard-pressed drug treatment workers in the UK.

**Depression and anxiety after myocardial infarction**

Dickens et al (2006) stress the importance of detection and treatment of anxiety and depression for quality of life after myocardial infarction and point to the mediating role of energy and fatigue.
We agree that depression following myocardial infarction predicts long-term quality of life and we recently showed that this effect persists after controlling for cardiac condition and quality of life at 3 months post-myocardial infarction (de Jonge et al., 2006). However, it is unclear whether and how detection and treatment of depression can counter these effects. In the SADHART study Glassman et al. (2002) found that the effects of sertraline were modest and appeared to be restricted to depression with an onset before the infarction, but Dickens et al. found that depression and anxiety which were present before myocardial infarction did not predict quality of life. In the ENRICHD trial (Berkman et al., 2003), cognitive—behavioural therapy had modest effects on depressive symptoms at 6 months post-infarction in patients with depression and social isolation, but these effects diminished over time. In the EXIT trial (Appels et al., 2005), where the focus of treatment was explicitly on vital exhaustion, only some intervention effects were observed and these were modified by the presence of a previous cardiac history.

We agree with Dickens et al. that there is a need for improved detection and treatment of depression and anxiety following myocardial infarction but several questions need to be addressed. These include: 'can the effects of depression and anxiety be linked to specific subgroups of emotional disorders based on symptoms and/or onset?'; 'can interventions that were developed in general psychiatry be applied to depression post-myocardial infarction or should they be adapted?'; and 'how can psychiatric interventions be integrated into regular cardiac aftercare?'


**de Jonge, P., Spijkerman, T. A., van den Brink, R. H. S., et al. (2006)** Depression following myocardial infarction is a risk factor for declined health-related quality of life and increased disability and cardiac complaints at 12 months. Heart, 92, 32–39.


**P. de Jonge** Department of Internal Medicine and Department of Psychiatry, University Medical Centre Groningen, University of Groningen, Hanzeplein 1, PO Box 30.001, Groningen, The Netherlands. Email: Peter.de.jonge@med.umcg.nl

**J. Ormel** Department of Psychiatry, University Medical Centre Groningen, The Netherlands doi: 10.1192/bjp.190.3.272b

**Authors’ reply:** We agree that although observational studies have shown that depression is associated with subsequent impairment in health-related quality of life in coronary heart disease, intervention studies have failed to provide convincing proof that treating depression improves this outcome. Previous intervention studies have not addressed this question satisfactorily because the SADHART study (Glassman et al., 2002) was not sufficiently powered to demonstrate the efficacy of antidepressants in coronary heart disease and the ENRICHD study (Berkman et al., 2003) did not anticipate very high rates of spontaneous remission of depression or unplanned prescription of antidepressants in the control group. The results of these trials, however, together with our own results are valuable for planning future treatment trials.

We also agree that there are many unanswered questions relating to the nature of the association between depression and negative outcomes in coronary disease. As mentioned by de Jonge & Ormel, the timing of the onset of depression (Dickens et al., 2004a), the specific aspects of depression or anxiety that are associated with poor outcome and the possibility of vulnerable sub-populations of patients (such as those without social support) (Dickens et al., 2004b) require further investigation. Furthermore, whether the association between depression and negative outcomes in coronary disease is the result of residual confounding by severity of heart disease (Dickens et al., 2005) remains unsolved. Further research is required to address these questions, although it is likely that most will only be convincingly resolved through intervention studies.


**High female suicide rates: ecological fallacy or sad reality?**

Yip & Liu (2006) present a demographic perspective of female suicide in China, the only country in which the suicide rate is higher among women than men. However, this reversed gender representation also exists in certain communities in other countries. In the Indian subcontinent suicide rates are higher in men than in women but the difference is lower than in most countries: the male:female suicide ratio in India is 1.3:1 (Cheng & Lee, 2000). Suicide among immigrants from the Indian subcontinent to Britain was higher among young married women than men (Soni Raleigh et al., 1990). Tadros & Salih (2006) also reported that significantly more Asian women than Asian men killed themselves in Birmingham and Solihull, a clearly reversed gender ratio compared with suicide in the White population and in other ethnic groups in Birmingham and the UK as a whole.

Suicide terrorism is not an egotistic suicide but none the less is a form of fatal self-harm in the legal and human sense and has a distinct underlying political, individual and social logic. The support of and acceptance by the attackers’ own communities ensure an endless supply of volunteers who seek ‘voluntary violent death’ in a bizarre act of so-called martyrdom, in order to promote what they firmly believe to be a just cause. Women carried out 15% (64) of such attacks over the past 25 years (Pope, 2005). Chechen women carried out 60% of all suicide bombings in Russia.
and 70% of such attacks were executed by Kurdish women in Turkey (Pope, 2005). There is also a high proportion of women suicide bombers in the Tamil Tigers (30%), al Qa’ida, which associates itself with Islamic fundamentalism, never used female suicide attackers from its formation in 1993 until the tragic attack in Jordan in 2005. In general, women are at a lower risk of suicide than men and a protective effect of child-bearing in terms of suicide risk has been postulated (Catalan, 2000). This does not appear to apply to female suicide bombers or to some countries and cultures in which gender representation in suicide is reversed. A higher female:male suicide ratio is not unique to China. The significantly higher rate of female suicide observed outside China is not an ‘ecological fallacy’ but a sad reality.


Honorary Senior Lecturer, Liverpool University

We also believe that socio-economic deprivation and poor social support – the ‘sad reality’ – faced by young women in rural China are underlying causes of the high suicide rates. Like the young married Indian women in Britain, there is some indication that young married women in rural China might be at high risk (Pearson et al, 2002). This reminds us that the lives of married women differ greatly across regions, countries, cultures and economies, and there is a need to avoid oversimplification when describing suicide in different countries; one size does not fit all.

Over 60% of the world’s suicides occur in Asian countries where low male:female ratios for suicide are common (Yip et al, 2000). Although the official male:female ratio for suicide in India was still greater than 1 (1.2:1 in 2002), the ratio was 0.8 among those aged 14 or below (World Health Organization, 2006). However, unlike China (Yip & Liu, 2006), the small size of this population subgroup meant that the national male:female ratio remained greater than 1. (This is the essence of our ecological fallacy argument.) In addition to specific social factors, the similarity in the methods of suicide used by males and females, together with the poor access to medical facilities, might explain the low male:female ratio in India and China. Restricting access to pesticides will prevent many suicides in Asia. In the long term improving economic and educational opportunities, especially for rural women in deprived areas, raising awareness of depression and better treatment will be pivotal for preventing suicides.

E. Salib Peasley Cross Hospital, St. Helens, and Honorary Senior Lecturer, Liverpool University Email: esalib@hotmail.com

G. Tadros Birmingham and Solihull Mental Health Trust, Birmingham, UK
doi: 10.1192/bjp.190.3.273a

Authors’ reply Salib & Tadros highlight the important issue of high female suicide rates among Indian migrants and the use of female suicide bombers. Like the high suicide rates among young females in rural China (Yip & Liu, 2006), these deviations from the general pattern should not be discounted as mere exceptions but should be considered as representative of the distressing situations faced by some women in Asia.


Bertolote et al (2006) report the global response to deaths from pesticide poisoning. Suicide rates in southern India have been reported to be high (Joseph et al, 2003; Aaron et al, 2004; Abraham et al, 2005; Prasad et al, 2006), with 1741 suicides documented in a population of about 100 000 from 1986 to 2005. Hanging (804 of 1741, 46.2%) and poisoning (746 of 1741, 42.8%) were the methods most commonly employed. Although people under 40 years tended to use poisons, older people tended to choose hanging (χ²=36.71, d.f.=4, P<0.001). Significantly more males (465 of 984, 47.3%) than females (281 of 757, 37.1%) (χ²=17.6, d.f.=1, P<0.001) chose death by poisoning. There was no significant change in the overall rate of suicide or the method employed during the period. Detailed analysis of the data from 2001 to 2005 revealed that only 68% of the fatal episodes of self-poisoning were a result of ingestion of pesticides.

Self-poisoning with pesticides is a significant public health problem in low- and middle-income countries. The majority of such poisoning occurs in rural agrarian households. Some suggestions to reduce such deaths are currently difficult to implement. Enforcing the hazardous chemicals and wastes conventions to restrict and control the sale and use of pesticides in such regions is no small task and requires major political, administrative, financial and social commitment. Given the many competing demands on limited governmental resources in low- and middle-income countries, such protocols are difficult to implement. In addition, the improved recognition and treatment of mental illness may not have a significant impact on the overall suicide rate as many people in the low- and middle-income countries who die by suicide do not have severe mental illness. Rather, the majority of such attempts are impulsive and follow stressful life events. Although reducing accessibility to pesticides will decrease such impulsive attempts and consequent deaths, social, economic and cultural factors must also be addressed to make a real difference. Thus, although the World Health Organization’s intersectoral global initiative is a step in the right direction, it is imperative that practical issues related to its implementation are discussed. It is necessary to consider
strategies to encourage governments to set up suicide prevention programmes to reduce suicide rates in populations as a whole rather than method- and site-specific rates.

Suicide in low- and middle-income countries is not only a medical and public health problem but is also related to economics and culture. A coordinated and a comprehensive response is needed to make any impact.


R. Alex, J. Prasad, A. Kuruvilla, K. S. Jacob Christian Medical College, Vellore 632002, India. Email: sanju@cmcvellore.ac.in doi: 10.1192/bjp.190.3.274a

One hundred years ago

The increase of temperance

The Inland Revenue returns show a steadily progressive decrease in the consumption of beer and spirits in the United Kingdom since 1899; that is in encouraging contrast with the equally steady but more rapid increase up to that date.

The beer consumption in 1899–1900 was 32.2 gallons per head of the population, making a total of 36.5 million barrels, but in 1905–1906 this had fallen to 27.9 gallons per head and to 33.5 million barrels.

The spirit consumption has also fallen each year from 1.17 gallons per head and a total of 48 million gallons for 1889–1900 to .90 gallons per head and 39.1 million gallons in 1905–1906.

The reduction in the consumption of spirits is very striking, and in addition to the reduction in the total quantity of beer consumed there is to be added the large increase in the proportion of the lighter beers of home and foreign manufacture.

Pauperism, crime, and insanity are so largely attributable to the abuse of alcoholic drinks that the statistics of each should be carefully watched during the next few years for any indication of an improvement. It is, of course, possible that this reduction may be due only to the greater moderation from necessity or improved habits of the middle and upper classes only, although it would appear to be too large to be thus explained.

Abuse of alcohol, in the statistics of the causes of insanity, has fluctuated very little for many years past, so that any distinct diminution would be very significant, and should encourage a still more vigorous crusade in favour of true temperance – the use without abuse of the cup that cheers and may inebriate.

REFERENCE

Journal of Mental Science, January 1907, 146.

Research by Henry Rollin, Emeritus Consultant Psychiatrist, Horton Hospital, Epsom, Surrey
doi: 10.1192/bjp.190.3.275
Lynette Jackson’s book is an excellent addition to the growing number of publications on the history of psychiatry in colonial Africa. Other books have dealt with the contribution of colonial psychiatrists to psychiatric theories (McCulloch, 1995), the history of asylums in south-west Nigeria (Sadowsky, 1999) and how race and culture acted as markers of difference in colonial asylums (Vaughn, 1991).

Jackson’s contribution to this subject is unique. It is a detailed study of one institution, Ingutsheni Asylum, tracing its origins, examining the conditions of daily life there and investigating the different pathways to care for men and women. Built on the site of the harem of the last independent Ndebele king, Lobengula Khumalo, it was opened in 1908 and was an addition to asylums in Robben Island, Kissey, Victoriaborg, Calabar, Accra and Yaba. From its inception, Ingutsheni received both African and European patients. The intention was that the African patients remained there until they were discharged or recovered (or escaped). For the European patients, it was simply a staging post on the way to one of South Africa’s asylums.

The vast majority of male African patients (inmates) were migrant labourers, men forced by poverty from rural areas far beyond Southern Rhodesia. A migrant labourer’s existence was harsh and the working conditions brutal. Jackson examines the routes into the asylum for these men, drawing attention to how behavioural transgression came to the notice of the colonial authorities. She also describes the pathway of admission for females. Mobile and unpaired, or ‘stray’, women, were regarded as exhibiting deviant behaviour and liable to be admitted to the asylum merely for being unaccompanied within colonial urban centres. Jackson describes the extent to which the colonial authorities attempted to maintain within the asylum the divisions present in wider society. This was most salient with regard to White women.

This book demonstrates how profitable it is to re-examine the institutions of colonial Africa. Asylums recreated the divisions within wider society, amplified differences and provided case material for ethnopsychiatrists to construct theories about African culture, the so-called ‘African mind’, often in an effort to sustain the colonial project.

addressed and twelve different treatment options presented. There is also discussion of investigational strategies for the treatment of rapid cycling, mixed episodes and atypical bipolar mood disorder, focusing on study design and offering suggestions for study methodology for this challenging condition.

The book is an easy and comprehensive read. Providing insight into present knowledge of the diverse manifestations along the full spectrum of bipolar disorder, it gives an overview of gaps in knowledge remaining to be studied. The book’s strength is that it not only determines issues that are weakly presented in the research arena but it also looks at methodological and study design issues that can help to improve future research. The weakness is that when addressing effectiveness of specific pharmacotherapeutic possibilities it does not always fully address the issue of adverse events. This would include those capable of resulting in patient- or physician-initiated discontinuation of treatment, problems that could arise from using specific drug combinations and the genetic differences that may be important in this field. This is an interesting book worthwhile for both clinicians and researchers in the field of bipolar disorders.

Ingeborg Wilting  Utrecht University, Utrecht, Holland. Email: omsmails@zoonet.nl  
doi: 10.1192/bjp.bp.105.019992

Alcohol, Gender and Drinking Problems: Perspectives from Low and Middle Income Countries

Alcohol consumption is enmeshed in social and cultural norms and problems relating to drinking can not be considered in isolation from these. Because the place of gender within social and cultural interactions is unique and has important ramifications for the behaviour and habits of men and women, this book addresses a topic of central significance to the consideration of how alcohol consumption may affect health, defined here in the broadest sense.

The book presents the reports of the project, ‘Gender, Alcohol and Culture: an International Study’ (GENACIS), conducted in Argentina, Brazil, Costa Rica, India, Mexico, Nigeria, Sri Lanka and Uganda. The first of its ten chapters provides an educative discussion of why it is important to examine gender differences in the use of alcohol. Chapters two through nine are reports of the project from the participating sites. Chapter ten attempts an integration of the main findings of the project. It shows that, when it comes to the use of alcohol, grouping of countries on the basis of development or region may be rather simplistic as the differences within these groupings are often large and complex.

Epidemiological evidence suggests that differences between males and females in regard to their use of alcohol are narrowing, with more recent birth cohorts showing closer similarities than earlier ones. Such narrowing of gender differences is probably more pronounced in low- and middle-income countries, especially those of Africa and Asia, where traditional restrictions on female drinking are beginning to wane as a result of various social changes, not least those related to urbanisation and globalisation. Anyone wishing to examine how social changes influence alcohol consumption will benefit from an understanding of the trend in gender patterning of drinking provided this book.

The book is enriched by the broad cultural contexts in which the studies were conducted. However, it has to be read within the constraints of the methodology of GENACIS. The focus of the surveys was on documenting gender differences in drinking patterns in the various study sites. Even though several of those sites used epidemiological approaches to sample respondents, and their results can be considered as representative of the regions where the studies were conducted, readers need to be aware that the data presented are not national profiles of drinking behaviour and are certainly not meant to highlight cross-national comparisons of alcohol use. What the book provides is a rich source of information about the dynamics of alcohol use in which the gender of those who drink and those who do not offers an opportunity for us to understand the social influences shaping trends in alcohol consumption. That information should interest policy makers across the globe.

Oye Gureje  University of Ibadan, Ibadan PMB 5116, Nigeria. Email: ogureje@comui.edu.ng  
doi: 10.1192/bjp.bp.106.029702

The science of well-being
Edited by Felicia Huppert, Nick Bayliss & Barry Keverne. Oxford University Press. 2005. 546pp. £80.00 (hb); £35.00 (pb). ISBN 0198567510 (hb); 0198567519 (pb)

This timely book coincides with initiatives (National Institute for Mental Health in England, 2005) for mental well-being at individual, neighbourhood and societal levels.

Argyle (1992) pioneered UK research on ‘happiness’ and health, developing the Oxford Happiness Inventory. Influential commissioners scorned positive outcome measures for funding psychiatric services, claiming ‘It’s not our job to make people happy’. Cronin de Chavez et al (2005) reviewed weaknesses in past research: single disciplines working in isolation cannot capture the multi-faceted nature of wellness. A trans-disciplinary meeting on the sciences of well-being took place in 2003 at the Royal Society. Unforgettable talks delivered there by Americans Fredrickson, Seligman and Putnam can be found among the 20 chapters of this new book.

The Science of Well-Being overlaps, but only in part, the transactions of that seminal event (Royal Society, 2004). Since 2003, interest in wellness has grown. Kahnemann contributes an excellent
Developments in international, cultural perspectives (‘Living, and thinking about it’). American and British authors dominate this book, but examples from low- and middle-income nations are included in Delle Fave and Massimini’s chapter ‘The relevance of subjective well-being to social policies’. Their insight into well-being for people with disabilities, related to the policies’. Their insight into well-being for middle-income nations are included in this book, but examples from low- and ‘Living, and thinking about it’).

While unlikely to convert many heretics, this book does make clear what is needed if social capital research is ever to deliver on the promises of its evangelists. If nothing else, the concept has encouraged people to look afresh at social ills and has inspired genuine efforts to relieve suffering and hardship.

Scott Weich, Health Sciences Research Institute, Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK.
Email: s.weich@warwick.ac.uk
doi: 10.1192/bjp.bp.106.026070

Social Capital and Mental Health
ISBN 1843103559

Social capital, embraced by New Labour and the World Bank, is currently the intellectual equivalent of Marmite, loved and loathed in equal measure. Cause of, and solution to, all of life’s problems – or Emperor’s new clothes? Like most agnostics, the editors of this book are not entirely sure but hedge their bets, just in case. Arguments for and against are marshalled in the best British tradition of fair play. On the one hand, social capital is the ‘missing link’ of social epidemiology, the glue required for effective societal functioning; on the other, it is poorly theorised, inadequately defined and unreliably measured. Up until now an industry in social capital research has been driven, and sometimes undermined, by the intuitive appeal and conceptual elasticity of the term itself. While this makes for great sound bites, the rhetoric is rarely matched by high-quality evidence. And when the findings do not fit, the theory is changed instead. If you cannot explain something by too little social capital then maybe the problem lies in too much of the wrong kind?

This neat, readable little book confronts these inconvenient truths head on, and makes an excellent starting point for sceptics, too. The editors have sandwiched reports from five innovative studies in between balanced reviews of current theory and practice. Even if there is a hint that the best bits of cutting-edge research have been saved for peer-reviewed publication, these primary research narratives make compelling reading. Where else would you find an evocative account of an idyllic summer’s afternoon on a city farm in Camden (complete with blackberry picking) juxtaposed with a death-defying description of working with teenagers in urban Columbia, homicide capital of the world?

Scott Weich, Health Sciences Research Institute, Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK.
Email: s.weich@warwick.ac.uk
doi: 10.1192/bjp.bp.106.021576

Social Capital and Mental Health
ISBN 1843103559

Social capital, embraced by New Labour and the World Bank, is currently the intellectual equivalent of Marmite, loved and loathed in equal measure. Cause of, and solution to, all of life’s problems – or Emperor’s new clothes? Like most agnostics, the editors of this book are not entirely sure but hedge their bets, just in case. Arguments for and against are marshalled in the best British tradition of fair play. On the one hand, social capital is the ‘missing link’ of social epidemiology, the glue required for effective societal functioning; on the other, it is poorly theorised, inadequately defined and unreliably measured. Up until now an industry in social capital research has been driven, and sometimes undermined, by the intuitive appeal and conceptual elasticity of the term itself. While this makes for great sound bites, the rhetoric is rarely matched by high-quality evidence. And when the findings do not fit, the theory is changed instead. If you cannot explain something by too little social capital then maybe the problem lies in too much of the wrong kind?

This neat, readable little book confronts these inconvenient truths head on, and makes an excellent starting point for sceptics, too. The editors have sandwiched reports from five innovative studies in between balanced reviews of current theory and practice. Even if there is a hint that the best bits of cutting-edge research have been saved for peer-reviewed publication, these primary research narratives make compelling reading. Where else would you find an evocative account of an idyllic summer’s afternoon on a city farm in Camden (complete with blackberry picking) juxtaposed with a death-defying description of working with teenagers in urban Columbia, homicide capital of the world?

Scott Weich, Health Sciences Research Institute, Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK.
Email: s.weich@warwick.ac.uk
doi: 10.1192/bjp.bp.106.021576

Social Capital and Mental Health
ISBN 1843103559

Social capital, embraced by New Labour and the World Bank, is currently the intellectual equivalent of Marmite, loved and loathed in equal measure. Cause of, and solution to, all of life’s problems – or Emperor’s new clothes? Like most agnostics, the editors of this book are not entirely sure but hedge their bets, just in case. Arguments for and against are marshalled in the best British tradition of fair play. On the one hand, social capital is the ‘missing link’ of social epidemiology, the glue required for effective societal functioning; on the other, it is poorly theorised, inadequately defined and unreliably measured. Up until now an industry in social capital research has been driven, and sometimes undermined, by the intuitive appeal and conceptual elasticity of the term itself. While this makes for great sound bites, the rhetoric is rarely matched by high-quality evidence. And when the findings do not fit, the theory is changed instead. If you cannot explain something by too little social capital then maybe the problem lies in too much of the wrong kind?

This neat, readable little book confronts these inconvenient truths head on, and makes an excellent starting point for sceptics, too. The editors have sandwiched reports from five innovative studies in between balanced reviews of current theory and practice. Even if there is a hint that the best bits of cutting-edge research have been saved for peer-reviewed publication, these primary research narratives make compelling reading. Where else would you find an evocative account of an idyllic summer’s afternoon on a city farm in Camden (complete with blackberry picking) juxtaposed with a death-defying description of working with teenagers in urban Columbia, homicide capital of the world?

Scott Weich, Health Sciences Research Institute, Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK.
Email: s.weich@warwick.ac.uk
doi: 10.1192/bjp.bp.106.021576

Social Capital and Mental Health
ISBN 1843103559

Social capital, embraced by New Labour and the World Bank, is currently the intellectual equivalent of Marmite, loved and loathed in equal measure. Cause of, and solution to, all of life’s problems – or Emperor’s new clothes? Like most agnostics, the editors of this book are not entirely sure but hedge their bets, just in case. Arguments for and against are marshalled in the best British tradition of fair play. On the one hand, social capital is the ‘missing link’ of social epidemiology, the glue required for effective societal functioning; on the other, it is poorly theorised, inadequately defined and unreliably measured. Up until now an industry in social capital research has been driven, and sometimes undermined, by the intuitive appeal and conceptual elasticity of the term itself. While this makes for great sound bites, the rhetoric is rarely matched by high-quality evidence. And when the findings do not fit, the theory is changed instead. If you cannot explain something by too little social capital then maybe the problem lies in too much of the wrong kind?

This neat, readable little book confronts these inconvenient truths head on, and makes an excellent starting point for sceptics, too. The editors have sandwiched reports from five innovative studies in between balanced reviews of current theory and practice. Even if there is a hint that the best bits of cutting-edge research have been saved for peer-reviewed publication, these primary research narratives make compelling reading. Where else would you find an evocative account of an idyllic summer’s afternoon on a city farm in Camden (complete with blackberry picking) juxtaposed with a death-defying description of working with teenagers in urban Columbia, homicide capital of the world?
We live in a knowledge-creating culture, where most of us regularly flip-flop between being learner and teacher. But we inhabit a bit of a vacuum when it comes to knowledge about knowledge. Few of us read medical education journals and the new generation of educational givens has something of the liberation movement about it, reacting against the conformity of lectures rather than emerging from an evidence base. Knowing how we get to know has suddenly become important, since, in common with North America and The Netherlands, we are about to embark on a massive reform of postgraduate medical education partly marketed as being based on contemporary learning theory.

So, Keith Sawyer’s multi-author book on the relatively new discipline of ‘learning science’ is timely. Most of us would certainly benefit from being more literate about what goes on in the learning environment and how this contributes to improved learner performance. A limitation of Sawyer’s book is that it is aimed mainly at school educationalists, with no reference to learning in medical environments. On the other hand it is a rich source of ideas and evidence on effective learning, with many lessons for psychiatry.

For example, there is a review of the nature of expertise. If you aspire to turn novices into experts, you have to have a pretty good idea of what an expert is. Also how they differ from novices, how you describe that difference in ways that can be translated into measurable outcomes and how you construct a learning environment in which the necessary skills can be acquired. Differences that emerge across a range of disciplines are the expert’s ability in ‘noticing’ and in pattern recognition; skills which affect the ability to rapidly identify problems and opportunities to act on them. The knowledge of experts also tends to be connected and organised around important ideas for their discipline. The lessons for learning include helping novices reflect on their own thinking, enabling them to gain deeper conceptual understanding and exposing them to increasingly complex social and technical environments.

A review of the evidence for what works best underlines the importance of learners being active participants. They need to build new knowledge onto existing knowledge, externalise and articulate their unformed and still developing understanding (articulating and learning go hand in hand), and engage in reflection, which is so critical for professionals working in complex practice. There is also a review of the importance of ‘arguing to learn’, based on the idea that science advances not by the accumulation of facts, but through debate and argumentation. The justification is that arguing involves elaboration, reasoning and reflection, all of which contribute to deep conceptual learning. Exposure to this so-called ‘collaborative argumentation’ helps professionals think critically and independently about important issues and contested values.

An interesting issue, particularly for psychiatrists, is the role of social context on learning. This is one area that medicine has always done well in, mainly because it has retained a largely apprenticeship model, the preferred way of learning expertise throughout recorded history. The so-called ‘cognitive apprenticeship’ model emphasises a focus on cognitive skills and processes and involves learning conceptual and factual knowledge used in a variety of contexts to solve real-world problems. Importantly, the apprenticeship model remains intact in the new foundation and run-through training programmes, where for the first time it is reinforced by a formal workplace-based assessment programme.

The Cambridge Handbook of The Learning Sciences is a bit like the Econo-mist; a little over-inclusive for most, but a really good review of those parts of the world you are interested in. Curiously, although there are sections on information technology and learning communities on the internet, as well as a fascinating piece on the importance of enabling conceptual change to take place in learners, there is virtually nothing on assessment. However, there is a tantalising plea to avoid the tendency to assess well-established routines and schemas and instead to examine ‘adaptive expertise’, or the ability to learn in a knowledge-rich environment. This connects with the idea that all students and trainees are seen as members of a knowledge-building community, where the aim is idea improvement as an alternative to progress towards a ‘truth’. And of course in a post-modern world there is no such thing as truth, only different versions of it. One of which is John Dewey’s version written in 1938 and quoted in the Cambridge Handbook: ‘Education is not preparation for life; education is life’.

Nick Rose  Department of Psychiatry. University of Oxford,Wormsden Hospital, Oxford OX3 7X, UK. Email: nick.rose@oxmhc-tr.nhs.uk

doi: 10.1192/bjp.bp.106.029678
From the Editor’s desk

PETER TYRER

INEQUALITY, STIGMA AND MENTAL HEALTH

Society is unfair and discriminatory. Last month I was trying to improve a patient’s home environment as part of the treatment called nidoetherapy (Tyrer et al, 2007) and was attempting to talk to a well-built neighbour whom my patient had identified as difficult and interfering. She did not particularly want to engage with me and after a few minutes, in what I remember as an unnecessarily loud voice, threatened to ‘knock me down’. I agreed with her that I was being just a little irritating but added that, while her proposed behaviour might seem justified to her at that moment, ‘unfortunately it would have a more serious consequence than if you knocked someone else down. Knocking a senior doctor down in the course of his community duties would be judged harshly by the courts and you may, quite unfairly in my opinion but society is like that, receive a long custodial sentence’. This made her pause and we ended our conversation on reasonably good terms with a promise to work hard on less aggravating behaviour.

Inequalities in society feed mental ill-health (Skapinakis et al, 2005) through a variety of means, including both societal and familial stigma (Lee et al, 2005), and are likely to be heavily involved in prematurity mortality (Ran et al, pp.237–242). Stigma is a fuzzy term that needs a good deconstructor, and we make a fine start with Thornicroft et al (pp.192–193), who note that we have for too long confined it to attitudes (where we can keep it at arm’s length) and not focused enough on behaviour (which makes it happen) and which, they argue, makes stigma into discrimination. Nevertheless, I hasten to add, this does not mean that the stigma scale developed by King et al (pp.248–254) is redundant, and the authors make a distinction between ‘felt’ and ‘enacted’ stigma that deals with the same concept. These papers will help in bringing stigma and all its nasty associations out into the spotlight of open enquiry. Unfortunately, psychiatric diagnosis shows equal evidence of stigma, inequality and, indeed, inefficiency, despite being much more exposed to scrutiny, and Baca-Garcia et al (pp.210–216) demonstrate in their commendably large sample just how often the labels we pin to our patients get changed over time, rather like the prices of goods in a supermarket altering with the weather.

We have always lived in an Orwellian society where some are ‘more equal than others’ and, in the case of severe mental illness, resources have been shifted from relatively expensive neglect in hospital to (equally expensive) intensive treatment in the community. Assertive outreach teams are more equal than community mental health teams as they are much better resourced and there have long been grumblings of inequity; the findings of Burns et al (pp.217–222) support both camps, as there seems to be no fundamental difference between assertive outreach and other community teams, although assertive teams do provide more comprehensive care as they reduce their case-load size. And so the unequal dance goes on, with odd penetrating genes in schizophrenia (McClellan et al, pp.194–199), selection of controls in case-control studies (Lee et al, pp.204–209), and the allocation of consultant care (Taylor et al, pp.268–269) all subject to the same lottery of variation. As for me, in my travels through the geographical inequality of London, I can give some comfort to the angry neighbour when I next see her. If she does knock me down and goes to a medium secure unit for her heinous offence, her chances of reconviction are much less than if she were a man (Coid et al, pp.223–229).

AN EDITOR DEPARTS

This month marks the retirement of Alan Lee, the second editor of Advances in Psychiatric Treatment, who took over from the founding editor Andrew Sims in 2003, after co-editing the journal since March 2002. At a moving farewell reception Alan explained the reasons for his departing this post in the prime of his prestige. He confessed openly his secret, hidden to all until now. He was a workaholic – he did not mind who knew as disclosure was his only hope. He asked those present, quietly but firmly, not to tempt him by little offers of work after he retired. He had continued craving for tasks of all sorts and even a small degree of re-exposure to them would quickly send him on a downward path. I promised to promulgate the message so that all others in the College and its wider community would be aware and help him to avoid temptation.

He leaves a journal in rude health and with great prospects; it is just about to be launched in North America as a major educational initiative, and its combination of independent thought, pertinent advice and clear exposition will go down well there. We welcome Joe Bouch as the new Editor, who has a gardening job on his hands, indicated by a verse from the departing doggerel I gave to Alan as he left with only idleness on his mind.

So how will we do without him
Now he’s decided to hang up his boots
We’ll just have to weed round his plants grown from seed
And make sure we look after the shoots
