Contents: Volume 186, March 2005

Psychiatry in pictures

Highlights of this issue

EDITORIALS

REVIEW ARTICLES

PAPERS

SHORT REPORTS

COLUMNS

Correspondence

One hundred years ago

Corrigenda

Book reviews

From the Editor's desk

Psychiatry in pictures:

Psychiatry in pictures


Highlights of this issue:

Highlights of this issue


EDITORIALS:

NICE, but will they help people with depression? The new National Institute for Clinical Excellence depression guidelines

Paula Whitty and Simon Gilbody


Current strategies for investigating the genetic and environmental risk factors for affective disorders

Anne Farmer, Thalia C. Eley, and Peter McGuffin

‘Psychiatric comorbidity’: an artefact of current diagnostic systems?
Mario Maj

REVIEW ARTICLES:

Randomised controlled trials relevant to aggressive and violent people, 1955-2000: a survey
Sharon Cure, Wan Lian Chua, Lorna Duggan, and Clive Adams

PAPERS:

Personality and comorbidity of common psychiatric disorders
Amir A. Khan, Kristen C. Jacobson, Charles O. Gardner, Carol A. Prescott, and Kenneth S. Kendler

Reduced hippocampal volumes and memory loss in patients with early- and late-onset depression
Ian Hickie, Sharon Naismith, Philip B. Ward, Keelin Turner, Elizabeth Scott, Philip Mitchell, Kay Wilhelm, and Gordon Parker

Hypofrontality in men with first-episode psychosis

Neural correlates of syntax production in schizophrenia
Tilo T. J. Kircher, Tomasina M. Oh, Michael J. Brammer, and Philip K. McGuire

Nicotine dependence and symptoms in schizophrenia: Naturalistic study of complex interactions
M. Carmen Aguilar, Manuel Gurpegui, Francisco J. Diaz, and Jose De Leon

Escitalopram in the treatment of social anxiety disorder: Randomised, placebo-controlled, flexible-dosage study
Siegfried Kasper, Dan J. Stein, Henrik Loft, and Rico Nil
Assessing adolescent personality pathology
   Drew Westen, Lissa Dutra, and Jonathan Shedler
   Br J Psychiatry 2005 186: 227-238.

Suitability and utility of the CORE-OM and CORE-A for assessing severity of
   presenting problems in psychological therapy services based in primary and secondary
care settings
   Michael Barkham, Naomi Gilbert, Janice Connell, Chris Marshall, and Elspeth Twigg

Psychiatric advance directives: qualitative study of informed deliberations by
   mental health service users
   Michaela Amering, Peter Stastny, and Kim Hopper

Repeated self-poisoning: increasing severity of self-harm as a predictor of
   subsequent suicide
   Greg Carter, David M. Reith, Ian M. Whyte, and Michelle McPherson

**SHORT REPORTS:**

Risk of puerperal and non-puerperal recurrence of illness following bipolar
   affective puerperal (post-partum) psychosis
   Emma Robertson, Ian Jones, Sayeed Haque, Roger Holder, and Nick Craddock

Suicide by hanging: multicentre study based on coroners’ records in England
   Olive Bennewith, David Gunnell, Navneet Kapur, Pauline Turnbull, Sue Simkin,
   Lesley Sutton, and Keith Hawton

**Correspondence:**

Cognitive processing in schizophrenia • Author’s reply
   F. Oyebode, J. Hall, and S. M. Lawrie

Early interventions for psychosis • Author’s reply
   K. H. Marlowe and A. P. Morrison
Integration of psychiatric and physical health  
F. J. Huyse  

Quality of life and ECT • Author’s reply • Author’s reply  
L. Andre, W. V. McCall, and L. Andre  

One hundred years ago:  
Insanity and murder  
Henry Rollin  
Br J Psychiatry 2005 186: 266.

Corrigenda:  
Correction for Volume 185, p. 318  
Br J Psychiatry 2005 186: 266.

Book reviews:  
Autism and Creativity: Is There a Link between Autism in Men and Exceptional Ability?  
Sabina Dosani  

Psychotherapy for Borderline Personality Disorder, Mentalization-based Treatment  
Duncan McLean  

Early Intervention for Trauma and Traumatic Loss  
Ronan J. McIvor  

Practical Psychiatric Epidemiology  
James Kirkbride  

From the Editor's desk:  
Peter Tyrer  

Copyright © 2005 The Royal College of Psychiatrists.
The National Institute for Clinical Excellence (NICE) guidelines on managing depression in primary and secondary care, having been in gestation since 2001, were finally delivered in December 2004. They should be widely welcomed by practitioners and people with depression alike. The guidelines are clear, concise and cover the key aspects of diagnosis and management: screening (high-risk groups only); distinguishing between mild, moderate and severe depression; assessing risk of self-harm and suicide; the benefits of ‘watchful waiting’ for mild depression; when to offer cognitive-behavioural therapy (mild or moderate depression); the use of selective serotonineruptake inhibitors as first-line therapy in most circumstances; and the early use of combined treatment for severe depression. There is also guidance on maintenance treatment, and what to watch out for when discontinuing antidepressants. These are all areas of clinical uncertainty, which have provoked substantial debate in recent years. Given that the majority of people are treated in primary care, the guidelines will be most relevant there.

There is plenty of work to be done to improve care for people with depression. Although depressive disorders are common – in the UK, depression affects between 5% and 10% of individuals and is the third most common reason for consultation in general practice (Singleton et al, 2001) – they may go unrecognised. Eighty per cent of patients with depression can consult with non-specific physical complaints, without spontaneously divulging the psychological nature of their problems (Kirmayer et al, 1993). It has been reported that depressive symptoms are not recognised in UK general practice in about half of attending patients with depressive disorders, ascertained by research diagnostic interview rather than questionnaire (Dowrick & Buchan, 1995). Unrecognised major depression is associated with poor treatment outcomes (Rost et al, 1998), whereas there is evidence that early and vigorous intervention for depression improves outcome (AHCPR Depression Guideline Panel, 2000). Hence the importance of the NICE guidelines, which aim to put this research into day-to-day practice.

However, will the guidelines lead to this improvement in outcomes for people with depression? The short answer is: no, not unless organisational support in primary care for treating the vast majority of people is considerably extended, which includes enhancing the working relationship between primary and secondary care (Gilbody et al, 2002, 2003). Our systematic review of educational and organisational interventions to improve the management of depression in primary care provides the evidence base for how these guidelines should be implemented (Gilbody et al, 2002, 2003). Although the studies reviewed included sophisticated educational interventions (for example, ‘academic detailing’), guideline implementation strategies were successful only when educational interventions were accompanied by an organisational intervention to enhance care, such as nurse case management (Rost et al, 2001) or collaborative and stepped care (Katon et al, 1999; Lin et al, 2000). Collaborative care includes patient and clinician education, along with shared care between primary care physicians, psychiatrists and psychologists, i.e. a substantial enhancement in the working relationship between primary and secondary care. An almost uniform feature of positive studies was the incorporation of some form of case management – usually by primary care nurses or graduate psychologists – to improve the delivery of care. In some studies, nurse involvement was of low intensity, and involved little more than brief medication counselling (Peveler et al, 1999) or psychosocial support over the telephone (‘nurse telehealth care’; Hunkeler et al, 2000). In others, case management was a core ingredient of an effective complex strategy. For example, in the QuEST study (Rost et al, 2001) non-psychiatrically trained practice nurses were given training in the management of depression, and provided a level of ongoing support and monitored therapy, out-patient attendance and treatment response according to well-established algorithms.

The NICE guidelines do include a recommendation on telephone support for monitoring antidepressant regimens (paragraph 1.5.6.1), but not on any of the other organisational interventions that have been identified as being effective. Further, they soften this recommendation with the preambles that ‘In primary care the following strategies can improve the effectiveness of treatments offered’ (our italics; paragraph 1.5.6). A new section of the UK National Health Service workforce has recently come into existence – one thousand graduate primary mental health workers, as promised in the National Service Framework for mental health (Secretary of State for Health, 1999).

There is a clear opportunity and evidence base to guide their role in primary care, by implementing case management and acting as a link between primary and secondary care. Although the full NICE guidelines raise the question of whether these new workers might ‘potentially and significantly affect this situation’ (5.1, p. 82), the summary guideline avoids making specific recommendations for primary care trusts and mental health trusts as to what they should do. There is therefore a danger that they will sink into overstretched services without a clear vision of their role, thus making little impact.

The implication is clear for the NICE depression guidelines: local publicity and educational events, audit of the guidelines and review by the Healthcare Commission will not be enough to help more people with depression get better. To do this, investment in case management and related organisational interventions in primary care are needed, together with an enhancement of the relationship between primary and secondary care services. The new graduate primary care health workers need to be part of this investment. Let us hope that policy makers will grasp this nettle: this is a unique opportunity to improve outcomes for people with depression.

DECLARATION OF INTEREST

None.
REFERENCES


Correspondence: Paula Whitty, Centre for Health Services Research, School of Population and Health Sciences, University of Newcastle, Newcastle upon Tyne; SIMON GILBODY, MB ChB, DPhil, MRCPsych, Academic Unit of Psychiatry, University of Leeds, Leeds, UK

(First received 16 November 2004, accepted 7 December 2004)
Current strategies for investigating the genetic and environmental risk factors for affective disorders*

ANNE FARMER, THALIA C. ELEY and PETER McGUFFIN

It is probable that the genetic components of affective disorders (bipolar affective disorder, major depressive disorder and anxiety states) result from multiple genes that confer a susceptibility or liability to develop the disorder when other (environmental) risk factors are also present. In general, bipolar affective disorder has been found to have the highest heritability (i.e. the proportion of variance explained by additive genetic factors) of around 80% (McGuffin et al, 2003), followed by major depression (between 40% and 70%, depending on the stringency of the definition; Kendler et al, 1993; McGuffin et al, 1996) and then by anxiety disorders, with heritabilities of around 40–50% (Eley et al, 2002).

For affective disorders in adult life, the role as precipitants of certain proximal factors such as severe and threatening life events has been well replicated (Brown & Harris, 1978). There is also much evidence of distal factors such as childhood adversity contributing to vulnerability (Gilman et al, 2003). Important developmental aspects include the continuities between childhood depressive symptoms and adult depression and the changing contributions of genes and environment throughout the life span. For example, recent findings support and extend earlier work that has shown increasing genetic influence on depressive symptoms as children grow into adolescence (Scourfield et al, 2003).

AFFECTIVE DISORDERS: ONE OR MORE SETS OF GENES?

The classification of affective disorders has often been the subject of debate, generating considerable controversy as to whether schemes of subtyping are of any use (Kendell, 1976; Farmer & McGuffin, 1989). One typology that has stood the test of time and seems clinically useful is the unipolar/bipolar subdivision. Until recently, a common view (Gershon et al, 1982) was that bipolar affective disorder and major depressive disorder lie on the same severity of liability continuum contributed to by the (mainly) additive effects of genetic and environmental risk factors. The theory suggests that the two phenotypes differ only in respect of severity of liability, with bipolar affective disorder representing the more severe, less common subtype and major depressive disorder the more common, less severe subtype. However, this model has been recently refuted by McGuffin and colleagues, who also applied structural equation model-fitting methods to further the aetiological overlap between the two disorders, using a twin design (McGuffin et al, 2003). These authors showed that although there is a substantial genetic correlation between mania and depression, most of the genetic variance in liability to mania is specific to the manic syndrome. That is, the main clinically relevant subtypes of affective disorder show a large overlap in their genetic aetiology, but bipolar disorder is also contributed to by a set of genes that are specific to the manic state.

A similar model-fitting approach has been applied to the genetic and environmental overlap between schizophrenia, schizoaffective disorder and bipolar disorder (Cardno et al, 2002). Although there was evidence of genetic overlap between the three disorders there was also evidence for specific genetic components for schizophrenia and bipolar disorder (but not schizoaffective disorder). This goes some way towards explaining the (otherwise puzzling) findings from linkage studies that have implicated some of the same genomic regions in schizophrenia and bipolar disorder and studies that have implicated a positional candidate gene G72 in both disorders (Elkin et al, 2004). Thus, the molecular and the quantitative genetic findings appear to be convergent in suggesting three sets of genes: one conferring liability to both schizophrenia and bipolar disorder, and two that are specific for each of the two main Kraepelinian syndromes. Interestingly, in the quantitative analyses the environmental risk factors appeared to be specific to each psychiatric disorder.

A rather different type of analysis seeks to tease out dimensions within the broad category of recurrent depression. Using phenotypic data from participants in current large-scale genetic studies of depression (see below), this has produced some interesting early results. Factor analysis of psychopathology from worst and second-worst episodes of depression in sibling pairs has shown that a four-factor solution provides the best fit for the data. Three of the factors were found to be familial by examining the sib-pair correlations, and a confirmatory factor analysis of a large sample of unrelated patients suggested that the factor structure is stable, replicable and potentially useful for analyses exploring the relationship with genetic markers (Korszun et al, 2004).

FINDING GENES AND EXPLORING FUNCTION

Linkage studies use genetic marker allele sharing between family members (most commonly, affected siblings) to find regions within the genome where susceptibility genes might be located. Once such regions have been found, they can then be explored in greater detail using association studies (either case–control or within-family designs). Both approaches require large datasets, drawn from hundreds or thousands of subjects, and the application of these methods to affective disorders has somewhat lagged behind other disorders such as schizophrenia. However, several large affected sib-pair and case–control collections of DNA for recurrent major depression and bipolar affective disorder, as well as depressive and anxiety symptoms that occur in the general population, have now been collected and results of genome scans are beginning to emerge (Nash et al, 2004).

Finding genetic polymorphisms associated with affective disorders is only the first step on the path to understanding what
the genes do. Currently the functional effects of common polymorphisms in genes in candidate pathways such as the serotonin pathway are under investigation. This includes studies of gene expression in post-mortem human brain tissue (Sugden et al, 2004) and in animal models (Fernandes et al, 2004).

GENE – ENVIRONMENT INTERPLAY

Although the relationship between adverse life events and the onset of depression is well established, the nature of the relationship between this and genetic vulnerability to depression is not yet understood. Do genes predispose some individuals to encounter adversity or do genes make some individuals more susceptible to the effects of adversity when it occurs? Several studies have shown that life events aggregate in families (e.g. McGuffin et al, 1988; Rijsdijk et al, 2001). In a nearest-aged sib-pair study of depression, Farmer et al (2000) found no significant difference between the number of threatening life events experienced by the siblings of individuals with depression and the siblings of healthy controls. However, the siblings of probands with depression who had experienced a threatening event were significantly more likely to develop depression than the siblings of controls. These findings suggest that genes make individuals susceptible to adversity rather than influence their exposure to it. Nevertheless, many individuals exposed to adversity do not develop depression. Examination of resilience, whatever protects individuals from developing psychosis-pathology in the presence of environmental risk factors, can also be informative (Farmer & McGuffin, 2003). These findings also suggest that interaction effects (not just simple additive effects) between genes and environment are probably more common than previously thought. Caspi et al (2003) demonstrated that a functional polymorphism in the promoter region of the 5-hydroxytryptamine transporter gene (5HTTLPR) moderates the impact of adversity. This finding concerning 5HTTLPR has recently been replicated in a sample of adolescent females (Eley et al, 2004b), and the detection of gene–environment interactions will be the subject of a later editorial.

DEVELOPMENTAL ASPECTS: SAD AND ANXIOUS CHILDREN

One of the methods for disentangling the changes in the relationship between adversity and mood states throughout childhood into adult life is to carry out longitudinal cohort studies. The Twins Early Development Study (TEDS) (Trouton et al, 2002) the largest-ever twin study of its kind, has undertaken regular assessments of twins born in the UK between 1994 and 1995. Several ‘spin-off’ studies have been performed, one of which examined the phenotypic and genetic structure of anxiety in young children (Eley et al, 2003). Another study currently being undertaken will examine mother–twin and twin–twin social interactions during slightly stressful or mildly anxiety-provoking tasks. Emotional responses and measures of anxious cognitive style in both twins as well as the quality of the interactions are being identified. As the study design crosses two time points in middle childhood, a developmental hypotheses regarding aspects of cognitive style and anxiety symptoms can be tested (Eley et al, 2003).

Another ‘spin-off’ study from a large sib-pair study of depression and anxiety in a general population adult sample (The GENESiS study) (Sham et al, 2000) recruited the study participants’ adolescent offspring (around 1900 children, of whom half are sibling pairs), who have been examined at three time points in adolescence. The sib-pair sample was combined with a sample of adolescent twins identified by the Office for National Statistics, and a comprehensive series of assessments have investigated socio-economic factors, education, employment, parenting style and friendships. As this age group is at increasing risk of developing depression, the evolution of affective symptoms and disorder has been evaluated along with cognitive risks associated with these disorders. The study team has shown that regarding risk of depression, there is an interaction between familial vulnerability to adolescent depression and parental lack of education (Eley et al, 2004a). They also found an interaction between negative life events, parental disciplinary style and genetic risk for depression (Lau et al, 2004a). Furthermore, attributional style – a cognitive risk factor for depression traditionally thought of as a learned trait – has been shown to be heritable, and to share genetic influence with both depression and parental disciplinary style (Lau et al, 2004b).

CONCLUSIONS

With the completion of the sequencing of all the base pairs in the human genome earlier in the decade, we are now entering a ‘post-genomic’ era, although identifying the genes involved in the aetiology of affective disorders remains a major research preoccupation. However, many geneticists as well as researchers from other disciplines are now turning their attention to environmental risk factors and how these interact and co-act with genes to lead to the expression of pathological phenotypes such as depression. Although genetic variation in humans can now be determined relatively easily from a single DNA sample derived from blood or even scrapings from the inside of the cheek, experimental manipulation of the environment of human subjects is clearly not possible. Consequently, alternative methods are required to measure the ‘environment’. One is to examine the genotypes of individuals who have all been exposed to a specific risk factor, such as childhood adversity or severe threatening life events, comparing those who have expressed the phenotype, for example by becoming depressed, and those who have not (resilient individuals). Some of the longitudinal and twin studies described above, as well as others currently being conducted by various research groups around the world, will lend themselves to this type of analysis.

DECLARATION OF INTEREST

A.F. and P.M. have received remuneration for presentations and chairing symposia, travel and subsistence costs from Eli Lilly, GlaxoSmithKline and Bristol-Myers Squibb. Both have received consultancy fees from GlaxoSmithKline.

REFERENCES


P. McGuffin, MD, FRCPsych, THALIA C. ELEY, PhD, PETER McGUFFIN, PhD, FRCPsych, Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, London, UK.

Correspondence: Professor Anne Farmer, Social, Genetic and Developmental Psychiatry Centre, PO Box 80, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK. E-mail: A.Farmer@iop.kcl.ac.uk

(First received 29 October 2004, final revision 16 November 2004, accepted 16 November 2004)


The term ‘comorbidity’ was introduced in medicine by Feinstein (1970) to denote those cases in which a ‘distinct additional clinical entity’ occurred during the clinical course of a patient having an index disease. This term has recently become very fashionable in psychiatry to indicate not only those cases in which a patient receives both a psychiatric and a general medical diagnosis (e.g. major depression and hypertension), but also those cases in which a patient receives two or more psychiatric diagnoses (e.g. major depression and panic disorder). This co-occurrence of two or more psychiatric diagnoses (‘psychiatric comorbidity’) has been reported to be very frequent. For instance, in the US National Comorbidity Survey (Kessler et al, 1994), 51% of patients with a DSM–III-R/DSM–IV (American Psychiatric Association, 1987, 1994) diagnosis of major depression had at least one concomitant (‘comorbid’) anxiety disorder and only 26% of them had no concomitant (‘comorbid’) mental disorder, whereas in the Early Developmental Stages of Psychopathology Study (Wittchen et al, 1998) the corresponding figures were 48.6% and 34.8%. In a study based on data from the Australian National Survey of Mental Health and Well-Being (Andrews et al, 2002), 21% of people fulfilling DSM–IV criteria for any mental disorder met the criteria for three or more concomitant (‘comorbid’) disorders.

This use of the term ‘comorbidity’ to indicate the concomitance of two or more psychiatric diagnoses appears incorrect because in most cases it is unclear whether the concomitant diagnoses actually reflect the presence of distinct clinical entities or refer to multiple manifestations of a single clinical entity. Because ‘the use of imprecise language may lead to correspondingly imprecise thinking’ (Lilienfeld et al, 1994), this usage of the term ‘comorbidity’ should probably be avoided.

However, the fact remains that the co-occurrence of multiple psychiatric diagnoses is now more frequent than in the past. This is certainly in part a consequence of the use of standardised diagnostic interviews, which helps to identify several clinical aspects that in the past remained unnoticed after the principal diagnosis had been made—a development that is obviously welcome because it is likely to lead to more comprehensive clinical management and more reliable prediction of future disability and service utilisation. But this is only one part of the story. The other part is that the emergence of the phenomenon of ‘psychiatric comorbidity’ has been to some extent a by-product of some specific features of current diagnostic systems. Artificially splitting a complex clinical condition into several pieces may prevent a holistic approach to the individual, encouraging unwarranted polypharmacy, and may represent a new source of diagnostic unreliability because clinicians may focus their attention on one or other of the different ‘pieces’, especially in those clinical contexts in which coding of only one diagnosis is allowed.

**‘PSYCHIATRIC COMORBIDITY’ AS A BY-PRODUCT OF RECENT DIAGNOSTIC SYSTEMS**

A powerful, usually unrecognised, factor contributing to the emergence of the phenomenon of ‘psychiatric comorbidity’ has been ‘the rule laid down in the construction of DSM–III (American Psychiatric Association, 1980) that the same symptom could not appear in more than one disorder’ (Robins, 1994). This rule (never made explicit, to my knowledge, in DSM-related publications), probably explains why the symptom ‘anxiety’ does not appear in the DSM–IV criteria for major depression, although the text of the manual acknowledges that patients with major depression frequently present with anxiety. Lee Robins, the only author who, as far as I know, has mentioned the above rule in the literature, stated: ‘I thought then, as I still do, that the rule was not a good one’ (Robins, 1994). Actually, DSM–IV does not allow the presence of anxiety in a patient with major depression to be recorded either as a symptom or, as allowed for delusions, a specifier for the diagnosis. The concomitant diagnosis of major depression and panic disorder is encouraged (being one of the most common forms of ‘psychiatric comorbidity’), whereas the concomitant diagnosis of major depression and generalised anxiety disorder is not allowed (unless generalised anxiety occurs also when the patient is not depressed). The latter exclusion criterion seems to be an acknowledgement of the implausibility of the idea that anxiety and depression, when they occur simultaneously, are two separate clinical entities, but it actually contributes to leaving the presence of anxiety in a patient with major depression (with its significant prognostic and therapeutic implications) totally unrecorded. Not surprisingly, both the elimination of the above exclusion criterion (Zimmerman & Chelminski, 2003), which would be consistent with the logic of the system but would multiply the cases of ‘psychiatric comorbidity’, and the introduction of a mixed depressive–anxiety diagnostic category (Tyrer, 2001) have been proposed.

A second, obvious, determinant of the emergence of the phenomenon of ‘psychiatric comorbidity’ has been the proliferation of diagnostic categories in recent classifications. If demarcations are made where they do not exist in nature, the probability that several diagnoses have to be made in an individual case will obviously increase. The current classification of anxiety and personality disorders is a good example of this. It is rare to see a patient with a diagnosis of an anxiety (or a personality) disorder who does not fulfil the criteria for at least one more anxiety (or personality) disorder. The fact that ‘neuroses and abnormal personalities’ do not have clear boundaries either among themselves or with normality was clearly recognised by Jaspers (1913; see below), and would argue in favour of a dimensional approach to their classification. Paradoxically, the attempt by the DSM to

---

1 See pp. 190–196, this issue.
characterise 'pure' disorders in these areas seems to be the first step towards the identification of several 'dimensions'. However, how a dimensional approach would actually work in clinical practice (e.g. in what cases a disorder would finally be diagnosed, and how the diagnosis would be expressed) remains unclear.

A third relevant characteristic of current diagnostic systems is the limited number of hierarchical rules. A consolidated tradition in psychiatry was to establish a hierarchy of diagnostic categories so that, for example, if a psychotic disorder were present, the possibly concomitant neurotic disorders would not be diagnosed because they would be regarded as part of the clinical picture of the psychotic condition. One could argue that the current possibility of diagnosing a panic disorder in the presence of a diagnosis of schizophrenia represents a useful development, because this additional diagnosis provides information that may be useful for clinical management. But are we sure that the occurrence of panic attacks in a person with schizophrenia should be conceptualised as the 'comorbidity' of panic disorder and schizophrenia? Is the panic of a person with agoraphobia, of a person with major depression and of a person with schizophrenia the same psychopathological entity that simply 'co-occurs' with the other three? I am not aware of any research evidence on this issue.

A fourth relevant feature of our current diagnostic systems is the fact that they are based on operational diagnostic criteria. Because of this, they are regarded as more precise and reliable than the traditional ones based on clinical descriptions. However, the old clinical descriptions provided a gestalt of each diagnostic entity, which is often not provided by current operational definitions. This was probably due in part to the different emphasis laid on the various clinical aspects (whereas in current operational definitions the various clinical features are usually given the same weight), as well as to the inclusion of some aspects regarded as essential (e.g. autism in the case of schizophrenia) that do not appear in current diagnostic systems because they are regarded as not sufficiently reliable. Traditional clinical descriptions encouraged differential diagnosis, whereas current operational definitions encourage multiple diagnoses, probably in part because they are less able to convey the 'essence' of each diagnostic entity. Is this an intrinsic limitation of any operational definition, or a remediable flaw of our current operational definitions? Was the above-mentioned gestalt (for instance, in the case of schizophrenia) a fact or an illusion? Are we sure that we have used all the resources of the operational approach in typifying, for instance, the disorder of social and interpersonal functioning in schizophrenia?

**PSYCHIATRIC COMORBIDITY AND THE NATURE OF PSYCHOPATHOLOGY**

Most of the recent debate about psychiatric comorbidity has been remarkably atheoretical, focusing on the practical usefulness of one or the other approach in terms of treatment selection and prediction of outcome and service utilisation. However, the emergence of the phenomenon of 'psychiatric comorbidity' has obvious theoretical implications. The frequent co-occurrence of the mental disorders included in current diagnostic systems has recently been regarded as evidence against the idea that these disorders represent discrete disease entities (e.g. Cloninger, 2002). The point has been made that the nature of psychopathology is intrinsically composite and changeable, and that what is currently conceptualised as the co-occurrence of multiple disorders could be better reformulated as the complexity of many psychiatric conditions (with increasing complexity being an obvious predictor of greater severity, disability and service utilisation). From the psychodynamic viewpoint, the idea seems to be reinforced that the interaction of congenital predisposition, individual experiences and the type and success of defence mechanisms employed may generate an infinite variety of combinations of symptoms and signs. From the psychobiological viewpoint, the hypothesis seems to be supported that 'noxious stimuli…perturb a variety of neuronal circuits…The extent to which the various neuronal circuits will be involved varies individually, and consequently psychiatric conditions will lack symptomatic consistency and predictability' (van Praag, 1996). From the evolutionary viewpoint, the concept seems to be corroborated that mental disorders are the expression of preformed response patterns shared by all humans, which may be activated simultaneously or successively in the same individual by noxae of various nature – a view endorsed by Kraepelin himself in one of his later works, in which he dismissed the model of discrete disease entities even for dementia praecox and manic–depressive insanity (Kraepelin, 1920).

However, the emergence of the phenomenon of 'psychiatric comorbidity' does not necessarily contradict the idea that psychopathology consists of discrete disease entities. An alternative possibility is that psychopathology does consist of discrete entities, but these entities are not appropriately reflected by current diagnostic categories. If this is the case, then current clinical research on 'psychiatric comorbidity' may be helpful in the search for 'true' disease entities, contributing in the long term to a rearrangement of present classifications, which may involve a simplification (i.e. a single disease entity may underlie the apparent 'comorbidity' of several disorders), a further complication (i.e. different disease entities may correspond to different 'comorbidity' patterns) or possibly a simplification in some areas of classification and a further complication in other areas.

There is, however, a third possibility: that the nature of psychopathology is intrinsically heterogeneous, consisting partly of true disease entities and partly of maladaptive response patterns. This is what Jaspers (1913) actually suggested when he distinguished between 'true diseases' (such as general paresis), which have clear boundaries among themselves and with normality; 'circles' (such as manic–depressive insanity and schizophrenia), which have clear boundaries with normality but not among themselves; and 'types' (such as neuroses and abnormal personalities), which do not have clear boundaries either among themselves or with normality. Recently, it has been pointed out (Nesse, 2000) that throughout medicine there are diseases arising from a defect in the body’s machinery and diseases...
arising from a dysregulation of defences. If this is true also for mental disorders – for example, if a condition such as bipolar disorder is a disease arising from a defect in the brain machinery, whereas conditions such as anxiety disorders, or part of them, arise from a dysregulation of defences – then different classification strategies may be needed for the various areas of psychopathology.

DECLARATION OF INTEREST
None.

REFERENCES


Randomised controlled trials relevant to aggressive and violent people, 1955–2000: a survey

SHARON CURE, WAN LIAN CHUA, LORNA DUGGAN and CLIVE ADAMS

Background Randomised trials remain the gold standard for evaluating health interventions. This applies to the criminal justice system as well as to health.

Aims To identify and survey randomised trials relevant to forensic mental health services.

Method We searched 29 electronic bibliographic databases and acquired randomised trials involving sex offenders, arsonists or people clearly and actively aggressive, or abusive of children or spouse. Two researchers reliably extracted data.

Results Of 409 studies found, we were able to acquire 300 for further inspection. They all involved particularly violent people (total n=28 669), mostly adult men; the mean study size was 197 (median 52, mode 60, range 1–1200). In these 300 randomised trials over 700 interventions were evaluated and short-term outcomes were recorded on 345 different scales.

Conclusions Wider collaboration, rationalising treatments and simplifying outcomes could further strengthen the tradition of trialling in forensic psychiatry. Systematic reviews of these studies are overdue.

Declaration of interest None.

The management of aggression and of potentially aggressive people forms a large part of the workload of forensic mental health services (Taylor & Gunn, 1999). This work is a priority at the highest political levels and society is becoming increasingly intolerant of aggression perpetrated by those with mental health difficulties. In the UK the government has acted to introduce new legislation (Department of Health, 2001). In this context of increasing public concern it is imperative that public policy is informed by the entirety of high-quality research rather than by a proportion.

Although often imperfect (Chalmers et al, 1983; Thornley & Adams, 1998), randomised controlled trials remain the gold standard for the evaluation of mental health interventions (World Health Organization Scientific Group on Treatment of Psychiatric Disorders, 1991). This applies equally to research into the criminal justice system (Farrington & Petrosino, 2001).

There are strong arguments for collecting and disseminating a regularly updated register of all randomised trials relevant to this area of work (Davies & Boruch, 2001). In mainstream healthcare the need of both providers and those receiving interventions to have ready access to all relevant high-quality research has been recognised, and the Cochrane Collaboration provides a structure by which this is undertaken.

More recently, those working in education, social welfare and the criminal justice system have formed the Campbell Collaboration to address the needs of – among others – forensic mental health services (Farrington & Petrosino, 2001). However, forensic mental health straddles many professions and this fragmentation makes it difficult for healthcare professionals, criminal justice system workers, consumers, researchers and policy-makers to access relevant information. Anticipating this, Petrosino compiled a database of social, psychological, educational and criminological randomised and possibly randomised studies (Petrosino et al, 2000). Our work benefits from, supersedes and expands Petrosino’s initiative. We created and surveyed a register of randomised controlled trials relevant to the management of violent and aggressive people.

METHOD

We searched 29 accessible electronic bibliographic databases (see Table 1) thought to be of relevance to the area. None of the relevant databases that we knew of was inaccessible. Published strategies for identifying randomised control trials were adapted as necessary. Participant-specific searches were then constructed (further details available from the author upon request). These broad electronic searches identified approximately 22 000 unique reports. One author (S.C.) inspected each electronic report and discarded irrelevant material; she then noted the participant group. Another author (C.A.) selected and recoded a random 10% sample. A total of 2184 reports of possibly randomised studies relevant to aggressive or potentially aggressive people remained.

A priori, we defined a subgroup of these studies as being of higher priority to forensic mental health services. These involved people who were clearly and actively aggressive, people abusive of children or spouse, sex offenders and arsonists, irrespective of age and whether they had underlying disorders. Studies of people at risk of becoming aggressive, for example juvenile offenders with no record of a specified aggressive act, were not included in this higher-priority group. Full copies of these high-priority studies were obtained and, using a data extraction sheet, S.C. recorded information on participants’ diagnoses, problematic behaviour, stage in criminal justice system, interventions and outcomes; C.A. checked the reliability of the coding by recoding a 10% random sample again. Methodological quality was scored according to the Jadad scale (Jadad et al, 1996).

This rates the quality of reporting of randomisation (0–2), the quality of reporting of masking (0–2) and the quality of reporting of withdrawals (0–1). Low scores indicate poor reporting of methods and are linked with estimates of effect substantially greater than when a study is rated as good on the Jadad scale (Moher et al, 1998). This overestimate of effect from studies in which methodology is poorly reported is in
Keeping with other studies using different parameters to measure study quality (Juni et al., 2001). Data were stored in ProCite (Adept Scientific, Letchworth, UK) and then exported to Epi Info version 6.04d (Centers for Disease Control, Atlanta, Georgia, USA) for analysis.

**RESULTS**

None of the 29 databases we searched stood out as a definitive source of forensic studies (Table 1). We identified 2184 electronic reports of trials of aggressive and potentially aggressive people. These were included in 481 different journals, books or dissertations (all dissertations counted as one source). Many of the reports identified but not included in our detailed survey will nevertheless be of interest to the forensic mental health services; these lower-priority studies focused on possibly...
or potentially aggressive or violent people and involved groups such as juvenile offenders or prisoners for whom the level of aggression or violence was not explicit (Table 2).

Because of time constraints and despite our best efforts, we were only able to acquire and survey 300 of the 409 studies that we had identified as being of higher priority. There was an approximately 30% false-positive rate, so we estimate that about 70 studies remain outstanding. These proved inaccessible even through the British Library and direct approaches to the relevant people or institutions.

The reliability of most coding was good, with 90–100% agreement for type of publication, country of origin, year of publication, language, participants’ gender, age and previous offences, intervention, number finishing trial, duration of trial, description of randomisation, description of masking and description of withdrawal. Agreement was between 50% and 90% for number randomised, problematic behaviour and diagnosis. Outcomes were not rated reliably (10% full agreement), probably because data were difficult to identify and involved many variables. Each rater found additional outcomes. The proportion of papers for which raters agreed on most (>70%) outcomes was 95%, but the numbers of scales listed below is likely to be an underestimate.

### Detailed survey of high-priority reports

The final column of Table 1 shows the proportion of unique high-priority studies identified in each database as it was searched in turn. For example, after SPECTR (Social, Psychological, Educational and Criminological Trials Register) was searched, a Medline search still found 19% of the 300 studies. After 14 other databases had been searched the Cochrane Library still found 11% of the total, and Dissertation Abstracts, despite being 18th to be searched, also found 11% of the total. Most of the 300 reports we were able to acquire were fully published papers in academic journals (105 different journals), but no core set of journals deserves a reputation for having a special interest in this area, and 20% of reports were found only in dissertations or conference proceedings.

Three-quarters (76%) of randomised controlled trials relevant to the management of very aggressive people originate from the USA. Of the remaining studies, 7% were from the UK, 4% from Europe and 12% from rest of the world (1% not specified). From 1995 there has been a steady increase in the number of relevant studies (1 per month 1991–2000).

A total of 28,669 people had been randomised within the 300 trials (mean sample size 197, median 52, mode 60, range 1–1200), and 280 studies clearly reported both the numbers starting and finishing the trial: the average attrition rate was 19% (95% CI 15–27%). The great majority of reports involved men; only 15 trials (5%) solely randomised women. Most studies dealt with aggression in adulthood, although one-third focused on adolescents.

It was often difficult to ascertain diagnoses from reports, and when they were specified, often several were described in a single report. Specified diagnoses were categorised and frequencies tallied: psychotic disorders were the most commonly reported (178; 59%), followed by personality disorder (85; 18%), affective disorder (34; 11%), substance misuse (31; 10%), sexual disorders (30%; 10%), behaviour disorders (30; 10%), neurotic problems (26; 9%), problems of organic origin (21; 7%),
Table 3  Top ten problematic behaviours stipulated in the trials

<table>
<thead>
<tr>
<th>Specific problem</th>
<th>Number of different reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aggression</td>
<td></td>
</tr>
<tr>
<td>Assault</td>
<td>37</td>
</tr>
<tr>
<td>Destruction of property</td>
<td>18</td>
</tr>
<tr>
<td>Hostility</td>
<td>18</td>
</tr>
<tr>
<td>Murder</td>
<td>19</td>
</tr>
<tr>
<td>Non-sexual child abuse</td>
<td>17</td>
</tr>
<tr>
<td>Sexual child abuse (high-risk groups)</td>
<td>11</td>
</tr>
<tr>
<td>Exhibitionism</td>
<td>18</td>
</tr>
<tr>
<td>Paedophilia</td>
<td>22</td>
</tr>
<tr>
<td>Rape</td>
<td>36</td>
</tr>
<tr>
<td>Unspecified</td>
<td>11</td>
</tr>
<tr>
<td>Spouse abuse</td>
<td>22</td>
</tr>
<tr>
<td>Threatens to harm others</td>
<td>15</td>
</tr>
<tr>
<td>Unspecified</td>
<td>153</td>
</tr>
<tr>
<td>Behaviour</td>
<td></td>
</tr>
<tr>
<td>Agitation</td>
<td>31</td>
</tr>
<tr>
<td>Disruptiveness</td>
<td>11</td>
</tr>
<tr>
<td>Impulsivity</td>
<td>20</td>
</tr>
</tbody>
</table>

Despite the considerable limitations of even the best electronic search (Adams et al, 1994) and the inaccessibility of 25% of the high-priority sample, this survey suggests that there may be hundreds and even thousands of randomised studies directly relevant to the forensic mental health services. These trials are published in a broad range of journals, and many do not seem ever to be published except as a dissertation of a doctoral student or a presentation at a conference. Although one relevant study from the high-priority group is published per month, it is impossible to predict where that report will appear. These multiple sources are indexed in many databases. Enormous effort went into identification of these studies, and almost every database searched yielded reports of previously undiscovered trials. This underlines the need for registration of trials at inception and for a central repository of such trials (Dickersin, 1988; Hetherington et al, 1989; Stern & Simes, 1997).

The 300 studies surveyed in detail are likely to be a biased sample. Reports in English are easier to find than similar work in other languages (Niemenen & Isohanni, 1999). Work with statistically significant results tends to be more accessible than trials with equivocal findings (Egger et al, 1997). It seems unlikely, however, that a significant body of higher-quality, larger studies has gone unnoticed. Reliability of coding of the variables used in this report is high, so results should reflect the subpopulation of studies surveyed.

The overall quality of reporting was mediocre. This is also the case in other branches of psychiatry (Thornley & Adams, 1998) and medicine (Gotzsche, 1989; Vanderkerckhove et al, 1993; Fahey et al, 1995; Schulz et al, 1995a; Cheng et al, 2000). This poor quality of reporting is likely to be associated with exaggerated estimates of effect (Schulz et al, 1995b). It is hoped that with CONSORT (Moher et al, 2001), the quality of trial reporting should improve.

People in the trials prioritised for this study commonly had psychosis or personality disorder and exhibited extremely aggressive behaviour. The range of interventions that have been trialled is bewildering, but few studies focus on similar interventions for similar participants. Pioneers have undertaken these important and often ground-breaking studies, but there is little evidence of collaboration between individuals or institutions to rationalise interventions and increase the power of their evaluative studies. Most studies are grossly underpowered for clinically relevant outcomes. Without widespread collaboration this is likely to remain the case.

One in three schizophrenia trials contain a new outcome rating scale (Thornley & Adams, 1998). More than a third of these scales are not validated and produce biased estimates of effect (Marshall et al, 2000). The 300 high-priority studies in this survey contain 1.2 new scales per report. The proportion not validated is likely to be high. Considering the limited clinical usefulness of much scale-derived data, this seems a remarkable waste of resources in a sub-speciality in which concrete and relevant outcomes may be more plentiful than in general psychiatry.

All trials identified by the project were made available within the Cochrane Controlled Trials Register and also offered to the Campbell Collaboration to build on their SPECTR database of trials. It is hoped that this database will allow people in a range of disciplines to have ready access to trial-based information relevant to offenders and potential offenders, and to learn from past practice in order to inform future work.
This broad overview suggests that wider collaboration, rationalising treatments and simplifying outcomes could further strengthen the tradition of trialling in forensic psychiatry. Systematic reviews of these studies are overdue.

ACKNOWLEDGEMENT

This work would not have been possible without the support of the National Health Service Research and Development Programme for Forensic Mental Health (grant HO01PN3X). We are grateful for their vision, help and patience.

REFERENCES


CLINICAL IMPLICATIONS

■ Often randomised trials involving participants and interventions of interest to the forensic services do exist.

■ These studies have been difficult to find but are now available within the Cochrane Controlled Trials Register and have been offered to the Campbell Collaboration to add to their SPECTR (Social, Psychological, Educational and Criminological Trials Register) database.

■ Collaborative work is needed to evaluate practices common in forensic mental health services.

LIMITATIONS

■ The sample of studies included are the most accessible of those identified.

■ Additional studies are likely to exist in different databases or journals, or as unpublished manuscripts.

■ In the period between undertaking this research and publication of the present report many other relevant studies may have been performed.

SHARON CURE, MA, MCLIP, GenCert Hort, Sainsbury Library, Said Business School, Oxford; WANDUAN CHUA, BMedSci, BMBS, MRCPsych, Airedale General Hospital, Keighley; LORNA DUGGAN, MBChB, MRCPsych, St. Andrew’s Hospital, Northampton; CLIVE ADAMS, MBChB, BAQ, MRCPsych, MSci(Epid), MD, Academic Unit of Psychiatry and Behavioural Sciences, University of Leeds, Leeds, UK

Correspondence: Professor Clive Adams, Academic Unit of Psychiatry and Behavioural Sciences, University of Leeds, 15 Hyde Terrace, Leeds LS2 9LT, UK. Tel: +44 (0)113 343 2730; fax: +44 (0)113 343 2723; e-mail: ceadams@cochrane-sz.org

(First received 13 November 2003, accepted 26 October 2004)
Personality and comorbidity of common psychiatric disorders

AMIR A. KHAN, KRISTEN C. JACOBSON, CHARLES O. GARDNER, CAROL A. PRESCOTT and KENNETH S. KENDLER

Background  We know little about the degree to which comorbidity, so commonly seen among psychiatric disorders, arises from variation in normal personality.

Aims  To study the degree to which variation in normal personality accounts for the comorbidity of eight common psychiatric and substance use disorders.

Method  Internalising disorders (major depression, generalised anxiety and panic disorders, phobias), externalising disorders (alcohol and drug dependence, antisocial personality and conduct disorders) and personality dimensions of neuroticism, extraversion and novelty seeking were assessed in 7588 participants from a population-based twin registry The proportion of comorbidity explained by each personality dimension was calculated using structural equation modelling.

Results  Neuroticism accounted for the highest proportion of comorbidity within internalising disorders (20–45%) and between internalising and externalising disorders (19–88%). Variation in neuroticism and novelty seeking each accounted for a modest proportion (10–12% and 7–14%, respectively) of the comorbidity within externalising disorders. Extraversion contributed negligibly.

Conclusions  High neuroticism appears to be a broad vulnerability factor for comorbid psychiatric disorders. Novelty seeking is modestly important for comorbid externalising disorders.

Declaration of interest  None.

Funding detailed in Acknowledgements.

High comorbidity among psychiatric disorders is consistently reported (Kessler et al, 1994; Merikangas et al, 1996). Among many proposed explanations, one possibility is that personality mediates part of this comorbidity (Jardine et al, 1984; Clark et al, 1994; Battaglia et al, 1996; Bienvenu et al, 2001; Krueger & Markon, 2001). This study examines the association of variation in personality traits of neuroticism, extraversion and novelty seeking and the comorbidity among eight disorders: major depression, generalised anxiety disorder (GAD), panic disorder, any phobia, alcohol dependence, drug dependence, antisocial personality disorder and conduct disorder. This study not only attempts to replicate previous work using a large epidemiological sample, including more comprehensive diagnostic categories and different statistical methodology, but also attempts to quantify the proportion of comorbidity among psychiatric disorders explained by individual personality dimensions.

METHOD

Participants

Our sample derives from two related projects utilising the population-based Virginia Twin Registry, which was formed from a systematic review of all birth certificates in the Commonwealth of Virginia and now constitutes part of the Mid-Atlantic Twin Registry. The female–female (FF) twin pairs used in this study come from birth years 1934–1974. Twin pairs became eligible to participate if both members had responded previously to a mailed questionnaire, the response rate to which was 64%. Eighty-eight per cent of our sample were first interviewed face to face in 1987–1989 (wave 1) and subsequently have participated in up to three additional telephone interviews (waves 2–4).

The male–male and male–female (MM/MMF) twin pairs, covering the birth years 1949–1974, were ascertained in a separate study beginning in 1993. We interviewed 72% of the eligible sample, usually by telephone, in our wave 1 study. This sample was followed up in a second wave of face-to-face interviews (1994–1998) that were completed with 79.4% of eligible participants.

We examine here the results of combined data from the MM/MMF and FF samples, based on the second and fourth wave of interviews, respectively, because these were the most recent waves in which we had measured both personality and psychiatric diagnoses. Our sample consisted of 7388 individual twins, with 4240 males (55.9%) and 3348 females (44.1%). All participants were Caucasian, ranging in age from 20 to 38 years (mean = 36.8, s.d. = 8.9) at the time of the interview. Informed consent was obtained from all participants prior to assessment.

Measures

Psychiatric disorders

The outcome measures of interest, as outlined in the introduction, were lifetime diagnoses of common psychiatric disorders. In order to facilitate the discussion, we will use the concepts of internalising (propensity to express distress inwards, including major depression, GAD, panic disorder, any phobia) and externalising (propensity to express distress outwards, including alcohol and drug dependence, antisocial personality disorder, conduct disorder) disorders as described by Krueger et al (Krueger, 1999; Krueger & Markon, 2001). With the exception of ‘any phobia’, all disorders were assessed using the Structured Clinical Interview for DSM–III–R (Spitzer & Williams, 1985). Diagnostic algorithms for GAD, panic disorder and alcohol dependence were modified to reflect DSM–IV criteria (American Psychiatric Association, 1994), whereas major depression, drug dependence, antisocial personality disorder and conduct disorder were based on DSM–III–R criteria (American Psychiatric Association, 1987) owing to the lack of items corresponding to DSM–IV criteria. The drug dependence diagnosis included dependence on marijuana, cocaine, opiates, hallucinogens, stimulants, sedatives or other drugs. Phobias were assessed with an adaptation of the phobic disorders section of the
Diagnostic Interview Schedule, version III-A (Robins & Helzer, 1985), and the diagnosis of 'any phobia' included agoraphobia, social, situational, animal, blood and miscellaneous phobias. The diagnostic algorithm for phobias has been described in detail previously (Kendler et al., 2002).

Interviewers were carefully trained and supervised, and had at least a master's degree in a mental health-related field or a bachelor's degree in such a field and two years of clinical experience. Diagnoses for conduct disorder and antisocial personality disorder were based on self-report questionnaires; all other diagnoses were assessed using personal interview. Interrater reliability for diagnosis (based on a subsample of FF twins) was high (e.g. for major depression, mean (s.d.), κ=0.96 (0.04)), and test–retest reliability (based on an average interval of 4.5 weeks, range 2–8 weeks, between base and reliability interview) was also acceptable for most diagnoses (range=0.23–0.74, average κ=0.52). Finally, the comorbidity of antisocial personality disorder and conduct disorder was not examined because the diagnosis of antisocial personality disorder requires the onset of conduct disorder before age 15 years. Table I describes the prevalence of psychiatric disorders in our sample.

### Personality

Neuroticism and extraversion, as conceptualised by Eysenck (Eysenck & Eysenck, 1975; Hirschfeld et al., 1983), have been identified cross-culturally as major personality traits by nearly all subsequent investigators (Pervin, 1990). Neuroticism reflects emotional instability, vulnerability to stress and anxiety proneness, whereas extraversion measures sociability and liveliness. Novelty seeking, another personality dimension, measures exploratory excitability, impulsiveness, extravagance and regimentation (Cloninger et al., 1991). Personality measures of neuroticism and extraversion were obtained by self-report questionnaire in the MM/MF sample and were part of the telephone interview in the FF sample. Novelty seeking was assessed by self-report questionnaire only, in both samples. Neuroticism and extraversion were assessed with 12 and 8 items, respectively, from the shortened version of the Eysenck Personality Questionnaire – Revised (EPQ–R; Eysenck et al., 1985; Heath et al., 1992). Novelty seeking was evaluated by 18 items from the abbreviated 54-item version of the Tridimensional Personality Questionnaire (TPQ) of Cloninger (Cloninger et al., 1991; Heath et al., 1994). For statistical analyses we used composite personality measures derived from individual items for each dimension, respectively.

### Missing data

Valid data on all three personality measures and all eight psychiatric disorders were available for the vast majority (85.6%; n=6499) of the sample. Missing data for major depression, GAD, any phobia and alcohol and drug dependence were minimal (<0.6%). Rates of missing data for conduct disorder and antisocial personality disorder were somewhat higher (approximately 1–3%, 7–16%) because these diagnoses were assessed using a separate self-report questionnaire. Rates of missing data for the three personality measures were 2–16%, also due primarily to lower response rates for the self-report questionnaire. Preliminary analyses revealed no significant differences in mean levels of personality or psychiatric diagnosis due to missing data on other variables (results available from the authors upon request).

### Statistical analysis

We performed logistic regression analyses to estimate the association of each personality dimension with each psychiatric disorder. Correction for the correlated structure of our twin data was done using generalised estimating equations (Li & Zeger, 1986) as implemented in the Statistical Analysis System (SAS) procedure GENMOD. Multiple logistic regression analyses were performed with all three personality measures as independent variables. Age, zygosity and gender were used as covariates. Scores for all personality measures were standardised to a mean of 0 and a variance of 1 to facilitate the direct comparison of their effects on the disorder of interest. Odds ratios with 95% confidence intervals and their statistical significance are reported. An odds ratio of >1 represents the increase in risk of disorder associated with each standard deviation (s.d.) increase in the score of the personality dimension. An odds ratio of <1 represents the decrease in risk associated with each s.d. increase in personality dimension score.

In order to calculate the proportion of comorbidity attributed to variation in normal personality, we conducted structural equation modelling analyses using the software program Mx (Neale et al., 1999). As depicted in Fig. 1, the model we used allowed us to calculate the total covariance (i.e. comorbidity) between the disorders of interest. This covariance was broken down into the covariance attributed to personality and the residual covariance, which represents any remaining

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Males (n/sample N)</th>
<th>Females (n/sample N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major depression</td>
<td>29.5% (1252/4240)</td>
<td>41.0% (1374/3348)</td>
</tr>
<tr>
<td>Generalised anxiety disorder</td>
<td>1.7% (71/4226)</td>
<td>2.8% (95/3340)</td>
</tr>
<tr>
<td>Panic disorder</td>
<td>1.1% (48/4214)</td>
<td>3.2% (108/3393)</td>
</tr>
<tr>
<td>Any phobia</td>
<td>22.0% (929/4215)</td>
<td>31.7% (1055/3329)</td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>25.9% (1092/4213)</td>
<td>9.8% (326/3332)</td>
</tr>
<tr>
<td>Any drug dependence</td>
<td>7.3% (309/4240)</td>
<td>4.3% (144/3325)</td>
</tr>
<tr>
<td>Antisocial personality disorder</td>
<td>3.9% (154/3947)</td>
<td>0.4% (12/2776)</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>10.1% (402/3962)</td>
<td>1.7% (48/2781)</td>
</tr>
</tbody>
</table>

1. Total sample size varies because of missing data.

Table I: Prevalence of psychiatric disorders by gender.
comorbidity was modelled across personality measures by allowing thresholds for neuroticism (N), extraversion (E) and novelty seeking (NS) to vary by gender. Because the overall correlation across personality measures was low to moderate (between neuroticism and extraversion = −0.19, neuroticism and novelty seeking = 0.04, extraversion and novelty seeking = 0.34) indirect effects of personality are ignored in this analysis. Covariance due to comorbidity attributable to personality was constrained to be equal for men and women and evaluated the overall fit of the model (using Akaike’s information criteria, AIC) compared with the model where thresholds were allowed to vary by gender. Models with the lowest AIC values were considered to be the best-fitting models. We also obtained gender differences in the overall pattern of covariance by constraining the parameter estimates to be the same in males and females, and comparing the pattern of covariance with a model where parameters were allowed to vary by gender. Because Mx currently lacks the capability to analyse continuous and ordinal traits simultaneously, the continuously measured personality traits were divided into categories based on the maximum number of responses possible, and thresholds corresponding to the proportions of individuals in each category were estimated. For example, scores on the neuroticism variable were in the range 0–12. Thus, we used 12 thresholds to estimate the proportion of individuals within each response category.

RESULTS

Logistic regression for the effects of personality on psychiatric disorders

Table 2 shows the odds ratios from the logistic regression analyses for each of the three personality measures. Higher scores on neuroticism significantly increased the risk for all the disorders examined. For each s.d. increase in neuroticism, the highest (130%) risk increase was for GAD and the lowest (26%) for conduct disorder. Extraversion’s impact was modest overall, with no consistent pattern across internalising and externalising disorders. Specifically, one s.d. increase in extraversion was associated with a 24% increased risk for drug dependence, with a smaller increase for GAD, alcohol dependence and major depression. Novelty seeking was most strongly associated with externalising disorders (alcohol and drug dependence, antisocial personality disorder, conduct disorder), with increase in risk ranging from 37% to 83%. Inspection of covariates revealed that internalising disorders (major depression, GAD, panic disorder and any phobia) were more prevalent in females whereas externalising disorders (alcohol and drug dependence, antisocial personality disorder, conduct disorder) were more prevalent in males (Table 1). Age was positively associated with internalising disorders (i.e. older subjects reported a higher prevalence of major depression, GAD, panic disorder and any phobia) and was negatively associated with the externalising disorders (i.e. younger subjects had higher rates of alcohol and drug dependence, antisocial personality disorder and conduct disorder). Zygosity was not associated with any of the psychiatric disorders.

We also tested for interactions between gender and each of our three personality measures for each of the disorders. Out of 24 possible interactions (8 disorders × 3 interactions), only the interaction between gender and neuroticism for alcohol dependence was significant (β = 0.06, s.e. = 0.02, Wald χ² = 5.22, P < 0.05). In this case, the relationship between neuroticism and alcohol dependence was stronger for females than for males. However, it should be noted that this significant interaction may be a stochastic effect. Thus, for the structural equation modelling analyses of personality and comorbidity, males and females were combined into a single sample, although thresholds corresponding to psychiatric disorder were estimated separately for males and females.

Structural equation modelling of personality effects on comorbidity

For ease of interpretation, the results of the structural equation modelling analyses are
The overall pattern of results, as shown in Fig. 2 and Table 3, indicates that neuroticism accounts for the highest proportion of comorbidity within internalising disorders (20–45%, arithmetic average=31%) and between internalising and externalising disorders (19–88%, arithmetic average=36.8%). Neuroticism also explained 10–12% of the comorbidity within externalising disorders. Extraversion explained only a very small proportion of the comorbidity (−4.9 to 7.4%). Novelty seeking accounted for a negligible proportion of comorbidity within internalising disorders (−0.8 to 0.7%) and between internalising and externalising disorders (−13.2% to 5.8%); however, novelty seeking did account for 7.4–14% of the comorbidity within externalising disorders. Residual covariance (i.e. due to factors other than personality) accounted for most of the comorbidity, with an arithmetic average of 65%. Negative values in Fig. 2 and Table 3 reflect the effects of low extraversion (introversion) and low novelty seeking on comorbidity, although the majority of these effects are quite small.

Although the models where thresholds for psychiatric disorders were allowed to vary by gender consistently fit the data better than models assuming equal thresholds, there were no significant gender differences in the covariance structure (results available from the authors upon request). Thus, the pattern of comorbidity accounted for by personality was similar in males and females, despite the significant differences in the rates of psychiatric disorders.

**DISCUSSION**

**Neuroticism**

Our results suggest that normal personality dimensions of neuroticism not only contributed to individual diagnoses but also accounted for a significant part of the lifetime comorbidity of common psychiatric disorders. The most striking finding was that neuroticism, on average, accounted for 26% of the comorbidity among the disorders included in the study (range=12–88%). This finding is consistent with previous research (Clark et al, 1994; Sher & Trull, 1994, Krueger & Markon, 2001; Bienvenu et al, 2001) and suggests neuroticism as a potential general underlying vulnerability factor for psychopathology.

**Extraversion**

Although extraversion was significantly, albeit weakly, associated with four of the eight psychiatric disorders in the logistic regressions, it explained very small proportions of comorbidity. This pattern of weak
Table 2 Association of personality measures and common psychiatric disorders

<table>
<thead>
<tr>
<th>Personality dimension</th>
<th>Major depression</th>
<th>Generalised anxiety disorder</th>
<th>Panic disorder</th>
<th>Any phobia</th>
<th>Alcohol dependence</th>
<th>Any drug dependence</th>
<th>Antisocial personality disorder</th>
<th>Conduct disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N (%)</td>
<td>E (%)</td>
<td>N (%)</td>
<td>E (%)</td>
<td>N (%)</td>
<td>E (%)</td>
<td>N (%)</td>
<td>E (%)</td>
</tr>
<tr>
<td>Neuroticism</td>
<td>1.95***</td>
<td>2.30***</td>
<td>1.93***</td>
<td>1.62***</td>
<td>1.57***</td>
<td>1.64***</td>
<td>1.44***</td>
<td>1.26***</td>
</tr>
<tr>
<td>Extraversion</td>
<td>1.07*</td>
<td>1.21*</td>
<td>0.95</td>
<td>1.18***</td>
<td>1.24**</td>
<td>0.92</td>
<td>0.93</td>
<td>0.83–1.04</td>
</tr>
<tr>
<td>Novelty seeking</td>
<td>1.14***</td>
<td>0.92</td>
<td>1.10</td>
<td>1.37***</td>
<td>1.58***</td>
<td>1.83***</td>
<td>1.39***</td>
<td>1.26–1.54</td>
</tr>
<tr>
<td>Any phobia</td>
<td>0.30</td>
<td>0.77–1.10</td>
<td>0.88–1.36</td>
<td>1.03–1.16</td>
<td>1.28–1.46</td>
<td>1.42–1.76</td>
<td>1.57–2.13</td>
<td>1.26–1.54</td>
</tr>
<tr>
<td>Alcohol dependence</td>
<td>0.34</td>
<td>0.74–1.62</td>
<td>0.72</td>
<td>0.25–3.35</td>
<td>0.35–4.9</td>
<td>0.25–7.6</td>
<td>0.25–6.6</td>
<td>0.25–5.6</td>
</tr>
<tr>
<td>Any drug dependence</td>
<td>0.40</td>
<td>0.85–2.72</td>
<td>0.63–1.47</td>
<td>0.21–1.27</td>
<td>1.30–6.3</td>
<td>1.30–12.3</td>
<td>1.30–8.6</td>
<td>1.30–8.6</td>
</tr>
<tr>
<td>Antisocial personality disorder</td>
<td>0.29</td>
<td>0.78–1.87</td>
<td>0.28–1.87</td>
<td>0.19</td>
<td>0.78–2.03</td>
<td>0.19</td>
<td>0.78–2.03</td>
<td>0.19</td>
</tr>
<tr>
<td>Conduct disorder</td>
<td>0.19</td>
<td>0.78–2.03</td>
<td>0.28–1.87</td>
<td>0.19</td>
<td>0.78–2.03</td>
<td>0.19</td>
<td>0.78–2.03</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Values are odds ratios, with 95% confidence intervals in parentheses; *P < 0.05, **P < 0.01, ***P < 0.001.

Table 3 Covariance between personality measures and comorbid psychiatric disorders, represented as percentage of total covariance (r\text{tot})

<table>
<thead>
<tr>
<th>Variable</th>
<th>Major depression</th>
<th>Generalised anxiety disorder</th>
<th>Panic disorder</th>
<th>Any phobia</th>
<th>Alcohol dependence</th>
<th>Any drug dependence</th>
<th>Antisocial personality disorder</th>
<th>Conduct disorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>N (%)</td>
<td>0.41</td>
<td>0.39</td>
<td>0.61</td>
<td>0.72</td>
<td>0.53</td>
<td>0.61</td>
<td>0.53</td>
<td>0.61</td>
</tr>
<tr>
<td>E (%)</td>
<td>0.09</td>
<td>0.09</td>
<td>0.10</td>
<td>0.12</td>
<td>0.11</td>
<td>0.12</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td>NS (%)</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>r\text{res} (%)</td>
<td>0.41</td>
<td>0.39</td>
<td>0.61</td>
<td>0.72</td>
<td>0.53</td>
<td>0.61</td>
<td>0.53</td>
<td>0.61</td>
</tr>
<tr>
<td>r\text{ind} (%)</td>
<td>0.41</td>
<td>0.39</td>
<td>0.61</td>
<td>0.72</td>
<td>0.53</td>
<td>0.61</td>
<td>0.53</td>
<td>0.61</td>
</tr>
</tbody>
</table>

N, neuroticism; E, extraversion; NS, novelty seeking; r\text{res}, residual covariance; IND, indirect covariance.
effects of extraversion on psychiatric disorders and comorbidity is inconsistent with previous research (Sher & Trull, 1994) and probably stems from the restrictive definition of our extraversion scale, which only reflects sociability. Eysenck revised the extraversion scale in the EPQ–R and items that measured impulsivity were largely moved to the psychoticism scale (Nyborg, 1997).

**Novelty seeking**

High novelty seeking increased the risk for externalising disorders significantly (Table 2) when these disorders were examined individually. Novelty seeking also accounted for the largest proportion of comorbidity between externalising disorders (7–14%, arithmetic average=11.9%). Not surprisingly, novelty seeking was unrelated to the comorbidity within internalising disorders and, for the most part, between internalising and externalising disorders. However, somewhat surprisingly, the contribution of neuroticism to the comorbidity within externalising disorders was comparable with the effects of novelty seeking.

These results further support the existence of broader, underlying dimensions of core psychopathological processes. Neuroticism appears to be a robust underlying dimension not only for the comorbidity within internalising disorders but also between internalising and externalising disorders and within externalising disorders. This leads us to reconsider the issue of psychiatric classification and an age-old question of splitting neurosis (Tyre, 1985). Our previous research has indicated that the comorbidity between major depression and GAD and, to some extent, between major depression and alcohol dependence largely results from common genetic factors (Kendler et al, 1992, 1993a) with notable gender differences (Prescott et al, 2000). In a previous report, we also found that over 50% of the genetic liability for major depression was shared with neuroticism (Kendler et al, 1993b). Thus, the possibility of common genetic liability between personality and comorbid disorders appears to be a reasonable hypothesis and will be the subject of future investigation.

**Limitations**

The results of this study should be interpreted in the context of four potential methodological limitations.

---

**Clinical Implications**

- Comorbidity among psychiatric disorders is a common and consistently reported finding.
- The normal personality dimension of neuroticism appears to be a broad vulnerability factor for the comorbid psychiatric disorders.
- Novelty seeking is modestly important for the comorbidity between externalising disorders only.

**Limitations**

- The normal personality dimensions used were from two different scales.
- The cross-sectional nature of the data used has a potential to confound state, trait and scar effects.
- The sample was limited to Caucasian individuals so the results might not be generalisable to other ethnic groups.

---

First, we used scales of neuroticism and extraversion from the EPQ–R and novelty seeking from the TPQ. Although neuroticism and extraversion represent widely accepted higher dimensions of personality, there is no agreement about the lower-order dimensions among different personality researchers. Moreover, some would argue that these two scales provide an incomplete description of the structure of heritable personality differences (Heath et al, 1994). How much more of the covariation among disorders would have been explained if we used the complete EPQ–R (neuroticism, extraversion, psychoticism and lie scale) or the complete TPQ (novelty seeking, harm avoidance and reward dependence) is speculative. Similarly, although interrater agreement for diagnosis was high, test-retest reliability for some of the lower-prevalence disorders (i.e. GAD, panic disorder and antisocial personality disorder) was low (0.23–0.42). This lower reliability may have increased the variance due to random errors of measurement, lowering the strength of associations of comorbidity with personality.

Second, the cross-sectional nature of the data made it difficult to establish causality and had a potential to confound state, trait and scar effects. However, the use of lifetime diagnosis provided some assurance that the confounding effects were likely to be minimal.

Third, because of some relatively young individuals in our sample, the risk period for certain psychiatric disorders was not over. As a result, true prevalence may be underestimated in the present sample, with concomitant effects on covariance.

Fourth, the sample was limited to Caucasian individuals so the results may not be generalisable to other ethnic groups.

**Acknowledgements**

Supported by National Institutes of Health (NIH) grants T32MH-20030, MH-40828, MH/AA/DA-49492, DA-11287 and AA-09095. The authors thank Dr Steve Aggen for his help in statistical analysis. We
also acknowledge the contribution of the Virginia Twin Registry, now part of the Mid-Atlantic Twin Registry (MATR), to the ascertainment of subjects for this study. The MATR, directed by Drs L. Corey and L. Eaves, has received support from the National Institutes of Health, the Carman Trust and the W. M. Keck, John Templeton and Robert Wood Johnson Foundations.

REFERENCES


Background  Hippocampal volume reduction has been reported inconsistently in people with major depression.

Aims  To evaluate the interrelationships between hippocampal volumes, memory and key clinical, vascular and genetic risk factors.

Method  Totals of 66 people with depression and 20 control participants underwent magnetic resonance imaging and clinical assessment. Measures of depression severity, psychomotor retardation, verbal and visual memory and vascular and specific genetic risk factors were collected.

Results  Reduced hippocampal volumes occurred in older people with depression, those with both early-onset and late-onset disorders and those with the melancholic subtype. Reduced hippocampal volumes were associated with deficits in visual and verbal memory performance.

Conclusions  Although reduced hippocampal volumes are most pronounced in late-onset depression, older people with early-onset disorders also display volume changes and memory loss. No clear vascular or genetic risk factors explain these findings. Hippocampal volume changes may explain how depression emerges as a risk factor to dementia.

Declaration of interest  Supported by National Health and Medical Research Council Program Grant No. 953208.

The hippocampus plays a key role in the regulation of mood and cognition and has been the subject of increased evaluation in people with mood disorders (Hickie et al, 1997a). To date, structural imaging studies of hippocampal volumes have returned mixed results. A significant number of negative studies (Vakil et al, 2000; Posener et al, 2003) have been interspersed with reports of unilateral (Shah et al, 1998) or bilateral (Sheline et al, 1999) volume reduction. Consequently, it is unclear whether hippocampal changes are restricted to older people with mood disorders, key clinical subgroups (e.g. late-onset, melancholia) or those with other vascular or genetic risk factors (e.g. isoforms of apolipoprotein E (ApoE) or the methylenetetrahydrofolate reductase (MTHFR) enzyme; Hickie et al, 2001). In this study, we sought to examine the interrelationships between hippocampal volume changes, visual and verbal memory function and key clinical, vascular and genetic risk factors in older persons with major depression.

METHOD

Participants
As part of a wider study of clinical, genetic and neuropsychological correlates of major depression (Hickie et al, 2001), 66 individuals with primary major depressive disorders (age range 28–82 years; mean=53.5, s.d.=13.5) were recruited from specialist service centres. These facilities attract somewhat older patients who have failed to respond to treatment in primary care services. Twenty healthy control participants (age range 40–74 years; mean=55.8, s.d.=10.0) were recruited via newspaper advertisement.

Potential participants were excluded if there was any indication of neurodegenerative disorder, history of stroke, head injury, substance misuse or medical contraindications to magnetic resonance imaging (MRI) scanning. Individuals who had received electroconvulsive therapy within the preceding 3 months also were excluded. All participants gave written informed consent prior to participation.

Clinical assessment
Psychiatrists performed structured clinical assessments (Hickie et al, 2001) generating DSM-IV (American Psychiatric Association, 1994) diagnoses. Additionally, severity of psychomotor change was evaluated using the CORE scale (Parker et al, 1994) and depression severity was rated using the 21-item Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960). Duration of current episode (maximum=104 weeks) was recorded, and duration since onset of illness (total years since onset) was calculated by subtracting age of depression onset from current age.

Participants with depression were subclassified into DSM-IV (American Psychiatric Association, 1994) non-melancholic (n=19, 29%) or melancholic (n=47, 71%; including 13 individuals with psychotic features) subtypes. Those who had their first episode of depression prior to age 50 years were classified as having ‘early-onset’ depression (n=49, 74%) whereas those who first experienced depression at age 50 years or later were classified as having ‘late-onset’ depression (n=17, 26%). Fourteen (88%) of those with late-onset depression also had a diagnosis of melancholia in comparison with 32 (65%) of those with early-onset depression (χ²=3.2, NS). The total years since onset of illness ranged from 0 to 60, with an average duration of 15 years (s.d.=15.8). Participants with early- and late-onset depression had mean lifetime illness duration of 19.3 (s.d.=16.4) and 3.5 (s.d.=2.8) years, respectively. Those with late-onset depression were significantly older (mean age=63.7 years, s.d.=10.4) than those with early-onset depression (mean age=50.1 years, s.d.=12.7; F=15.6, P<0.001).

Neuropsychological assessment
All participants were administered the Mini-Mental State Examination (MMSE; Folstein et al, 1975). As part of a wider neuropsychological assessment (Naismith et al, 2003), a subset of control participants (n=19) and participants with depression (n=46) were administered the Rey
Magnetic resonance imaging
Participants underwent high-resolution MRI scanning (124 × 1.5 mm coronal slices; time to repetition = 24 ms, time to echo = 5 ms, field of view = 26 cm, matrix 256 × 256) using a 1.5 T GE Signa machine. Data were transferred to a Silicon Graphics workstation and analysed using the BRAINS software package (Andreasen et al., 1993). Images were re-sampled digitally in the anterior commissure–posterior commissure plane to standardise anatomical orientation. Whole-brain volumes were traced using methods described previously (Levitan et al., 1999). All slices of the left and right hippocampi were traced manually by a rater masked to diagnosis. Although all traces were made in the coronal plane, additional traces were made on sagittal and axial views, and points from these were telegraphed to orthogonal planes, to be used as guidelines to tracing. Volumes (cm³) of each structure were summed across coronal slices to give total left and right hippocampal volumes. Definitions of anatomical boundaries and landmarks were derived from the literature (Cook et al., 1992; Watson et al., 1992), by consultation with a neuroanatomist and by use of a brain atlas (Duvernoy, 1991).

Vascular risk factors
Based on a combination of self-report and close informant questionnaires and medical review by a psychiatrist, the following vascular risks were recorded as present (1) or absent (0): diabetes; treated or untreated hypertension; smoking; cardiovascular disease; elevated cholesterol; and family history of at least two vascular disorders (including stroke and transient ischaemic attack). These six vascular risk factors were summed for each participant to give a total risk rating (range: 0–6; Hickie et al., 2001; Naismith et al., 2003).

Apolipoprotein E and methylenetetrahydrofolate reductase genotyping
As described in our previous studies (Hickie et al., 2001; Naismith et al., 2002, 2003), genotypes of ApoE and MTHFR were determined by polymerase chain reaction-based methods. Heterozygous (n = 31) and homozygous (n = 9) groups for the C677T MTHFR mutant allele were pooled to form a group ‘at risk’ (n = 40). Similarly, participants with at least one ApoE ε2 (ApoE2) or ε4 (ApoE4) allele were coded as either positive (n = 9 and n = 23, respectively) or negative (n = 71 and n = 57, respectively) for the allele.

Statistical analysis
Data were analysed using the Statistical Package for the Social Sciences (version 11.5 for PC). An α level of 0.05 was employed for all tests except those employing Bonferroni corrections.

RESULTS
As shown in Table 1, there was no difference in age or gender between persons with depression and control participants. Those with depression had more vascular risk factors, lower MMSE scores and poorer memory performance. They also had smaller whole-brain volumes and smaller left, right and total hippocampal volumes.

Table 1 Demographic, clinical, cognitive, vascular, genetic and magnetic resonance imaging (MRI) data for control participants and participants with depression

<table>
<thead>
<tr>
<th></th>
<th>Control participants (n = 20)</th>
<th>Participants with depression (n = 66)</th>
<th>F/χ²</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years): mean (s.d.)</td>
<td>55.8 (10.0)</td>
<td>53.5 (13.5)</td>
<td>0.5</td>
<td>NS</td>
</tr>
<tr>
<td>Gender: % female (n/N)</td>
<td>55 (11/20)</td>
<td>67 (44/66)</td>
<td>0.3</td>
<td>NS</td>
</tr>
<tr>
<td>MMSE score: mean (s.d.)</td>
<td>28.5 (1.4)</td>
<td>26.4 (3.7)</td>
<td>6.1</td>
<td>0.016</td>
</tr>
<tr>
<td>RAVLT score: mean (s.d.)</td>
<td>85.4 (16.7)</td>
<td>73.3 (18.1)</td>
<td>6.2</td>
<td>0.015</td>
</tr>
<tr>
<td>BVRT score: mean (s.d.)</td>
<td>7.6 (1.6)</td>
<td>5.5 (2.6)</td>
<td>11.6</td>
<td>0.001</td>
</tr>
<tr>
<td>HRSD score: mean (s.d.)</td>
<td>--</td>
<td>24.9 (9.2)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Age of onset (years): mean (s.d.)</td>
<td>--</td>
<td>38.4 (16.3)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>CORE score: mean (s.d.)</td>
<td>--</td>
<td>12.2 (7.6)</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Cumulative vascular risk</td>
<td>1.1 (1.1)</td>
<td>2.0 (1.5)</td>
<td>6.9</td>
<td>0.010</td>
</tr>
<tr>
<td>ApoE2: % positive (n/N)</td>
<td>0 (0/20)</td>
<td>15 (9/60)</td>
<td>3.4</td>
<td>NS</td>
</tr>
<tr>
<td>ApoE4: % positive (n/N)</td>
<td>40 (8/20)</td>
<td>25 (15/60)</td>
<td>1.6</td>
<td>NS</td>
</tr>
<tr>
<td>MTHFR: % positive (n/N)</td>
<td>45 (9/20)</td>
<td>33 (31/59)</td>
<td>0.3</td>
<td>NS</td>
</tr>
<tr>
<td>MRI volume (cm³): mean (s.d.)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Whole-brain volume | 1354.4 (171.8) | 1256.5 (118.6) | 8.3  | 0.005  |

Although there was no relationship between age and hippocampal volumes in control participants, increasing age was associated with smaller hippocampal volumes in the participants with depression (Table 2). However, the relationship between age and hippocampal volumes was particularly evident for those with late-onset (r = –0.3, P = 0.026) depression.

Importantly, there was no association between cumulative vascular risk factors and hippocampal volumes or whole-brain volumes for participants with or without depression (Table 2). Hippocampal volumes were not significantly associated with depression severity, clinician-rated psychomotor change, duration of depressive episode, total number of years since depression onset (Table 2) or bipolar disorder (Table 3).

Neuropsychological performance
There was no association between visual and verbal memory performance and hippocampal volumes in control participants (Table 2). However, for those with depression there were significant associations between smaller left and total
hippocampal volumes and poorer general cognition (i.e. as measured by the MMSE) and memory.

Analysis of covariance indicated a significant difference in memory scores between control participants and those with early and late-onset depression, even after controlling for age (BVRT: $F_{2,63}=8.4$, $P=0.001$; RAVLT: $F_{2,63}=4.2$, $P=0.021$). After Bonferroni correction, visual memory scores were poorer for both depression groups relative to controls ($P=0.002$ and $P=0.004$ for early and late-onset depression, respectively). However, within this lower subsample, verbal memory scores were significantly lower for those with early-onset ($n=36$, $P=0.017$) but not late-onset ($n=10$, NS) depression relative to control participants.

### Hippocampal volumes

For participants with depression, there were significant relationships between current age, age of depression onset and hippocampal volumes (Tables 2 and 3). After controlling for age and whole-brain volume, there was a significant effect of age of onset group (i.e. control and early- and late-onset depression groups) on total ($F_{2,80}=4.5$, $P=0.015$), left ($F_{2,80}=5.3$, $P=0.007$) and right ($F_{2,80}=3.2$, $P=0.045$) hippocampal volumes. As shown in Fig. 1, people with early-onset depression had smaller total hippocampal volumes than controls but larger volumes than those with late-onset depression. For the left hippocampus, age, whole-brain volume and Bonferroni-corrected analyses revealed that participants with both early- and late-onset depression had smaller ($P=0.021$ and $P=0.013$, respectively) left hippocampal volumes than control participants, although they did not differ from each other. For the right and total hippocampal volumes only the participants with late-onset depression differed significantly from controls ($P=0.045$ and $P=0.013$, respectively), whereas those with early-onset depression did not differ from either control participants or those with late-onset depression.

### DSM–IV subtype

After controlling for age and whole-brain volume, there was a significant difference between participants with melancholia and non-melancholic depression and controls in left ($F_{2,80}=3.8$, $P=0.025$) but not right ($F_{2,80}=2.2$, NS) hippocampal volumes. Bonferroni analyses revealed that only participants with melancholia differed significantly from controls (left: $P=0.006$; total: $P=0.021$), whereas those with non-melancholic depression did not differ significantly from controls or those with melancholia.

### Apolipoprotein E and MTHFR

There was no difference in age between those positive and negative for the ApoE4 allele. As shown in Table 3, there was no significant difference in hippocampal volumes for those with depression who were positive and negative for the ApoE2 allele, whereas those with the ApoE4 allele had larger (i.e. not smaller) volumes. Participants with depression and the MTHFR gene mutation did not have smaller hippocampal volumes than those without the mutation.

### Multivariate predictors of hippocampal volumes

In order to identify the best predictors of total hippocampal volumes, significant univariate predictors were entered into a stepwise regression model after controlling for whole-brain volume (forced entry). Hence, the entered variables were age of onset group (i.e. control and early- and late-onset depression), DSM–IV subtype group (i.e. control, non-melancholic and melancholic), ApoE4 and age. The resulting model, accounting for 53% of the variance in hippocampal volumes ($F_{4,74}=20.8$, $P<0.001$), included whole-brain volume ($t=6.3$, $P<0.001$), age of onset group ($t=−2.8$, $P=0.007$), age ($t=−2.6$, $P=0.010$) and presence of the ApoE4 allele ($t=−2.3$, $P=0.024$). These predictors uniquely contributed to 25%, 4.8%, 4.4% and 3.3% of the variance.

#### Table 2

Pearson correlations between hippocampal volumes and demographic, cognitive and vascular risk factors for control participants and participants with depression

<table>
<thead>
<tr>
<th></th>
<th>Total</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>r</em></td>
<td><em>P</em></td>
<td><em>r</em></td>
<td><em>P</em></td>
</tr>
<tr>
<td><strong>Control participants (n=20)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>−0.2</td>
<td>NS</td>
<td>−0.2</td>
</tr>
<tr>
<td>MMSE</td>
<td>0.1</td>
<td>NS</td>
<td>0.04</td>
</tr>
<tr>
<td>RAVLT (n=19)</td>
<td>−0.2</td>
<td>NS</td>
<td>−0.4</td>
</tr>
<tr>
<td>BVRT (n=19)</td>
<td>−0.2</td>
<td>NS</td>
<td>−0.3</td>
</tr>
<tr>
<td>Cumulative vascular risk</td>
<td>0.1</td>
<td>NS</td>
<td>−0.01</td>
</tr>
<tr>
<td><strong>Participants with depression (n=66)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>−0.4</td>
<td>&lt;0.001</td>
<td>−0.4</td>
</tr>
<tr>
<td>MMSE</td>
<td>0.3</td>
<td>0.027</td>
<td>0.3</td>
</tr>
<tr>
<td>RAVLT (n=45)</td>
<td>0.3</td>
<td>NS</td>
<td>0.036</td>
</tr>
<tr>
<td>BVRT (n=46)</td>
<td>0.4</td>
<td>0.016</td>
<td>0.4</td>
</tr>
<tr>
<td>Age of depression onset</td>
<td>−0.3</td>
<td>0.005</td>
<td>−0.25</td>
</tr>
<tr>
<td>Years since illness onset</td>
<td>−0.01</td>
<td>NS</td>
<td>−0.09</td>
</tr>
<tr>
<td>HRSD score</td>
<td>0.1</td>
<td>NS</td>
<td>0.1</td>
</tr>
<tr>
<td>CORE score</td>
<td>−0.2</td>
<td>NS</td>
<td>−0.2</td>
</tr>
<tr>
<td>Duration of episode</td>
<td>−0.1</td>
<td>NS</td>
<td>−0.1</td>
</tr>
<tr>
<td>Cumulative vascular risk</td>
<td>−0.02</td>
<td>NS</td>
<td>−0.03</td>
</tr>
</tbody>
</table>

**Notes:** Table 2. Pearson correlations between hippocampal volumes and demographic, cognitive and vascular risk factors for control participants and participants with depression. BVRT: Benton Visual Retention Test; HRSD, Hamilton Rating Scale for Depression; MMSE, Mini-Mental State Examination; MRI, magnetic resonance imaging; RAVLT, Rey Auditory Verbal Learning Test.
**DISCUSSION**

In this study, individuals with primary major depressive disorders recruited from specialist service settings demonstrated reduced whole-brain and left and right hippocampal volumes, impaired verbal and visual memory and an increased number of clinical risk factors to vascular disease. Reductions in hippocampal volumes in these individuals (but not control participants) were correlated with age, age of onset and general cognitive and memory decrements. Importantly, although reductions in hippocampal volumes were more significant in older patients, in those with late-onset depression and in those with melancholia, those with early-onset depression also had smaller hippocampal volumes. Consistent with recent longitudinal research examining ApoE (Steffens et al., 2002), hippocampal volume reduction was not predicted by specific genetic risk factors to neurodegeneration, or by clinical or genetic risk factors to vascular disease.

**Functional significance of reduced hippocampal volume**

The reductions in hippocampal volumes in people with depression are of considerable functional significance because of their relationship with visual and verbal memory decrements. Although it is well established that people with depression have impaired memory, such functional deficits are often attributed to poor encoding of information, poor effort (Elliot, 1998) or difficulties with executive functioning. This study, however, supports previous research in people with chronic depression (Shah et al., 1998) and in those with subjective memory problems (von Gunten et al., 2000) in suggesting that impaired memory may be a direct consequence of structural change within the hippocampus.

Depression has been recognised increasingly as a risk factor for later dementia (Steffens et al., 2002), and a variety of explanatory models have been proposed (Jorm, 2001). Importantly, in our study, hippocampal volumes were also reduced in persons with early-onset disorders, making it less likely that the onset of depression simply reflects an early phase of another dementing illness such as Alzheimer’s disease or vascular dementia. Consistent with this interpretation, hippocampal volume reductions were not predicted by the ApoE4 allele or at-risk isoforms of the MTHFR gene or clinical risk factors to vascular disease.

**Potential preventive strategies**

Because hippocampal atrophy was most pronounced in people with depression who were older at assessment or had late-onset disorders, potential risk factors that increase with age (e.g., neurodegeneration, vascular disease) still remain. Explanatory models have also been proposed (Hickie et al., 2003). Previously we have reported strong associations between both white matter and subcortical nuclei (i.e., caudate nucleus volume) structural brain changes and neurocognitive impairment, vascular risk factors, age, age of depression onset and poor response to treatment (Hickie et al., 1995, 1997, 1999; Shah et al., 2001), and reduced psychomotor speed in patients with depression (Naismith et al., 2002). Such studies do imply common pathophysiology underpinning the epidemiological association between at least late-onset depressions and dementia.

**Is depression associated with neurodegenerative changes?**

Importantly, it now also appears likely that hippocampal atrophy occurs directly as a consequence of early-onset depression (or other risk factors to that condition). Consistent with this view, lifetime duration of untreated depressive illness has emerged as
a predictor of such hippocampal changes (Sheline et al., 1999, 2003; MacQueen et al., 2003). Although we did not find a direct correlation with years since onset of the illness, we were not able to differentiate the importance of treated vs. untreated periods of illness. In our study, participants with melancholic disorders demonstrated more hippocampal atrophy. Such people are more likely to experience hypercortisol-emia, which is a possible mechanism for hippocampal atrophy (Sapolsky, 2000). An accumulation of evidence is also emerging suggesting that brain-derived neurotrophic factor (important for the development, maintenance and survival of neurons) is decreased in patients with depression and is enhanced by anti-depressant treatment (Duman et al., 1997; Dwivedi et al., 2003). This suggests another important mechanism whereby untreated depression may be detrimental to key brain structures such as the hippocampus, which in turn is likely to have prognostic significance (Steffens et al., 2002).

Important challenges arise from this research. First, we need to determine whether hippocampal atrophy is a risk factor for or a consequence of depressive disorders or to key subtypes (e.g. late-onset depression, melancholia). Second, we need to make greater use of population-based cohorts or other informative samples (e.g. twins, discordant sib-pairs). Third, more research needs to focus on longitudinal examination of at-risk groups and follow, in particular, the brain changes that may accompany either the transition to illness or the longer-term effects of its untreated or treated course.

REFERENCES


Hypofrontality in men with first-episode psychosis

V. MOLINA, J. SANZ, S. REIG, R. MARTÍNEZ, F. SARRAMEA, R. LUQUE, C. BENITO, J. D. GISPERT, J. PASCAU and M. DESCOS

Background Decreased metabolic activity in the prefrontal cortex during cognitive activation is a recurrent finding and a likely functional marker of schizophrenia.

Aims To investigate the occurrence of hypofrontality in patients with first-episode psychosis, with or without evolution to schizophrenia.

Method We used fluorodeoxyglucose positron emission tomography during the performance of an attention task and magnetic resonance imaging to study the dorsolateral prefrontal region in 13 men with a first episode of psychosis. Data from patients who progressed to schizophrenia were compared with those of patients who did not meet criteria for this diagnosis after 2 years.

Results Patients who developed schizophrenia demonstrated a significant hypofrontality in the dorsolateral prefrontal cortex in comparison with the non-schizophrenia and control groups.

Conclusions Our results suggest that hypofrontality could be a marker of schizophrenia at the time of the first psychotic episode, in agreement with neurodevelopmental theories of schizophrenia.

Declaration of interest None. Funding detailed in Acknowledgements.

Decreased metabolic activity in the prefrontal cortex during cognitive activation is a recurrent finding in schizophrenia (Volz et al., 1999). If it were not present in other psychotic disorders, hypofrontality could constitute a distinctive marker for schizophrenia. Although other groups have published data on hypofrontality in schizophreniform disorders (Andreasen et al., 1997), we know of none comparing metabolic patterns in these patients in the light of whether or not they progress to schizophrenia. To confirm hypofrontality as a feature typical of schizophrenia, it would be advisable to study patients in very early stages of the disease. This design requires a corroborative diagnosis of schizophrenia by a prospective follow-up. With this goal, we used magnetic resonance imaging (MRI) and [18F]fluorodeoxyglucose (FDG) positron emission tomography (PET) to investigate two groups of people with first-episode psychosis: one group in which the diagnosis of schizophrenia was confirmed 2 years later, and a group who did not meet the criteria for this diagnosis after the same period.

METHOD

Study participants The existence of gender-related differences in frontal activity cannot be ruled out, considering that structural alterations seem to be more pronounced in men (Nopoulos et al., 1997). For this reason, and to achieve maximum sample homogeneity, our study was limited to men. Thirteen right-handed men with a first psychotic episode were enrolled in a diachronic study of first-break schizophrenia; all of them were enrolled while hospitalised for the first time in the psychiatric unit of a general hospital. The PET and MRI scans were performed at the time of inclusion in the study, before any of the patients had met the criteria for schizophrenia. To be included in the protocol, patients must have presented with first-episode psychosis with symptoms lasting more than 1 week, not attributable to organic or toxic origin, and unrelated to any other Axis I disorder; the presence of any relevant stressors clearly related to the episode had to be ruled out. We used these criteria in order to avoid recruiting patients with transient psychotic symptoms, according to DSM-IV criteria (American Psychiatric Association, 1994).

After inclusion, all patients were followed for 2 years on an outpatient basis with monthly visits, to confirm or rule out a diagnosis of schizophrenia at the end of this period. After this follow-up, two experienced psychiatrists (V.M. and J.S.), masked to the results of the PET scans, diagnosed each patient and decided whether the index episode was a first break of schizophrenia or, instead, a single psychotic episode that did not evolve into schizophrenia 2 years after the enrolment. For this final diagnosis they used all the available follow-up information (including total duration of significant symptoms), a semi-structured interview, the Structured Clinical Interview for DSM-IV Access I Disorders – Clinician Version (SCID-CV; First et al., 1997) and the information provided by the families and clinical personnel. For diagnostic purposes, the duration of symptoms was calculated covering the period before and after initiation of treatment.

On the basis of their final diagnosis, the 13 patients were divided into two groups. The schizophrenia group comprised six people with a diagnosis of paranoid schizophrenia according to DSM-IV criteria. The non-schizophrenia group comprised six people with a diagnosis of schizophreniform psychosis and one person who had experienced a brief psychotic disorder (with a symptom duration of 3 weeks), likewise according to DSM-IV criteria. The latter diagnoses meant that these patients’ psychotic episodes were shorter than required to qualify as schizophrenia, no other psychotic episode occurred during follow-up and the patients did not meet schizophrenia criteria (including enduring negative symptoms) at the end of follow-up. The duration of the disease in the non-schizophrenia group was less than 3 months in all cases. The clinical scores calculated using the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1992) at the time of admission and at the end of follow-up are shown in Table 1.
All patients were antipsychotic-naive prior to the presenting episode; they were treated solely during the 2 days prior to the PET scan. This was a practical means of enrolling a sample representative of the usual presentation of schizophrenia, avoiding a selection bias related to the impaired ability of patients with acute psychosis to cooperate during neuroimaging procedures. All patients received the same treatment (haloperidol 10 mg per day liquid formulation) and the medication was discontinued for 12 h prior to the PET scan. No other treatment was administered before scanning. One patient from the schizophrenia group and two from the non-schizophrenia group had received benzodiazepines for insomnia before inclusion, but they had discontinued this medication for more than 2 weeks. No acute dystonic reaction occurred, thus there was no need to administer anticholinergic medication before scanning. Nursing personnel verified treatment compliance.

Participants were instructed to fast for at least 12 h prior to scanning. Each participant began the task, which was divided into four blocks of 5 min each, with a 1 min rest between blocks; FDG was administered 1 min after initiating the task. Participants were instructed to push a button if a letter T immediately followed the letter L, presented on a computer screen. The interstimulus interval was 1 s. During the PET scan a neuropsychologist supervised the appropriate response and full compliance with the performance test of all participants. After an intravenous line had been inserted for FDG administration, the participant began the task, which was divided into four blocks of 5 min each, with a 1 min rest between blocks; FDG was administered 1 min after initiating the task. Participants were instructed to fast for at least 6 h before the scan, with particular emphasis on avoiding coffee and other psychoactive beverages.

Tracer activity values were proportionally normalised to the global activity of each scan (Frackowiak et al, 1997) thus representing relative activity. Total metabolic activity for each region of interest was divided by the region’s volume, thus providing a measurement independent of the amount of tissue sampled.

**Image analysis**

To determine the metabolic activity in the dorsolateral prefrontal cortex and the occipital lobe, we used a procedure based on MRI/PET image fusion. This method uses the anatomical information of the MRI to allow detailed measurement of regional metabolic activity in the PET image (Molina et al, 2003).

**MRI protocol**

Magnetic resonance images were acquired on a Philips Gyroscan 1.5 T scanner (Philips Medical Systems, Reigate, UK) using a T₁-weighted three-dimensional gradient echo sequence with the following parameters: matrix size 256 × 256, pixel size 0.9 mm × 0.9 mm (field of view approximately 256 × 256 mm), flip angle 30°, echo time 4.6 ms, slice thickness 1.1–1.5 mm. In addition, T₂-weighted sequences were acquired for tissue segmentation and for other clinical purposes (turbo spin echo, turbo factor 15, echo time 120 ms, matrix size 256 × 256, slice thickness 5.5 mm).

**PET protocol**

The PET study used a Siemens Exact 47 tomograph (Siemens, Erlangen, Germany). Scans were obtained 20 min after injecting 370 MBq of [¹⁸F]FDG, while the participant performed a contingent continuous performance task (Rosvold et al, 1956). Participants were instructed to push a button if a letter T immediately followed the letter L, presented on a computer screen. The interstimulus interval was 1 s. During the PET scan a neuropsychologist supervised the appropriate response and full compliance with the performance test of all participants. After an intravenous line had been inserted for FDG administration, the participant began the task, which was divided into four blocks of 5 min each, with a 1 min rest between blocks; FDG was administered 1 min after initiating the task. Participants were instructed to fast for at least 6 h before the scan, with particular emphasis on avoiding coffee and other psychoactive beverages.

Tracer activity values were proportionally normalised to the global activity of each scan (Frackowiak et al, 1997) thus representing relative activity. Total metabolic activity for each region of interest was divided by the region’s volume, thus providing a measurement independent of the amount of tissue sampled.

**Segmentation**

To obtain metabolic measurements of the prefrontal cortex and occipital lobe, we used a method for semi-automated segmentation of the brain based on the Talairach reference system, similar to that described
by Andreasen et al (1996) and Kates et al (1999). We used a two-step procedure (Desco et al, 2001): the first step involved editing the MRI to remove skull and extracranial tissue, registration of the PET and MRI scans, and an initial segmentation of cerebral tissues into grey matter, white matter and cerebrospinal fluid; in the second stage we applied the Talairach reference system (Talairach & Tournoux, 1988) to define regions of interest and to obtain final metabolic activity data (Fig. 1).

The edited MRI (without extracranial tissue) was co-registered with the PET image using the Automated Image Registration (AIR) algorithm (Woods et al, 1993). Fusion results were visually checked in all cases and the observed co-registration was always optimal. An initial segmentation of cerebral tissue was performed using an automated method, currently included as a standard processing tool in the Statistical Parametric Mapping (SPM) program (Ashburner & Friston, 1997). This method classifies all MRI pixels into four tissue types: grey matter, white matter, cerebrospinal fluid and 'other tissues'. The algorithm also removes the effect of radiofrequency field inhomogeneities (Ashburner & Friston, 1997). This segmentation was checked for inconsistencies and manually corrected whenever necessary by an experienced radiologist masked to diagnosis.

In the second stage, the regions of interest were obtained by superimposing the three-dimensional tissue masks corresponding to the first three tissue types on to each individual’s Talairach reference grid (Fig. 1), where the regions of interest were defined as sets of cells. On the MRI/PET fused images, volume and activity were measured by summing the data from the grid cells associated with each region of interest (Desco et al, 2001).

Validity of the Talairach-based procedure as an automated segmentation tool suitable for schizophrenia research has been proved (Andreasen et al, 1996; Kates et al, 1999). A single operator performed all manual procedures, thus avoiding any potential interrater variability. Reliability of the entire segmentation procedure had been previously assessed by repeating the measurements three times for 15 cases selected at random. Repeatability of metabolic activity values ranged from 93% to 99% (Molina et al, 2003).

Region of interest variables included in the analysis were the dorsolateral prefrontal cortex and the occipital lobe, which was used only as a reference structure (see below). The dorsolateral prefrontal cortex was defined as the cortex encompassing Brodmann areas 8, 9, 10 and 46. The occipital lobe was defined according to boundaries previously established by other authors also using the Talairach method (Andreasen et al, 1996; Kates et al, 1999). We studied this brain region as a control for potential global effect bias in our measurements of metabolic activity due to the short-course treatment (Bartlett et al, 1994). Data from the occipital cortex allowed us to discard a potential difference in overall metabolic activity between patient groups. The absence of bias due to a global effect was confirmed by the similar values of metabolic activity in the occipital region found in all the groups studied.

Since there were slight age differences between groups, we used Spearman’s correlation coefficients to discard a significant effect of age on activity data. None of the correlation coefficients between age and metabolic activity variables was statistically significant.

**Statistical analysis**

The study hypothesis was approached by comparing the relative activity in both the right and left dorsolateral prefrontal cortex in the three groups: the schizophrenia patient group, the non-schizophrenia patient group and the control group. Because of the small sample sizes we used non-parametric statistics: the Kruskal–Wallis test to assess overall differences among the three groups, and the Wilcoxon rank sum test for independent samples to check the differences between the schizophrenia group and each of the other two comparison groups. All analyses were carried out using the Statistical Analysis System (SAS) version 6.12.

**RESULTS**

The mean activity for each region is shown in Table 2. In the comparison among the
groups, we found a significant effect of the diagnosis on the left side ($\chi^2=9.54$, d.f. = 2; $P=0.008$, Kruskal–Wallis test). We found the same tendency on the right side ($\chi^2=5.60$, d.f. = 2; $P=0.06$, Kruskal–Wallis test). The schizophrenia group showed hypofrontality in comparison with the control group in the Wilcoxon test ($Z=-2.51$, $P=0.01$, on the left side). The schizophrenia group also demonstrated hypofrontality compared with the non-schizophrenia group in the Wilcoxon test ($Z=-2.21$, $P=0.03$, on the left side; $Z=-2.50$, $P=0.01$, on the right side) (Fig. 2). There was no difference in the metabolic activity values of the occipital cortex among the three groups ($\chi^2=0.52$, d.f. = 2, $P=0.78$; $\chi^2=3.47$, d.f. = 2; $P=0.18$, respectively for the left and right sides, Kruskal–Wallis test).

**DISCUSSION**

### Hypofrontality in early stages of schizophrenia

Our results, obtained in a male sample population, agree with the hypothesis that people with a first episode of schizophrenia demonstrate hypofrontality during the performance of an attention task; people with acute psychosis who do not progress to schizophrenia do not demonstrate this anomaly. This finding fits with the neurodevelopment theories of schizophrenia, which postulate the existence of a set of alterations that appear to be present even in the earliest phases of the disease. This hypofrontality in the initial stages of first-episode schizophrenia is similar to that found in treatment-naïve patients and those with chronic schizophrenia (Buchsbaum et al., 1992; Schroeder et al., 1994). Although the participants in our schizophrenia group were slightly younger than those in the comparison groups, their hypofrontality cannot be attributed to age differences, since there was no association between age and activity in our sample. Besides, cortical metabolic activity in the normal population declines with age (Marchal et al., 1992).

**Pharmacological effects on hypofrontality**

The hypofrontality found in this study does not seem to be attributable to pharmacological effects, since both patient groups had received the same treatment before the PET scan. The higher activity in the dorsolateral prefrontal region in the psychosis patients without schizophrenia could be tentatively attributed to poorer treatment compliance by these patients than by those with schizophrenia. However, the nursing personnel carefully verified this aspect and the haloperidol was administered as a liquid formulation, so lack of compliance may reasonably be ruled out. On the other hand, there is no reason *a priori* to expect better treatment compliance from the schizophrenia group than from the non-schizophrenia group of patients. Also, in the event of better treatment compliance by the schizophrenia group, we should have observed a corresponding change in the metabolic activity of the occipital lobe (Bartlett et al., 1994), which was not the case (Table 2).

A different treatment response in each of the patient groups might explain the lower dorsolateral prefrontal cortical activity in the people who went on to develop schizophrenia than in the non-schizophrenia patient group. This possibility was suggested by a previous study (Bartlett et al., 1998), which showed that patients who did not respond to treatment demonstrated more marked changes in activity in this brain area than responders. This does not seem to be the case in our sample, since patients in our schizophrenia group did, in fact, respond to treatment, except for negative symptoms (Table 1). In addition, other studies have shown the same hypofrontal pattern in completely drug-free patients (Buchsbaum et al., 1992; Andreasen et al., 1997). On the other hand, if the hypofrontality were caused by differences in response or in compliance with haloperidol treatment, this would not...

---

Table 2  Metabolic activity values (per ml) for each group in the dorsolateral prefrontal and occipital regions (significant differences in comparison of the values of the schizophrenia group vs. the control and non-schizophrenia groups by a Wilcoxon rank sum test for two independent samples are marked with asterisks)

<table>
<thead>
<tr>
<th>Cortical region</th>
<th>Schizophrenia group (n=6)</th>
<th>Non-schizophrenia group (n=7)</th>
<th>Control group (n=8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (s.d.)</td>
<td>Mean (s.d.)</td>
<td>Mean (s.d.)</td>
</tr>
<tr>
<td>Left DLPF</td>
<td>93.5 (4.4)</td>
<td>102.1 (4.5)**</td>
<td>102.8 (7.7)**</td>
</tr>
<tr>
<td>Right DLPF</td>
<td>99.2 (3.2)</td>
<td>105.8 (5.2)*</td>
<td>105.0 (6.3)</td>
</tr>
<tr>
<td>Left occipital</td>
<td>104.0 (5.7)</td>
<td>102.4 (2.7)</td>
<td>104.9 (5.3)</td>
</tr>
<tr>
<td>Right occipital</td>
<td>104.3 (6.0)</td>
<td>100.0 (3.0)</td>
<td>103.4 (3.4)</td>
</tr>
</tbody>
</table>

DLPF: dorsolateral prefrontal. *p < 0.05; **p < 0.01.

---

Fig 2  Raw values of metabolic activity for the left and right dorsolateral prefrontal cortex in patients and controls. Mean values for each group are shown linked. Values for the schizophrenia group (n=6) were significant in the left and right dorsolateral prefrontal cortex compared with the non-schizophrenia group (n=7) and also in the left dorsolateral prefrontal cortex when compared with the control group (n=8).
explain the absence of differences in the occipital cortex, since the aforesaid decrease in activity – observed solely in the non-responders – was generalised in the cortex (Barlett et al., 1998). Finally, the Scottish Schizophrenia Research Group (Anonymous, 1998) reported the existence of hypofrontality prior to initiation of treatment and its persistence after 6 months of treatment. Analysed together, these data support the fact that no treatment-related factor would explain the differences between our patient groups.

**Hypofrontality in other types of psychosis**

There is little information on cerebral activity in schizophreniform psychoses with good prognosis. Nevertheless, clinical data support the theory that cerebral substrates may differ between schizophreniform and schizophrenic psychoses; for example, in a recent study persistent primary symptoms were not found in 61 people with schizophreniform psychosis, whereas these sympotms were found in 26% of 338 people with schizophrenia (Peralta & Cuesta, 2004). However, there are several studies of bipolar disorder, which also may commence clinically with first-episode psychosis and has a better prognosis than schizophrenia. There is evidence of decreased dorsolateral prefrontal metabolic activity under activation conditions (word generation and decision-making) during manic episodes (Blumberg et al., 1999; Rubinstein et al., 2001), although these studies were not performed in the initial stages of the disease, and participants were receiving various medications. In the case of bipolar mania, hypoactivation may be right-lateralised (Blumberg et al., 1999; Rubinstein et al., 2001) and accompanied by increased activity in the caudate and cingulate regions (Blumberg et al., 2000; Rubinstein et al., 2001). These findings are different from the pattern observed not only in our patients with first-episode schizophrenia but also in untreated and in treated patients (Siegel et al., 1993; McGuire et al., 1998; Shihabuddin et al., 1998), suggesting that a degree of caution is essential before considering hypofrontality to be a specific marker for schizophrenia. In particular, our results do not support the supposition that hypofrontality may be a specific marker of schizophrenia, but only suggest that it could distinguish between first psychotic episodes of good prognosis and first episodes evolving to schizophrenia.

**Methodological issues**

Our study has various limitations. One of them is the small sample size. However, the short duration of the disease and the prospective diagnosis in patients with first-episode psychosis worked in our favour. Only a prospective follow-up can separate recent-onset schizophrenia from other psychotic episodes. This approach offers the advantage of studying the patients very early, as cerebral changes due to the disease may take place even in the first months (Gur et al., 1998; Rapoport et al., 1999; Pantelis et al., 2003). Besides, our patients were not exposed to other potentially brain-damaging situations, such as drug misuse or severe behaviour disorganisation. In consequence, although our sample is small, it should offer greater homogeneity than other studies using a cross-sectional design on a sample of people with first psychotic episodes. Another problem is that the brief treatment administered could hinder interpretation of the results, even though it did allow the study of a group that was truly representative of the usual presentation of acute schizophrenia. Neuroimaging studies could easily be biased towards cases in which the patient is likely to cooperate well with the procedures, which might lead to certain types of patients being more strongly represented in the sample. This problem was partially overcome by administering the same treatment to all patients prior to the PET scans. Concerning the specificity of our performance task for the frontal area of the brain, we did not perform a control task to rule out non-frontal aspects, as it would have complicated the FDG protocol significantly. Nevertheless, despite this lack of a control reference, the frontal involvement related to our performance task seems unquestionable (Rosvold et al., 1956). Finally, our findings cannot be directly generalised to women with schizophrenia, because of the greater number of structural anomalies in male patients (Nopoulos et al., 1997).

Other groups have also enrolled patients recently diagnosed with schizophrenia or schizophreniform disorder (Rubin et al., 1991). Our study group has the advantage of the prospective follow-up, which allowed us to differentiate between patients who were subsequently diagnosed as having schizophrenia and those who were not. In conclusion, our data suggest that the existence of hypofrontality could contribute to predicting which patients with a first episode of psychosis are likely to progress to schizophrenia.

**ACKNOWLEDGEMENTS**

The authors thank Dr M. A. Pozo and Dr A. Maldonado from the Centro PET Complutense for their advice and collaboration in performing the PET scans. The study was supported in part by grants from Fondo de Investigaciones Sanitarias (02/1178, 02/3095, Red Tematica IM3), Ministerio de Ciencia y Tecnologia TIC (2001-3697-C03-03) and Fundacion La Caixa (99/042-00).

**REFERENCES**


Neuroanatomical abnormalities before and after Decision-making in mania: a PET study. MRI comparison. to clinical and neurobehavioral measures. schizophrenia. Relationship of neuroanatomical changes follow-up magnetic resonance imaging study of schizophrenia. Relationship of neuroanatomical changes to clinical and neurobehavioral measures.

LIMITATIONS

Our results are limited to a male population. Conclusions should be confirmed in a larger sample size including both men and women.

The brief treatment administered might have affected our results. However, the possible confounding effect should be minimal because both groups of patients received the same treatment.

Other forms of psychosis, such as bipolar disorder or psychotic depression, may also show hypofrontality at early stages.

V. MOLINA, PhD, Department of Psychiatry, Hospital Clinico Universitario de Salamanca; J. SANZ, MD, Department of Psychiatry, Hospital Doce de Octubre, Madrid; S. REIG, PhD, R. MARTINEZ, PhD, Department of Experimental Medicine, Hospital Gregorio Marañón, Madrid; F. SARRAMEA, PhD, R. IJUQUE, MD, Department of Psychiatry, Hospital Reina Sofia, Córdoba; C. BENITO, MD, Clinica Monterprinciple, Hospital de Madrid; D. GSEPRT, PhD, J. PASCAL, MSc; M. DESCQ, PhD, Department of Experimental Medicine, Hospital Gregorio Marañón, Madrid, Spain.

Correspondence: Dr V. Molina, Department of Psychiatry, Hospital Clínico de Salamanca, Paseo de San Vicente, 58-182, 37007b Salamanca, Spain. Tel: +34 923 291383; Fax: +34 923 229158; e-mail: vmolina@usal.es.

(First received 15 December 2003, final revision 23 September 2004, accepted 30 September 2004.)

CLINICAL IMPLICATIONS

- Hypofrontality seems to be a marker of schizophrenia even at the time of the patient’s first psychotic episode.
- If the predictive value of hypofrontality is confirmed, decisions about the most suitable treatment will be facilitated.
- The presence of hypofrontality in first-episode psychosis fits with the neurodevelopmental theories of schizophrenia.

- Our results are limited to a male population. Conclusions should be confirmed in a larger sample size including both men and women.
- The brief treatment administered might have affected our results. However, the possible confounding effect should be minimal because both groups of patients received the same treatment.
- Other forms of psychosis, such as bipolar disorder or psychotic depression, may also show hypofrontality at early stages.
Neural correlates of syntax production in schizophrenia

TILO T. J. KIRCHER, TOMASINA M. OH, MICHAEL J. BRAMMER and PHILIP K. McGUIRE

Background The production of grammatically complex sentences is impaired in schizophrenia. It has been suggested that impaired syntax processing reflects a risk for the disorder.

Aims To examine the neural correlates of syntax production in people with schizophrenia using functional magnetic resonance imaging (fMRI).

Method Six patients with schizophrenia and six healthy volunteers spoke about seven Rorschach inkblots for 3 min each while correlates of brain activation were measured with fMRI. Participants produced varying amounts of syntactically simple and complex sentences during each 3 min run. The number of simple and complex sentences was correlated separately with the BOLD contrast.

Results In the comparison between the control group and the patient group, the number of complex sentences produced was correlated with activation in the posterior portion of the right middle temporal (Brodmann area 21) and left superior frontal (BA 10) gyri in the control group but not in the patients.

Conclusions The absence of activation in the right posterior temporal and left superior frontal cortex in patients with schizophrenia might contribute to the articulation of grammatically more simple speech in people with this disorder.

Declaration of interest None.

Syntax refers to the way that words are combined grammatically to form meaningful phrases or sentences. Relative to healthy controls, patients with schizophrenia have deficits with both comprehension (Bagner et al., 2003) and production (Thomas et al., 1996; Oh et al., 2002) of syntactically complex sentences. These deficits are especially evident in patients with formal thought disorder (Rodriguez-Ferrera et al., 2001) and do not seem to be simply a function of impaired semantic processing or general intellectual functioning (Condray et al., 2002). It has been proposed that impaired syntax processing reflects a risk for the disorder (DeLisi, 2001).

In this study we examined the neural correlates of syntax generation on a sentence level, using functional magnetic resonance imaging (fMRI) to observe participants articulating grammatically complex sentences. As syntactic processing in healthy individuals involves the inferior frontal and lateral temporal cortices (Kaan & Swaab, 2002), we predicted that the production of complex sentences would be correlated with activation in these areas in the control group, and that patients with schizophrenia would exhibit attenuated engagement of the left inferior frontal cortex and both left and right lateral temporal cortex.

METHOD

In the present investigation we used data derived from an earlier study: full details of the participants, the stimuli used, speech recording, fMRI and analysis are given by Kircher et al. (2001b, 2002). The patient group comprised six right-handed men with a DSM–IV diagnosis of schizophrenia (American Psychiatric Association, 1994), stabilised on neuroleptic medication, with prominent positive formal thought disorder. The control group consisted of six healthy individuals who were matched for age, gender, verbal IQ, immediate memory recall, attention and socio-demographic variables; see Kircher et al. (2002) for details.

During scanning, seven Rorschach inkblot plates were presented consecutively on a screen. Participants were asked to speak about whatever came to mind on viewing the inkblot, starting as soon as the stimulus appeared; they spoke freely and no prompting was given if they paused or stopped. Each of the seven plates was presented for 3 min (one ‘run’) with breaks of about 1 min between each presentation. The participants’ speech during scanning was recorded on a computer.

Linguistic analysis

Acoustic noise generated by image acquisition was filtered from the recordings of the participants’ speech, which was transcribed verbatim from these recordings and subsequently analysed from the transcripts. The number of simple sentences (e.g. ‘I see a moth’) and sentences with subordination (e.g. ‘I see a moth that is blue’) were evaluated using Brief Syntactic Analysis (Thomas et al., 1996a) by one of the authors (T.O.), who was masked to the speaker’s identity. Brief Syntactic Analysis is a linguistic analysis that rates spontaneous speech for syntactic complexity. In this analysis, sentences were classified as one of the following:

(a) Simple sentences: sentences consisting of a subject and predicate, and containing one finite verb, e.g. ‘It looks like a brain’.

(b) Complex sentences: sentences consisting of one independent clause together with one or more subordinate (embedded) clauses (e.g. ‘I jumped at the sound of his voice, which was inordinately loud in the silence’). In such sentences the relationship between the independent clause and its subordinate clauses is a hierarchical one; a complex sentence may therefore have one or more levels of embedding, depending on the relationship between the subordinate clauses and the main one.

Each 3 min scanning run was broken down into nine epochs of 20 s, and a total score for simple and complex sentences was obtained for each epoch.
**Image acquisition and data analysis**

The method of image acquisition and data analysis has been fully described by Kircher et al. (2002). Gradient-echo echoplanar MR images were acquired using a 1.5 T GE Signa System (General Electric, Milwaukee, Wisconsin, USA): 14 planes, 60 $T_2^*$-weighted MR images; time to echo (TE) 40 ms, repetition time (TR) 3000 ms, 0°–90°, in-plane resolution 3.1 mm, slice thickness 7 mm, slice skip 0.7 mm.

For fMRI data analysis (Brammer et al., 1997), the behavioural data (number of simple and complex sentences per 20 s epoch in the two runs per individual with the highest variance and highest number of maxima and minima) were interpolated using a cubic spline to obtain a behavioural value corresponding to each fMRI volume acquired (one value per TR). The dependent variable (observed time series at each voxel) was regressed on the independent variable (number of simple and complex sentences). Further analysis was performed as described by Kircher et al. (2002). For group comparisons, we computed a mask composed of all voxels activated at a stringent voxel-wise threshold (0.001) in either of the two groups or conditions that we needed to compare. We then carried out between-condition and between-group comparisons within this masked region to reduce the multiple comparison problem and thus enhance sensitivity. Comparison of groups was performed by computing the difference between the mean levels of activation at each voxel in standard space over all participants in each group. The significance of this difference was then assessed by reference to a null distribution obtained by repeated permutation of group membership and recomputation of the mean difference.

**RESULTS**

**Behavioural measures**

In the two Rorschach plates per individual used for the fMRI analysis, the number of simple sentences spoken by the patients ranged from 3 to 26 (mean 13.8, s.d. = 7.4); in the volunteers the range was 2–28 (mean 14.2, s.d. = 8.2; P = 0.9, Mann–Whitney $U = 71.5$). The number of complex sentences spoken by the patients ranged from 5 to 13 (mean 8.5, s.d. = 2.8), whereas in the volunteers the range was 6–20 (mean 13.5, s.d. = 4.1; P = 0.003, $U = 21.0$). There was no correlation between the number of complex and simple sentences produced in each group. The groups did not differ in the total number of sentences produced (controls: median 31.5, maximum 36.0; patients: median 29.5, maximum 39.0; $U = 71.0$, P = 0.9).

**Head movement**

Analysis of the estimated time series of displacements in three dimensions revealed no evidence of major head motion in any individual. The maximum amount of head movement in the three dimensions $x$, $y$ and $z$ in the runs per patient was $x = 0.5$ mm (s.d. = 0.3), $y = 0.7$ mm (s.d. = 0.3), $z = 1.4$ mm (s.d. = 1.6) and in the controls it was $x = 0.4$ mm (s.d. = 0.3), $y = 0.3$ mm (s.d. = 0.4) and $z = 0.8$ mm (s.d. = 0.8).

**Cerebral activation**

**Control group**

In the control group the number of simple sentences produced per 20 s period was positively correlated ($P < 0.001$) with the BOLD response in the left thalamus and primary visual cortex. The main correlations for complex sentences ($P < 0.001$) were with responses in the posterior part of the left middle temporal gyrus, Brodmann area (BA) 39, and in the posterior portion of the right superior temporal sulcus, BA 21/22 (Table 1). Between-condition differences in median correlation coefficient were examined using analysis of variance (ANOVA) at each voxel with probability of type I error at $P < 0.05$. Significantly stronger correlations for complex versus simple sentences were evident in the right posterior superior temporal sulcus (BA 21) and the left superior frontal gyrus (BA 10; Table 1).

**Patient group**

In the patient group the number of simple sentences produced per 20 s period was positively correlated ($P < 0.001$) with the BOLD response in the cerebellum bilaterally (vermal lobule VI; Schmahmann et al., 1999). The main correlations for complex sentences ($P < 0.001$) were with responses in the middle temporal gyri bilaterally (BA 21), in a part anterior and inferior to that activated in the control group. Between-condition differences in median correlation coefficient were examined using an ANOVA at each voxel. This indicated

---

**Table I** Main foci of signal changes in the control group during production of continuous speech: correlations with simple and complex sentences are shown ($P < 0.001$), followed by statistical differences between the two conditions ($P < 0.05$)

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>Coordinates (mm)</th>
<th>Voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$x$ $y$ $z$</td>
<td></td>
</tr>
<tr>
<td>Simple sentences</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thalamus</td>
<td>7</td>
<td>L</td>
<td>$-5$ $-8$ $1$</td>
<td>7</td>
</tr>
<tr>
<td>Primary visual cortex</td>
<td>18/19</td>
<td>L</td>
<td>$-30$ $-84$ $8$</td>
<td>6</td>
</tr>
</tbody>
</table>

**ANOVA**

<table>
<thead>
<tr>
<th>Simple sentences</th>
<th>Complex sentences</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; complex sentences</td>
<td></td>
</tr>
<tr>
<td>Middle occipital gyrus</td>
<td>19</td>
</tr>
<tr>
<td>Complex sentences</td>
<td>&gt; simple sentences</td>
</tr>
<tr>
<td>Caudate nucleus</td>
<td>R</td>
</tr>
<tr>
<td>Superior temporal sulcus</td>
<td>21/22</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>10</td>
</tr>
</tbody>
</table>

ANOVA, analysis of variance; BA, Brodmann area; L, left; R, right. I. See Talairach & Tournoix (1988).
that there were significantly stronger correlations for complex sentences than for simple sentences with responses in the right middle temporal gyrus (BA 21; $P < 0.05$), again in a region anterior and inferior to the site of the analogous differences in the control group (Table 2; Fig. 1).

**Differences in activations for simple sentences between groups**

Signal changes correlating with simple sentences were compared between groups ($P < 0.05$). Participants in the control group activated the left primary visual cortex and the thalamus more than the patients, whereas those in the patient group engaged the right cerebellar cortex more than the controls (Table 3).

**Differences in activations for complex sentences between groups**

Signal changes correlating with complex sentences were compared between groups ($P < 0.05$). The control group activated the right posterior middle temporal gyrus (BA 21) and the supramarginal gyrus (BA 40) more than the patients, whereas the patient group engaged the left fusiform/parahippocampal gyrus more than the controls (Table 4, Fig. 1).

**DISCUSSION**

In this study, participants spoke continuously about Rorschach inkblots while their regional brain activity was being measured with fMRI. In accordance with previous behavioural studies on syntax production in schizophrenia, we found that the patient group produced fewer complex sentences than controls (Thomas et al., 1996b; Oh et al., 2002). However, the numbers of simple sentences produced by the two groups were similar.

**Syntax production in the control group**

The rate of production of each sentence type was correlated with the BOLD response in each group. In members of the control group there was a positive correlation between the number of complex sentences and activation in the left posterior middle temporal gyrus, which is part of Wernicke’s area, and in the right superior temporal sulcus. Both these areas have consistently been activated in functional imaging studies of syntax comprehension in

### Table 2

Main foci of signal changes in participants with schizophrenia during production of continuous speech: correlations with simple and complex sentences are shown ($P < 0.001$), followed by statistical differences between the two conditions ($P < 0.05$)

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>Coordinates (mm)</th>
<th>Voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$x$</td>
<td>$y$</td>
</tr>
<tr>
<td>Simple sentences</td>
<td></td>
<td></td>
<td>-15</td>
<td>-57</td>
</tr>
<tr>
<td>Complex sentences</td>
<td></td>
<td></td>
<td>10</td>
<td>-53</td>
</tr>
<tr>
<td>Cerebellum</td>
<td></td>
<td>L</td>
<td>-35</td>
<td>-29</td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>21</td>
<td>R</td>
<td>54</td>
<td>-28</td>
</tr>
<tr>
<td>Precentral gyrus</td>
<td>6</td>
<td>L</td>
<td>-48</td>
<td>-40</td>
</tr>
<tr>
<td>Primary visual cortex</td>
<td>18/19</td>
<td>L</td>
<td>-31</td>
<td>-71</td>
</tr>
<tr>
<td>Inferior frontal gyrus</td>
<td>46</td>
<td>L</td>
<td>-40</td>
<td>37</td>
</tr>
<tr>
<td>Posterior insula</td>
<td></td>
<td>L</td>
<td>-26</td>
<td>-15</td>
</tr>
</tbody>
</table>

**ANOVA**

Simple sentences > complex sentences

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>Coordinates (mm)</th>
<th>Voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$x$</td>
<td>$y$</td>
</tr>
<tr>
<td>Cerebellum</td>
<td></td>
<td>L</td>
<td>-17</td>
<td>-50</td>
</tr>
</tbody>
</table>
| Complex sentences > simple sentences
| Postcentral gyrus              | 1/3 | R          | 49  | -14 | 37   | 9     |
| Primary visual cortex         | 19  | R          | -32 | -69 | 9    | 6     |
| Middle temporal gyrus         | 21  | R          | 52  | -33 | -7   | 4     |

ANOVA, analysis of variance; BA, Brodmann area; L, left; R, right.

**Fig. 1** Brain activation maps: functional magnetic resonance imaging signal changes correlated with the amount of syntactically complex sentences produced per 20 s epoch across two 3 min runs, showing significant differences in power of response between the patient and control groups ($P < 0.05$). Blue voxels indicate greater power of response in the control participants (right posterior middle temporal and left superior front gyrus, right nucleus accumbens); red voxels indicate greater power of response in the participants with schizophrenia (left parahippocampal gyrus). The left side of the brain is shown on the right side of the image; Talairach z coordinates (Talairach & Tournoux, 1988) are given below each slice.
Table 3 Differences in activation (P < 0.05) for simple sentences between patient and control groups

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>Coordinates (mm)</th>
<th>Voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>Simple sentences: controls &gt; patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary visual cortex</td>
<td>18/19</td>
<td>L</td>
<td>−29</td>
<td>−83</td>
</tr>
<tr>
<td>Thalamus</td>
<td></td>
<td>L</td>
<td>−3</td>
<td>−6</td>
</tr>
<tr>
<td>Simple sentences: patients &gt; controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebellum</td>
<td></td>
<td>R</td>
<td>6</td>
<td>−47</td>
</tr>
</tbody>
</table>

BA, Brodmann area; L, left; R, right.

Table 4 Differences in activation (P < 0.05) for complex sentences between patient and control groups

<table>
<thead>
<tr>
<th>Region</th>
<th>BA</th>
<th>Hemisphere</th>
<th>Coordinates (mm)</th>
<th>Voxels</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td>Complex sentences: controls &gt; patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posterior middle temporal gyrus</td>
<td>21</td>
<td>R</td>
<td>58</td>
<td>−50</td>
</tr>
<tr>
<td>Nucleus accumbens</td>
<td></td>
<td>R</td>
<td>9</td>
<td>6</td>
</tr>
<tr>
<td>Superior frontal gyrus</td>
<td>10</td>
<td>L</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Supramarginal gyrus</td>
<td>40</td>
<td>R</td>
<td>52</td>
<td>−44</td>
</tr>
<tr>
<td>Complex sentences: patients &gt; controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusiform/parahippocampal gyrus</td>
<td>35/36</td>
<td>L</td>
<td>−35</td>
<td>−25</td>
</tr>
</tbody>
</table>

BA, Brodmann area; L, left; R, right.

Healthy volunteers (Kaan & Swaab, 2002). However, whereas previous studies investigated syntax in the context of sentence comprehension, the neural correlates in our study were those of syntactic production during continuous speech. We thus interpret the signal changes in the left middle temporal gyrus and right superior temporal sulcus as being related to the production of syntactically complex sentences.

Syntax production in the patient group

In the patient group the production of complex sentences was also correlated with bilateral activation in the temporal cortex, but in regions anterior and inferior to those engaged in the controls. Direct comparison of the correlations in the two groups revealed significantly greater activation in the right posterior middle temporal gyrus, right supramarginal gyrus and left superior frontal gyrus in the control group compared with the patients. This weaker engagement of the posterior temporal cortex and adjacent inferior parietal lobule may be related to the impairment in producing complex sentences that we observed in the patients at the behavioural level. However, we cannot entirely exclude the possibility that these functional differences were related to semantic processing. Complex sentences are more complicated than simple sentences in terms of syntax as well as their semantic content. Reading sentences involves the temporal cortex (Kuperberg et al., 2000), and patients with schizophrenia show deficits in semantic processing at the behavioural level (Bagner et al., 2003). However, if semantic content were solely responsible for the temporal activations, we would also expect at least small signal changes in these areas for the correlations with simple sentences. An increased number of simple sentences per unit of time also conveys increased semantic content. We did not find such a correlation in either group and would therefore suggest that, in line with recent research (Kaan & Swaab, 2002), the temporal activations reflect syntactic complexity.

The impaired posterior temporal activation we observed in the patients is in line with other data demonstrating structural and functional abnormalities in this area in schizophrenia. Thus, the production of thought-disordered speech in another study of the same patients correlated negatively with signal changes in the left superior temporal gyrus (Kircher et al., 2001b), and lateralisation changes in the same areas have been described during a sentence completion task, also in the same sample (Kircher et al., 2001a). The temporal cortices have been implicated in the generation of auditory verbal hallucinations (Dierks et al., 1999; Shergill et al., 2000). Structural imaging studies (for review, see Shenton et al., 2001) have described volumetric anomalies in the lateral temporal gyr in schizophrenia, particularly in patients with formal thought disorder. Our findings are also consistent with abnormalities in the planum temporale during simple acoustic processing, where abnormalities in lateralisation patterns have been described (Rockstroh et al., 1998; Kircher et al., 2004).

Frontal lobe

Apart from the temporal cortex, the only other region that was significantly more correlated with complex rather than simple sentence production in the control group was in the left prefrontal cortex, close to the frontal pole (BA 10). This difference was not evident in the patient group and direct contrast of the correlation maps for complex sentences in the two groups revealed significantly greater activation in the controls than in the patients. Other studies have found activation in the proximity of this region during syntax comprehension in healthy volunteers (Newman et al., 2001). Although we predicted impaired differential engagement of the left inferior frontal gyrus rather than this more anterior region in the two groups, these observations suggest that this might also be related to impaired syntactic processing in people with schizophrenia.

It has been proposed that Broca’s area is involved in syntactic processing only when the demands on verbal working memory are high (Kaan & Swaab, 2002). Although the articulation of syntactically complex sentences might have engaged verbal working memory to a greater extent than simple sentences, the difference in the memory load might have been too small to lead to differential left inferior frontal activation. The absence of correlations between the production of complex sentences and the engagement of motor-related areas more generally may seem surprising, given that the participants were articulating speech; however, as they were speaking at
a rate that was self-paced and felt natural for them, the demands on articular processing might have been relatively constant. Activation in motor areas might thus have been at ceiling and not have shown a measurable response to fluctuations in the amount of speech produced in association with the production of complex sentences (Kircher et al., 2002).

Methodological considerations

Although overt speech can be associated with artefacts in fMRI studies secondary to articulation-related head movement (Bullmore et al., 1999), we quantified head movement during image acquisition, and found it to be minimal in both groups. Artefacts can also be introduced by air volume changes in the pharynx during phonation. However, when scanning at 1.5 T these effects are likely to be small, except in areas close to the orbital frontal cortex (Barch et al., 1999). Functional MRI is also associated with significant scanner noise, but all our participants reported that they were able to hear themselves speak during the task.

Concluding remarks

Our data implicate the left posterior superior temporal sulcus (Wernicke’s area), the right posterior middle temporal gyrus and the left prefrontal cortex in the production of syntactically complex sentences in healthy individuals. Patients with schizophrenia failed to show this pattern of activation, with significantly weaker engagement of the right temporal and left prefrontal cortex, which may underlie the reduced production of grammatically complex sentences in these patients. Language abnormalities seem to be the only symptom that is genetically transmitted among relatives of people with schizophrenia (Cardno et al., 2001). The way in which language, genes, the brain and schizophrenia relate to each other is a very intriguing question (Crow, 2000) and needs further investigation.

ACKNOWLEDGEMENTS

TK. was supported by the German Research Council.

REFERENCES


Nicotine dependence and symptoms in schizophrenia

Naturalistic study of complex interactions

M. CARMEN AGUILAR, MANUEL GURPEGUI, FRANCISCO J. DIAZ and JOSE DE LEON

Background Smoking may have a beneficial effect on either schizophrenic symptoms or antipsychotic side-effects, but studies are hampered by the lack of control of confounding factors.

Aims To explore the self-medication hypothesis in a large sample of stable outpatients with schizophrenia.

Method Symptoms, assessed with the Positive and Negative Syndrome Scale (PANSS), and number of hospitalisations were compared in 250 outpatients with DSM–IV schizophrenia classified into three categories: highly dependent smokers, mildly dependent smokers and non-smokers. Log-linear analysis was used to control for potential confounding and interacting variables.

Results High PANSS total scores and positive symptoms were less frequent in mildly dependent smokers than in non-smokers or highly dependent smokers. The highly dependent smokers had the worst outcome.

Conclusions The data do not generally support the self-medication hypothesis but rather suggest a complex interaction between nicotine dependence and schizophrenic symptoms.

Declaration of interest None. Funding detailed in Acknowledgements.

Schizophrenia is associated worldwide with a higher rate of smoking than that observed among the general population or those with other severe mental illnesses (McCreacle, 2002; Llerena et al, 2003). This association persists after correcting for such confounding factors as antipsychotic medication, institutionalisation or alcohol and drug use (de Leon et al, 1995, 2002a; Llerena et al, 2003). Smoking might be a marker of a more severe illness or might have a beneficial effect in schizophrenia by improving its symptoms and/or decreasing extrapyramidal side-effects of antipsychotics - 'the self-medication hypothesis' (Lohr & Flynn, 1992; Ziedonis et al, 1994; Dalack et al, 1998). This study explores both the self-medication hypothesis and the hypothesis that severe forms of schizophrenia are associated with high levels of nicotine dependence.

METHOD

Patients

The study was located at two community mental health centres and their rehabilitation unit, covering the catchment area of the city of Granada (southern Spain). All participants received free psychiatric treatment from the national health system. The sample included the first 250 consecutive patients with a diagnosis of DSM–IV schizophrenia who provided written informed consent after a complete description of the study (18 of 278 patients refused to participate; 10 additional excluded patients had a chart diagnosis of schizophrenia but did not meet the DSM–IV diagnosis). The diagnosis was made with the clinician version of a structured diagnostic interview (First et al, 1994). All 250 patients were Caucasians. The mean age was 36.1 (s.d. = 9.5) years and the mean age at diagnosis was 21.9 (s.d. = 6.0) years. There were 195 males (78%); this male overrepresentation is typical of treatment samples from many countries (Hambrecht et al, 1993), including Granada (Salize et al, 1999).

Twenty per cent had not completed their primary education; 45% had completed primary, 25% secondary and 10% a university education. Most patients (94%, 236/250) were taking antipsychotics, with a mean dose of chlorpromazine equivalents of 350 mg/day (s.d. = 459). The frequency of patients taking depot antipsychotics was 45% (113/250); risperidone, 33% (82/250); olanzapine, 6% (14250); and clozapine, 4% (10250). There is no reason to believe that the self-reported smoking of these patients was unreliable, because until recently smoking has been socially acceptable in Spain. Moreover, a reliable self-report of smoking or non-smoking status was provided by a subsample of 99 participants (of the 250 studied) whose cotinine in saliva was measured by radioimmunoassay.

Variables All ratings were conducted by a research psychiatrist (M.C.A.). Table 1 describes the variables used in statistical analyses. In order to avoid bias in the assessment, the clinical evaluation was conducted first, and information concerning medication and substance use, including tobacco and nicotine dependence, was gathered afterwards.

All variables were dichotomised except nicotine dependence, which was given three categories. On the basis of the Fagerstrom Test for Nicotine Dependence (FTND), smokers were classified as very highly dependent (FTND > 7; smoking a median of 40 cigarettes/day) or not very highly dependent (FTND ≤ 7; median of 20 cigarettes/day) (Fagerstrom et al, 1990). The three categories will be called highly dependent smokers, mildly dependent smokers and non-smokers. Schizophrenic symptomatology was assessed with the Spanish version of the Positive and Negative Syndrome Scale (PANSS; Peralta & Cuesta, 1994). The PANSS total scores were divided into high (≥ 45) and low scores. The negative, positive, disorganised, excited, anxious and depressed factors of the PANSS were calculated by adding the scores of the items with a loading higher than 0.50 in the factor (Peralta & Cuesta, 1994), and dividing by the number of those items. Subjects with a score ≥ 2 for a factor
were considered to have clinically significant symptoms (except for the excited factor, see Table 1 footnote). In summary, the presence of symptoms with regard to a factor (e.g. positive symptoms) in these clinically stable out-patients suggests that despite treatment they continue to have sufficient positive symptoms to be identified through a standardised assessment.

The Simpson & Angus (1970) Neurological Rating Scale was used to measure parkinsonian side-effects. Akathisia was assessed with the Barnes Akathisia Scale (Barnes, 1989). Some patients might have extrapyramidal side-effects previously corrected by antiparkinsonian drugs, therefore vulnerability to extrapyramidal side-effects was defined as the occurrence of at least one of the three following conditions: current treatment with antiparkinsonian medication; a score of >0 on the neurological rating scale; or a score of >0 on the Barnes Akathisia Scale (presence of akathisia).

A high antipsychotic dose was defined as a chlorpromazine equivalent of >10 mg/kg per day. Current alcohol and caffeine intake was assessed by interview and verified by chart review and collateral information from the family.

Table 1  Description of variables in 250 patients with schizophrenia, according to their nicotine dependence status: non-smokers (NS), mildly dependent smokers (MDS) and highly dependent smokers (HDS)

<table>
<thead>
<tr>
<th>Variable</th>
<th>250 patients</th>
<th>77 NS</th>
<th>92 MDS</th>
<th>81 HDS</th>
<th>( \chi^2 ) (d.f.=2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>High total score on the PANSS (≥ 45)</td>
<td>115 (46%)</td>
<td>40 (52%)</td>
<td>32 (35%)</td>
<td>43 (53%)</td>
<td>7.4**</td>
</tr>
<tr>
<td>Presence of symptoms1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>93 (37%)</td>
<td>31 (40%)</td>
<td>29 (32%)</td>
<td>33 (41%)</td>
<td>2.0</td>
</tr>
<tr>
<td>Positive</td>
<td>76 (30%)</td>
<td>24 (31%)</td>
<td>18 (20%)</td>
<td>34 (42%)</td>
<td>10.3**</td>
</tr>
<tr>
<td>Disorganised</td>
<td>34 (14%)</td>
<td>10 (13%)</td>
<td>5 (5%)</td>
<td>19 (24%)</td>
<td>11.9**</td>
</tr>
<tr>
<td>Excited</td>
<td>49 (20%)</td>
<td>11 (14%)</td>
<td>12 (13%)</td>
<td>26 (32%)</td>
<td>11.9**</td>
</tr>
<tr>
<td>Anxious</td>
<td>117 (47%)</td>
<td>39 (51%)</td>
<td>38 (41%)</td>
<td>40 (49%)</td>
<td>1.8</td>
</tr>
<tr>
<td>Depressive</td>
<td>78 (31%)</td>
<td>30 (39%)</td>
<td>23 (25%)</td>
<td>25 (31%)</td>
<td>3.8</td>
</tr>
<tr>
<td>Vulnerability to extrapyramidal symptoms</td>
<td>172 (69%)</td>
<td>54 (70%)</td>
<td>57 (62%)</td>
<td>61 (75%)</td>
<td>3.6</td>
</tr>
<tr>
<td>Presence of akathisia</td>
<td>35 (14%)</td>
<td>13 (17%)</td>
<td>9 (10%)</td>
<td>13 (16%)</td>
<td>2.2</td>
</tr>
<tr>
<td>High antipsychotic dose (≥ 10 mg/kg per day of chlorpromazine equivalents)</td>
<td>50 (20%)</td>
<td>12 (16%)</td>
<td>12 (13%)</td>
<td>26 (32%)</td>
<td>11.1**</td>
</tr>
<tr>
<td>Use of typical antipsychotic</td>
<td>177 (71%)</td>
<td>48 (62%)</td>
<td>59 (64%)</td>
<td>70 (86%)</td>
<td>14.2***</td>
</tr>
<tr>
<td>Male gender</td>
<td>195 (78%)</td>
<td>49 (64%)</td>
<td>79 (86%)</td>
<td>67 (83%)</td>
<td>13.6**</td>
</tr>
<tr>
<td>Caffeine intake</td>
<td>147 (59%)</td>
<td>27 (35%)</td>
<td>54 (59%)</td>
<td>66 (82%)</td>
<td>35.1***</td>
</tr>
<tr>
<td>Alcohol intake</td>
<td>52 (21%)</td>
<td>7 (9%)</td>
<td>16 (17%)</td>
<td>29 (36%)</td>
<td>18.1***</td>
</tr>
<tr>
<td>High number of hospital admissions</td>
<td>88 (35%)</td>
<td>18 (23%)</td>
<td>26 (28%)</td>
<td>44 (54%)</td>
<td>19.6***</td>
</tr>
</tbody>
</table>

1. Presence of symptoms is defined as factor scores > 2, except for the excited factor, which is > 1.

\*P < 0.05; **P < 0.01; ***P < 0.001.

Statistics

The Statistical Package for the Social Sciences (version 11.0) was used for calculations (SPSS, 1997). Initially, the three groups were compared by univariate parametric or non-parametric tests, as appropriate. Then, log-linear analyses of the data were performed (Agresti, 1990; SPSS, 1997); the log-linear analyses had two main purposes: they tested the hypothesis of a significant association of nicotine dependence with schizophrenic symptomatology, as measured by either the PANSS total score (or each one of its factors) or the number of hospitalisations; and they described the strength and direction of such association across different combinations of levels of potential interacting variables such as gender, antipsychotic dose/type and caffeine and alcohol intake. Strength and direction of associations were measured with odds ratios and their 95% confidence intervals from cross-tabulations.

In a first analysis, the association between nicotine dependence and PANSS total score was tested while controlling for gender, antipsychotic dose/type and caffeine and alcohol intake. This was performed by including the seven variables in a saturated log-linear model. Table 2 shows the significant interactions that were obtained. The significances of interaction were tested using partial \( \chi^2 \) (SPSS, 1997). A significant interaction between two variables was interpreted as evidence that the two variables were associated, even when controlling for the other variables in Table 2  Log-linear analysis focusing on the association between the Positive and Negative Syndrome Scale (PANSS) total score and nicotine dependence

<table>
<thead>
<tr>
<th>Variables</th>
<th>Significant interactions</th>
<th>d.f.</th>
<th>( \chi^2 )</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group A1</td>
<td>Nicotine dependence v. PANSS total score</td>
<td>2</td>
<td>7.4</td>
<td>0.03</td>
</tr>
<tr>
<td>Antipsychotic dose v. nicotine dependence</td>
<td>2</td>
<td>5.9</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Antipsychotic type v. nicotine dependence</td>
<td>2</td>
<td>8.1</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Group B2</td>
<td>Nicotine dependence v. gender</td>
<td>2</td>
<td>9.2</td>
<td>(&lt; 0.01)</td>
</tr>
<tr>
<td>Caffeine intake v. nicotine dependence</td>
<td>2</td>
<td>21.9</td>
<td>(&lt; 0.01)</td>
<td></td>
</tr>
<tr>
<td>Alcohol intake v. nicotine dependence</td>
<td>1</td>
<td>16.5</td>
<td>(&lt; 0.01)</td>
<td></td>
</tr>
<tr>
<td>Gender v. alcohol intake</td>
<td>1</td>
<td>23.4</td>
<td>(&lt; 0.01)</td>
<td></td>
</tr>
</tbody>
</table>

1. Group A: variables influencing the association between PANSS total score and nicotine dependence.
2. Group B: variables not influencing the association between PANSS total score and nicotine dependence.

Model including vulnerability to extrapyramidal side-effects: PANSS total score v. nicotine dependence \( \chi^2 = 6.5, \text{ d.f.}=1, P = 0.04; \) PANSS total score v. vulnerability to extrapyramidal side-effects \( \chi^2 = 8.7, \text{ d.f.}=1, P < 0.01; \) antipsychotic type v. vulnerability to extrapyramidal side-effects \( \chi^2 = 9.7, \text{ d.f.}=1, P < 0.01. \)
the model. Analyses similar to that of the PANSS total score were repeated for number of admissions (Table 3) and for negative, positive, disorganised, excited, anxious and depressed PANSS factors (results not presented).

A second purpose of the statistical analyses was to describe the association between nicotine dependence and schizophrenic symptomatology across different combinations of levels of other variables. One difficulty was that the relatively high number of variables considered in this study produced many possible combinations. Some variables represented a small sample size; for instance, the five dichotomous variables (gender, antipsychotic dose and type and caffeine and alcohol intake) produced $2^5 = 32$ possible combinations of levels. It is not practical or statistically advisable to perform so many cross-tabulations. Because odds ratios are rather inaccurate with small sample sizes, a systematic methodology that discarded irrelevant variables was used (SPSS, 1997). This methodology was based on the collapsibility conditions (Agresti, 1990); the rationale behind the collapsibility conditions is that variables that do not affect an association can be excluded from the analysis of that association (group B of variables in Tables 2 and 3), even if those variables have an effect on one of the variables involved in the association. The above analyses were repeated by including vulnerability to extrapyramidal side-effects as an additional variable.

### RESULTS

Among the 250 patients there were 173 (69%) current smokers and 77 (31%) non-smokers (including 7 (4%) former smokers and 70 (27%) that had never smoked daily). As expected, the rate of smoking was higher in our sample than among the Spanish general population (Pinilla & González, 2001) for both males (75% v. 45%) and females (49% v. 27%).

Table 1 shows the variable distribution among the 250 patients. The number of hospital admissions was 2.8 (median) of cotinine (ng/ml) in saliva in the 99 participants for whom it was determined were: 551 in highly dependent smokers (n=31), 423 in mildly dependent smokers (n=29) and 0.6 in non-smokers (n=29).

### Table 3 Log-linear analysis focusing on the association between corrected number of hospital admissions and nicotine dependence

<table>
<thead>
<tr>
<th>Variables</th>
<th>Significant interactions</th>
<th>d.f.</th>
<th>$\chi^2$</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corrected number of hospital admissions vs. corrected nicotine dependence</td>
<td>2</td>
<td>17.3</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>Gender vs. nicotine dependence</td>
<td>2</td>
<td>7.8</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Antipsychotic type vs. nicotine dependence</td>
<td>2</td>
<td>6.6</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Antipsychotic dose vs. nicotine dependence</td>
<td>2</td>
<td>5.9</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Coffee intake vs. nicotine dependence</td>
<td>2</td>
<td>20.6</td>
<td>&lt;0.01</td>
<td></td>
</tr>
<tr>
<td>Alcohol intake vs. nicotine dependence</td>
<td>2</td>
<td>5.8</td>
<td>0.05</td>
<td></td>
</tr>
<tr>
<td>Caffeine intake vs. alcohol intake</td>
<td>1</td>
<td>16.3</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Gender vs. alcohol intake</td>
<td>1</td>
<td>23.4</td>
<td>&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Antipsychotic dose vs. antipsychotic type</td>
<td>1</td>
<td>6.3</td>
<td>0.01</td>
<td></td>
</tr>
</tbody>
</table>

1. Group B: variables not influencing the association between corrected number of hospital admissions and nicotine dependence.

### Table 2 Symptom score and nicotine dependence

Table 2 shows results from the log-linear model that included nicotine dependence and the variables listed in the first column. The interaction between nicotine dependence and PANSS total score was significantly different from zero, indicating that these two variables were significantly associated when controlling for the other variables (gender, antipsychotic dose and type, caffeine and alcohol intake). Nicotine dependence was significantly associated with other variables (gender, antipsychotic dose and type, and caffeine intake). Two groups of variables can be identified from Table 2: group A, comprising PANSS total score and antipsychotic dose and type; and group B, comprising gender and caffeine and alcohol intake. Nicotine dependence was associated significantly with all the variables of group A and some variables of group B. However, the variables of group A were not associated with the variables of group B. In fact, no significant interactions simultaneously involving variables of A and variables of B were found. By virtue of the collapsibility conditions, the strength and direction of the association between PANSS total score and nicotine dependence do not vary across the levels of caffeine or alcohol intake, gender or across combinations of those levels. Thus, this association can be studied by controlling only for antipsychotic dose and type.

The association between PANSS total score and nicotine dependence was therefore studied with cross-tabulations for each of the four combinations of antipsychotic doses and types. The association was most significant among those on a low dose of typical antipsychotics (Fig. 1). In these subjects, mildly dependent smokers included the lowest number of subjects with clinically meaningful symptoms in the total PANSS. Among those on a low dose of a typical antipsychotic, non-smokers have an odds ratio of 2.7 of having a high PANSS total score when compared with mildly dependent smokers (Fig. 1). In other words, the percentage of patients with high total scores was significantly lower in mildly dependent smokers than among non-smokers or highly dependent smokers.

When vulnerability to extrapyramidal side-effects was also included in the log-linear model, the significant interactions were the same as in Table 2; additionally,
a significant interaction between antipsychotic type and vulnerability to extrapyramidal side-effects, and a significant interaction between PANSS total score and vulnerability to extrapyramidal side-effects were found (see footnote to Table 2). The association between PANSS total score and nicotine dependence was close to significant for those on a low dose of typical antipsychotic medication who showed vulnerability to extrapyramidal side-effects (Fig. 1); among these, mildly dependent smokers included the lowest number with clinically meaningful symptoms in the PANSS total score.

The analysis of the positive factor was very similar to the analysis of the PANSS total score and supported the self-medication hypothesis for those on a low dose of typical antipsychotics who are mildly dependent smokers.

Negative, depressive or anxious symptoms were not significantly associated with nicotine dependence (Table 1). Thus, the analyses of these symptoms did not support the self-medication hypothesis. Neither the disorganised nor the excited PANSS factor, after analysis, supported the self-medication hypothesis. Moreover, disorganised residual symptoms were associated with heavy smoking (see next section).

Number of admissions and nicotine dependence

Table 1 shows that highly dependent smokers had the highest proportion of hospital admissions compared with mildly dependent smokers (odds ratio=3.0) and non-smokers (odds ratio=3.9) (see also Fig. 2).

When the disorganised symptom variable was considered in the analysis, the association between nicotine dependence and number of admissions was significant only for those without disorganised symptoms (Fig. 2).

**DISCUSSION**

**The self-medication hypothesis of smoking in schizophrenia**

There is a discrepancy in the literature, with numerous animal studies suggesting that nicotine should help negative symptoms but scarce clinical data suggesting that this may be true in those with schizophrenia (Hughes, 2000). Two main sub-hypotheses are usually included in the self-medication hypothesis: smoking reduces the side-effects of antipsychotics; and nicotine may improve schizophrenic symptoms, particularly the negative, cognitive and/or depressive symptoms (Taiminen et al, 1998).

Two mechanisms have been implicated in the reduction of antipsychotic side-effects: a release of dopamine resulting from the administration of nicotine, a notion supported by both acute administration of nicotine in animal models (Drew et al, 2000) and in vivo human studies (Salokangas et al, 2000); and a decrease in antipsychotic blood levels through enzymatic induction. Individuals with schizophrenia who smoke tend to receive consistently higher doses of antipsychotics than non-smokers (Ziedonis et al, 1994; de Leon et al, 1995, 2002a). The inductive effect of smoking in antipsychotic metabolism therefore is inadvertently corrected by psychiatrists, because smokers...
tend to be treated with higher daily doses of antipsychotics than non-smokers. When compared with others with severe mental illness in three epidemiological studies in psychiatric hospitals, the effect of antipsychotic medication did not explain the association between schizophrenia and smoking (de Leon et al., 1995, 2002a; Llerena et al., 2003). Some cross-sectional studies have suggested that smoking reduces antipsychotic side-effects and others have not (Dalack et al., 1998); yet all of these studies are hampered by the lack of control for confounding factors. Longitudinal studies with small samples suggest that, when compared with atypical antipsychotics, typical antipsychotics are associated with increased smoking in some individuals (McEvoy et al., 1995) and with a greater difficulty for quitting smoking (George et al., 2000). Anticholinergic medication was not associated significantly with smoking in this or in previous studies (de Leon et al., 1995, 2002a,b).

In spite of the hypothesis from animal studies (Drew et al., 2000), very limited clinical data support an association between smoking and a reduction in negative symptoms (Dalack et al., 1998). Data indicating that nicotine may improve sensory gating abnormalities and smooth pursuit eye movements in schizophrenia or cognitive abnormalities induced by antipsychotics are somewhat stronger. Nicotine may have antidepressant qualities in individuals with depression (Salin-Pascual et al., 1996), but this is not well established in those with schizophrenia.

The literature appears to suggest that those with severe forms of schizophrenia may smoke more frequently, and more heavily, than those with less severe forms (Lohr & Flynn, 1992). The possible beneficial effect of nicotine (and smoking) on schizophrenic symptoms and antipsychotic side-effects may be obscured by this association between smoking and severe forms of schizophrenia. In summary, a critical reading of the literature lends very limited support to the self-medication hypothesis, but this effect may be obscured by the association between severe forms of schizophrenia and heavy smoking.

**Limitations and strengths of this study**
The limitations of the cross-sectional design make it impossible to prove definitively or to deny that smoking has beneficial effects on schizophrenia. It is not ethical to conduct long-term studies by (ideally) randomising patients to heavy or mild smoking. However, our study – involving a great number of stable out-patients with schizophrenia – like other naturalistic studies, may help to select which individuals are more likely to improve their schizophrenic symptoms and/or extrapyramidal side-effects using nicotine patches or other nicotine agonists. Because experimental designs with randomisation (to different levels of smoking and lack of smoking) are not admissible, other naturalistic studies with large samples, refined assessments and sophisticated statistical techniques to control for confounders are needed to confirm these findings.

Our findings suggest that nicotine dependence and schizophrenic symptomatology might be statistically dependent in out-patients with schizophrenia, but the data imply a complex interaction between these two variables. Our large sample size and the use of a sophisticated statistical technique, log-linear analysis, made it possible to control for potential confounding and interacting variables (Agresti, 1990). In contrast with other statistical techniques such as multiple linear or logistic regression, the log-linear methodology provides a clearer way to identify and deal with multiple interactions among several variables. In addition, as explained in the statistics section, the methodology used allows the systematic identification of variables affecting the association between nicotine dependence and schizophrenic symptoms, and the identification of subgroups where this association exists. Finally, the log-linear methodology does not imply the assumption of linear relationships among the variables analysed. The results of this study in stable out-patients with low levels of symptoms suggest that such an assumption is unsustainable. One apparent drawback of log-linear methodology is the need to transform continuous variables into categorical ones. Nonetheless, the amount of information that a log-linear analysis produces compensates for the possible loss of some information in the transformation (Agresti, 1990).

One may argue that the association between increased frequency of hospitalisation and high nicotine dependence may be partly explained by institutionalisation. However, this is not likely. Our cross-cultural studies suggest that, although the prevalence of current smoking in the general population is influenced by social pressure, high nicotine dependence among smokers with severe psychiatric illness appears to be similar across countries and remarkably resistant to social pressure (de Leon et al., 2002b).

**Support for the self-medication hypothesis in mildly dependent smokers**
The analysis of the PANSS total score, as well as the analysis of its positive symptom factor, supported the self-medication hypothesis for mildly dependent smokers, especially for those taking low doses of typical antipsychotics. Those with low levels of total symptoms are overrepresented among mildly dependent smokers compared with non-smokers. It may be that mild levels of smoking were associated with a reduction of schizophrenic symptoms, especially among those taking a low dose of typical antipsychotics (and particularly those showing vulnerability to extrapyramidal side-effects). The differential effect of typical and atypical antipsychotics on smoking behaviour is consistent with other studies with different designs suggesting that those with schizophrenia on typical antipsychotics may smoke more (McEvoy et al., 1995) and have more difficulties in quitting smoking (George et al., 2000).

The possible alleviation of positive symptoms by chronic nicotine administration could be explained by a potential correction of the cortical–subcortical dissociation of dopamine activity, which may be associated with schizophrenia (Dalack et al., 1998). However, in certain cases of schizophrenia (perhaps the most severe), ‘self-medication’, even with higher amounts of nicotine – as in our highly dependent smokers – would not be effective.

In summary, if there is any beneficial effect of nicotine it may be restricted to mildly dependent smokers, and particularly to those on low dosages of typical antipsychotics who are sensitive to the extrapyramidal side-effects. Such a benefit appears to affect only certain symptoms. Our study does not support the self-medication hypothesis for highly dependent smokers, who have poorer outcomes despite their heavy smoking.

**Other symptom differences do not support the self-medication hypothesis**
In contrast to the positive symptoms, the analyses of negative, anxious and
depressive symptoms in our sample do not support the self-medication hypothesis. This does not necessarily refute the hypothesis; the assessment may not have been sensitive enough or the effect size too small to be apparent with the statistical power in our sample. Tamnen et al. (1998) similarly found no differences in negative symptoms according to smoking behaviour. Although these two naturalistic studies do not rule out the possibility of beneficial effects of nicotine on negative symptoms, they certainly suggest that the alleviation of other schizophrenic symptoms is more likely. Ziedonis et al. (1994) described lower levels of negative symptoms in heavy smokers (and higher levels of positive symptoms), in comparison with light smokers and non-smokers with schizophrenia. Goff et al. (1992) found higher levels of both negative and positive symptoms in smokers than in non-smokers, whereas Kelly & McCreaddie (1999) were not able to demonstrate significant differences between smokers and non-smokers.

The presence of disorganised symptoms was associated with a high dependence on nicotine in our sample and did not support the self-medication hypothesis.

**Poor outcome and heavy smoking**

Our results suggest that severe forms of schizophrenia with poor outcome, manifested by either residual disorganised symptoms or a greater number of hospital admissions without residual disorganised symptoms, were associated with heavy smoking. Certainly if nicotine has some beneficial influence in schizophrenia, it is not evident in those with a poor outcome. There are no clear-cut theories to explain the association between highly dependent smoking and poor long-term outcome in schizophrenia. The most likely underlying reason is that these individuals may have vulnerability to both high nicotine dependence and schizophrenia with poor outcome.

**ACKNOWLEDGEMENTS**

The authors are grateful to patients and staff of the Granada-South and Granada-North Community Mental Health Centres and their Out-patient Rehabilitation Unit, for their collaboration, and to the Institutional Review Boards of the San Cecilio and Virgen de las Nieves University Hospitals, who approved this protocol. Juan Rivero helped with data collection. M.C.A’s work in this study was supported by a grant from the Spanish Agency for International Cooperation (AECI).

**REFERENCES**


**Escitalopram in the treatment of social anxiety disorder**

Randomised, placebo-controlled, flexible-dosage study

**SIEGFRIED KASPER, DAN J. STEIN, HENRIK LOFT and RICO NIL**

**Background** Selective serotonin reuptake inhibitors are effective in the treatment of social anxiety disorder and are currently regarded as the pharmacotherapy of choice.

**Aims** To investigate the efficacy and tolerability of escitalopram in the treatment of generalised social anxiety disorder.

**Method** Patients with generalised social anxiety disorder were randomised to receive placebo (n=177) or 10–20 mg escitalopram (n=181) in a 12-week, double-blind trial. The primary outcome measure was the mean change from baseline to last assessment in the Liebowitz Social Anxiety Scale (LSAS) total score.

**Results** The study showed a statistically superior therapeutic effect for escitalopram compared with placebo on the LSAS total score (P=0.005). There were significantly more responders to treatment for escitalopram than for placebo (54% v. 39%; P<0.01). The clinical relevance of these findings was supported by significant reduction in the work and social components of the Sheehan Disability Scale and by the good tolerability of escitalopram treatment.

**Conclusions** Escitalopram was efficacious and well tolerated in the treatment of generalised social anxiety disorder.

**Declaration of interest** The study was sponsored by H. Lundbeck A/S. Other funding detailed in Acknowledgements.

Social phobia or anxiety disorder is increasingly recognised as a highly prevalent and chronic disorder with onset during the teenage years (Lépine & Pelissolo, 1996; Wittchen et al, 1999). Although the disorder is associated with significant disability (including educational and occupational) which has a negative impact on quality of life, it is both underdiagnosed and undertreated (Kasper, 1998). Early work demonstrated that monoamine oxidase inhibitors (e.g. phenelzine) were effective in the treatment of the disorder, but these agents are limited by their side-effect profile, the need for dietary precautions, and drug interactions (Versiani, 2000). More recent work has established the efficacy of several selective serotonin reuptake inhibitors (SSRIs) (Stein et al, 1999; Van Ameringen et al, 2001; Liebowitz et al, 2002) and these agents have been recommended as first-line pharmacotherapy (Ballenger et al, 1998). This study investigates the efficacy and tolerability of escitalopram in the treatment of generalised social anxiety disorder.

**METHOD**

**Study design and dosing schedule**

This multinational study was a randomised, parallel group, placebo-controlled trial involving 41 centres in eight countries (Austria, Canada, Denmark, Finland, Germany, Norway, South Africa and the UK). Patients who met selection criteria entered a 1-week, single-blind, placebo lead-in period before being randomised to 12 weeks of double-blind treatment with escitalopram or matched placebo capsules. Patients were contacted for a safety follow-up 30 days after their last dose. The initial dosage of escitalopram was 10 mg per day. The dosage could be increased to 20 mg per day after 4, 6 or 8 weeks of treatment in case of an unsatisfactory response, judged as a score above 5 on the Clinical Global Impression scale rating for severity (CGI–S; Guy, 1976) or no decrease in CGI–S score since baseline. The mean daily dose of escitalopram was 17.6 mg at week 12. Efficacy and tolerability were assessed at baseline and after 1, 2, 3, 4, 6, 8 and 12 weeks of treatment.

**Patient population**

The patient population comprised female and male out-patients with a primary diagnosis of generalised social anxiety disorder established by means of a diagnostic interview following DSM–IV criteria (American Psychiatric Association, 1994), using the Mini-International Neuropsychiatric Interview (MINI; Sheehan et al, 1998) to assist in the exclusion of disallowed comorbidity. The patients were mainly recruited through advertisements. At the screening visit, patients 18–65 years old were selected if they had a total score of at least 70 on the Liebowitz Social Anxiety Scale (LSAS; Liebowitz, 1987) with exhibited fear or avoidance traits in at least four social situations, and were otherwise healthy based on a physical examination. Patients were excluded if they had another Axis I disorder that was considered the primary diagnosis within the previous 6 months, if the investigator diagnosed a serious risk of suicide or if the Montgomery–Åsberg Depression Rating Scale (MADRS; Montgomery & Åsberg, 1979) total score was higher than 19. Patients were also excluded if they had a DSM–IV diagnosis of alcohol or drug misuse during the past 6 months, or if they had taken a psychoactive drug (including any type of antidepressant, beta-blocker, benzodiazepine, narcotic, analgesic, antipsychotic or herbal remedy) within 2 weeks (5 weeks for fluoxetine and 6 months for depot neuroleptics) before screening, or if the patient had a positive urine drug screen for opiates, methadone, cocaine, amphetamines or benzodiazepines. The only allowed concomitant use of a psychotropic drug during the study was chloral hydrate taken as a hypnotic but not for more than three consecutive nights. Furthermore, patients with a diagnosis of mania or hypomania, body dysmorphic disorder, schizophrenia/other psychotic disorder, eating disorders, mental retardation or any Axis II cluster diagnosis were also excluded. Patients with a known drug (including citalopram) allergy or hypersensitivity or a known lack of therapeutic response to an adequate trial with
citalopram were also excluded. Patients participating in a formal psychotherapy programme that went beyond medical counselling were not included.

**Efficacy assessments**

The primary efficacy measure was the mean change from baseline to the last assessment (carried forward) of the LSAS total score. This scale consists of 24 items, 13 describing performance situations and 11 describing social interaction situations (Liebowitz, 1987). Each of the items is separately rated for ‘fear’ and ‘avoidance’ using a four-point categorical scale. All investigators attended supervised group sessions in order to standardise the interview and rating techniques. Secondary efficacy measures included:

(a) mean change from baseline to each visit in the LSAS sub-scale scores for ‘fear/ anxiety’ and ‘avoidance’;

(b) CGI–S score per visit and change from baseline to visit;

(c) Clinical Global Impression – Improvement (CGI–I) score: proportion of responders to treatment, defined as patients achieving a score of 1 (very much improved) or 2 (much improved) on the CGI–I;

(d) Sheehan Disability Scale (SDS; Sheehan, 1983) score, for the three domains ‘work’, ‘social’ and ‘family’;

(e) change from baseline to each visit in MADRS total score (the MADRS consists of ten items, each rated on a scale from 0 to 6).

**Safety and tolerability**

Safety assessments were based on vital signs (in a sitting position after 5 min rest), body weight, clinical laboratory tests (including haematology and biochemistry) and electrocardiograms (ECGs), and were assessed at the screening visit and at week 12. Adverse events observed by the investigator, spontaneously reported by the patient or reported in response to non-leading questions were recorded at each visit. The investigator documented the relationship to treatment, onset duration and intensity (mild, moderate or severe). All adverse events were coded using the included term according to the World Health Organization Adverse Reaction Terminology.

**Statistical analysis**

Efficacy analyses were based on the full analysis set (corresponding to the intent-to-treat population), which comprised all randomised patients who took double-blind study product and had at least one valid post-baseline assessment of the primary efficacy measure. Safety analyses were based on the set of all patients treated, which included all patients who took at least one dose of double-blind study product.

A minimum of 135 patients per treatment arm was required to reach a power of 90% to detect a significant difference between treatment groups in mean change from baseline to final assessment in LSAS total score at the 5% significance level. A general linear model for analysis of covariance (ANCOVA) was applied to the primary and secondary efficacy measures with factors for treatment group and centres (all centres with fewer than four patients were collapsed into one collective centre), and with baseline LSAS total score as a covariate. The final CGI–S and CGI–I scores were also analysed using the non-parametric Cochran–Mantel–Haenszel mean score statistics. Between-group comparisons of the proportion of responders (CGI–I score of 1 or 2) to treatment were performed using chi-squared and Fisher’s exact tests. Descriptive statistics were used for absolute values and mean changes from baseline in laboratory values, ECG parameters, vital signs and body weight. All statistical tests were two-sided and were carried out at the 5% level of significance.

**RESULTS**

**Patient baseline characteristics**

A total of 358 patients were randomised into the study, 177 to placebo treatment and 181 to escitalopram treatment. Of these, 5 patients did not receive double-blind treatment. The full analysis set thus consisted of 177 patients in the escitalopram group and 176 patients in the placebo group. A total of 290 patients (81%) completed the study, 145 in each treatment group (Fig. 1). There were slightly more men than women in both treatment groups. Baseline characteristics were similar for the two treatment groups with the exception of age and duration of the disorder, both of which were slightly higher in the escitalopram group (Table 1). No between-group difference was seen for the severity of the disorder, as measured by the baseline LSAS total score and the CGI score. There was no difference with respect to medical history or physiological variables. Comorbidity with depressive symptoms was low, as judged by the baseline MADRS total score and the low number of patients with a diagnosis of comorbid depression or dysthymia (Table 1). The high baseline LSAS total score and the baseline SDS score between 6 and 7 (on a ten-point scale) for the work and social life items are in line with the average CGI–S score, indicating a markedly ill patient population.

**Patient withdrawals**

A total of 68 patients (19%) withdrew from the study, with no overall between-group difference (18% in the placebo group and 20% in the escitalopram group). However, numerically more patients in the escitalopram group (8.8%) than in the placebo group (4.5%) withdrew because of adverse events and numerically more patients in the placebo group (6.2%) than in the escitalopram group (2.2%) withdrew because of lack of efficacy, with the latter difference approaching statistical significance ($P=0.059$).
Mean change from baseline in Liebowitz Social Anxiety Scale (LSAS) total score (last observation carried forward; LOCF) by week, for the escitalopram and placebo groups (full analysis set), adjusted for baseline score and centre by least squares mean analysis of covariance (**P<0.005) (Fig. 2).

A total of 54% of escitalopram-treated patients and 39% of placebo-treated patients responded to treatment (LOCF, P<0.01). The corresponding figures for the observed case (OC) analysis were 63% of escitalopram-treated patients and 43% of placebo-treated patients (P<0.001).

**Safety results**

Table 3 shows all treatment-emergent adverse events with an incidence of more than 5% in either treatment group. No clinically relevant trend was observed in mean ECG or in clinical laboratory parameters.

**DISCUSSION**

**Patient population**

The typical onset of social anxiety disorder during adolescence, with its chronic course, its high level of psychiatric comorbidity and its low spontaneous remission rate, contributes to serious impairment of daily functioning in the professional and social life of those with this disorder (Lépine & Pelissolo, 2000). These epidemiological characteristics were reflected among our participants. The mean age of onset was 15 years and the chronicity of the disorder was evident from its average duration, which was more than 20 years. Sheehan Disability Scale mean baseline scores for ‘work’ and ‘social’ items (around 7 on the ten-point sub-scales) indicate the negative impact of the disorder on daily life functioning in this group.

In order to investigate the specific efficacy of escitalopram in the treatment of social anxiety disorder, the study selected a somewhat atypical patient population with a low level of comorbidity. The average MADRS total score of 7.5 indicates the absence of significant depressive symptoms. It can thus be concluded that the patient population in this study represents patients with relatively pure, generalised social anxiety disorder. The average LSAS total score at baseline of over 95 indicates a more severely ill patient population than that in other published clinical drug trials (Baldwin et al, 1999; Liebowitz et al, 2002).
et al. al.

With that reported in other studies of SSRIs, showed a decrease in the total LSAS score and 'avoidance'. The primary analysis of the secondary efficacy measures, including the week trial period on both the primary and the secondary response rate in the placebo groups with increasing size of trial, as found by Oosterbaan et al. (2001), is consistent with the substantial size of our trial.

Irrespective of the placebo response size, the clinical significance of the escitalopram treatment effects in this study was demonstrated by statistically significant effects on the global measures of severity of illness and improvement (CGI–S and CGI–I) and, importantly, also on the two Sheehan Disability Scale items 'work' and 'social'. A final score on the CGI–I scale of 1 or 2 (very much or much improved) has commonly been used as a response criterion in social anxiety disorder pharmacotherapy trials. In this trial, the escitalopram response rate (OC) was 63% compared with 43% in the placebo group. Again, the magnitude of response of the escitalopram-treated patients is consistent with that reported in other studies, whereas the placebo response rate is higher than that found previously (Liebowitz et al., 2002).

**Withdrawals**

The total withdrawal rate of 19% is clearly lower than that in a recently reported fixed-dose study with paroxetine (Liebowitz et al., 2002) and somewhat lower than the average rate of 23% based on the 15 studies reviewed by Oosterbaan et al. (2001). The latter review further reported a positive relation between withdrawal rate and the size of the trials. The withdrawal rates varied slightly between treatment groups in our study, with borderline statistical significance for the higher rate of withdrawals due to lack of efficacy in the placebo group, and a somewhat higher withdrawal rate due to adverse events in the escitalopram group.

**Tolerability**

Escitalopram was well tolerated in this study, with prevalence rates of single adverse symptoms comparable with those in studies of its use in depression (Wade et al., 2002). A favourable tolerability profile is important in the pharmacotherapy of this chronic disease, for which

---

**Table 2** Adjusted mean change from baseline to week 12 in LSAS and SDS scores and response rate

<table>
<thead>
<tr>
<th></th>
<th>Placebo group (n=176)</th>
<th>Escitalopram group (n=177)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adjusted mean change in score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LSAS2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>27.2</td>
<td>34.5**</td>
</tr>
<tr>
<td>Fear/anxiety sub-scale</td>
<td>12.7</td>
<td>16.9***</td>
</tr>
<tr>
<td>Avoidance sub-scale</td>
<td>14.4</td>
<td>17.6*</td>
</tr>
<tr>
<td>SDS2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Work sub-scale</td>
<td>2.03</td>
<td>2.90***</td>
</tr>
<tr>
<td>Social sub-scale</td>
<td>2.12</td>
<td>2.70*</td>
</tr>
<tr>
<td>Family sub-scale</td>
<td>1.55</td>
<td>1.59</td>
</tr>
<tr>
<td>Responders to treatment (%)1</td>
<td>39</td>
<td>54**</td>
</tr>
</tbody>
</table>

LSAS, Liebowitz Social Anxiety Scale; SDS, Sheehan Disability Scale.

1. Full analysis set, last observation carried forward.
2. Participants achieving a score of 1 or 2 on the Clinical Global Impression – Improvement scale.

---

**Therapeutic efficacy and placebo response**

This study of the SSRI escitalopram confirmed the efficacy of SSRIs in the treatment of generalised social anxiety disorder. Escitalopram had a significantly better effect than placebo at the end of the 12-week trial period on both the primary and the secondary efficacy measures, including the two LSAS sub-scales of ‘fear/anxiety’ and ‘avoidance’. The primary analysis showed a decrease in the total LSAS score of 34.4 points in the escitalopram group and a relatively large decrease of 27.2 points in the placebo group. The effect size in the escitalopram group is comparable with that reported in other studies of SSRIs in the treatment of generalised social anxiety (Stein et al., 1999; Baldwin et al., 1999). However, no other published study has reported a placebo response as high as 39% (LOCF) in social anxiety disorder. A review by Oosterbaan et al. (2001) analysed 15 placebo-controlled studies and concluded that a moderate placebo response is seen in this disorder which appears to be lower than that in depression or panic disorder. The review found no evidence of an increase in the placebo response in studies of social anxiety over the past decade, although this is seen for other disorders. There was, however, a trend towards a higher response rate in the placebo groups, but not in the active treatment groups, with increasing sample size. No relation was found between the baseline severity of social anxiety disorder and improvement during treatment, as measured by the mean change from baseline or the percentage of responders. This is somewhat in contrast to other studies of this disorder, in which placebo responders were generally less symptomatic (Montgomery, 1998) and where a better separation between active medication and placebo was seen among the more severely affected patients. The trend towards a higher response rate in the placebo groups with increasing size of trial, as found by Oosterbaan et al. (2001), is consistent with the substantial size of our trial.

---

**Table 3** Treatment-emergent adverse events with an incidence greater than 5%

<table>
<thead>
<tr>
<th></th>
<th>Placebo group (n=177) n (%)</th>
<th>Escitalopram group (n=181) n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headache</td>
<td>45 (25)</td>
<td>46 (25)</td>
</tr>
<tr>
<td>Nausea</td>
<td>21 (12)</td>
<td>39 (22)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>15 (9)</td>
<td>25 (14)</td>
</tr>
<tr>
<td>Somnolence</td>
<td>9 (5)</td>
<td>18 (10)</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>8 (5)</td>
<td>17 (9)</td>
</tr>
<tr>
<td>Insomnia</td>
<td>11 (6)</td>
<td>17 (9)</td>
</tr>
<tr>
<td>Dizziness</td>
<td>9 (5)</td>
<td>13 (7)</td>
</tr>
<tr>
<td>Rhinitis</td>
<td>9 (5)</td>
<td>13 (7)</td>
</tr>
<tr>
<td>Sweating increased</td>
<td>3 (2)</td>
<td>11 (6)</td>
</tr>
<tr>
<td>Ejaculation failure (men)</td>
<td>0 (0)</td>
<td>6 (6)</td>
</tr>
<tr>
<td>Libido decreased</td>
<td>2 (1)</td>
<td>10 (6)</td>
</tr>
</tbody>
</table>
lengthy treatment may be required. Headache was the adverse event with the highest incidence, and its incidence was similar in the two treatment groups. Nausea, increased sweating and sexual side-effects occurred with a higher incidence during escitalopram treatment. Also in agreement with depression studies, no clinically relevant mean change in ECG variables was seen in the escitalopram group.

ACKNOWLEDGEMENTS

The authors acknowledge the valuable contributions of all participating investigators in this study. Austria: Siegfried Kasper; Hans Georg Zapotoczky, Anton Tolk, Albert Wachszitz; Canada: Michael Van Ameringen; Jacques Flamandon, Kevin Dwight Kjernisted, Pratap Rao Chokka, Stanley P. Kutcher, Martin A. Katzman; Denmark: Henrik Lublin, Eva Jensen, Annemarie Lund Laurusen, Peter Kaare Østergaard, Flemming Bjornal, Pale Stensbæk Schrøder; Finland: Hannu Juhan Kopenen, Anna Elisa Savela, Anneli Timonen, Ulla Marjatta Lepola, Ritva Helena Jokinen, Markus Pulkkinnen; Germany: Dolores Backhaus, Johannes Böhninger, Karl Ditzler, P. Donat, Buckhard Flatoatto, Hinrich Hörlein-Rummel, Gerhard Roth, Christoph Schenk; Norway: Alv A. Dahl, Fred Holsten; South Africa: Dan J. Stein, William J. C. Verbeeck, Catherine Mary Maud, Dora Wynchank, Donald A. B. Wilson, Premakanthie R. Laban, Wally Lundberg, Chane Magnus; UK: Alan Wade.

R.N. and H.L. are full-time employees of H. Lundbeck AS. S.K. has received grants or research support from Eli Lilly, Lundbeck, Bristol-Myers Squibb, GlaxoSmithKline, Organon and Servier. He works or has worked as a consultant or on an advisory board for AstraZeneca, Bristol-Myers Squibb, GlaxoSmithKline, Eli Lilly, Lundbeck, Pfizer, Organon, Janssen Pharmaceutica and Novartis. His speaker’s bureaux include AstraZeneca, Eli Lilly Lundbeck and Janssen Pharmaceutica. D.S. has received research grants and/or consultancy honoraria from AstraZeneca, Eli Lilly, GlaxoSmithKline, Lundbeck, Orion, Pfizer, Pharmacia, Roche, Servier, Solvay, Sunitomoto and Wyyeth.

REFERENCES


Siegfried Kasper, MD, FRCP, Department of General Psychiatry, University of Vienna, Vienna, Austria; Dan J. Stein, PhD, FRCP, University of Stellenbosch, Cape Town, South Africa and University of Florida, Gainesville, Florida, USA; HENRIK LOFT, PhD, H. Lundbeck A/S, Copenhagen, Denmark; RICO NIL, PD, Dr.sc.nat., ETH, Lundbeck (Switzerland) Ltd, Glattbrugg, Switzerland

Correspondence: Professor Siegfried Kasper, Department of General Psychiatry, University of Vienna, Währinger Gürtel 18–20, 1090 Vienna, Austria. Tel: +43 1 40 400 3568; fax: +43 1 40 400 3099; e-mail: sc-genpsy@meduniwien.ac.at.

(First received 21 January 2004, final revision 16 September 2004, accepted 29 September, 2004)

CLINICAL IMPLICATIONS

- Escitalopram is efficacious in the treatment of generalised social anxiety disorder.
- The tolerability of escitalopram in the dosage range 10–20 mg was as favourable as that previously seen in the treatment of depression.
- The efficacy/tolerability profile of escitalopram and the low withdrawal rate in this study make escitalopram a valuable pharmacotherapeutic option in the treatment of patients with social anxiety disorder.

LIMITATIONS

- A high placebo response rate was found in this study.
- Given the chronic course of the disease, studies with a longer treatment duration may be warranted to assess further potential improvements.
- The patient sample was selected to minimise comorbidity with other psychiatric disorders in order to investigate effects of escitalopram specifically in social anxiety disorders.


Assessing adolescent personality pathology

DREW WESTEN, LISSA DUTRA and JONATHAN SHEDLER

Background Personality pathology constitutes a major form of psychopathology in adolescents.

Aims To examine the reliability and validity of a Q-sort instrument for assessing adolescent personality pathology designed for clinically experienced informants.

Method A sample of 294 randomly selected psychiatrists and psychologists each provided data on a current patient, aged 14–18 years. Clinicians completed several measures, including the Shedler–Westen Assessment Procedure for Adolescents (SWAP–200–A).

Results Factor analysis identified 11 dimensions of adolescent personality: psychopathy/malignant narcissism, dysphoria/inhibition, psychological health, histrionic sexualisation, schizotypy, sexual conflict, emotional dysregulation, anxious obsessivity, peer rejection, delinquent behaviour and attentional dysregulation. These correlated in predicted ways with a range of criterion variables, including measures of adaptive functioning, Axis II pathology, the Five Factor Model and the Child Behavior Checklist.

Conclusions The SWAP–200–A shows promise as an instrument for assessing personality pathology in adolescents. Trait dimensions such as delinquent behaviour and emotional dysregulation may prove useful additions to a classification of personality.

Declaration of interest None. Funding detailed in Acknowledgements.

A growing body of research suggests that personality pathology constitutes a major form of psychopathology in adolescents, as it does in adults (Ludolph et al., 1990; Johnson et al., 1999; Levy et al., 1999; Westen & Chang, 2000). At present there are no widely accepted assessment procedures or diagnoses designed specifically for adolescent personality pathology. Whereas previous research on personality disorder features in adolescents, including our own earlier study (Westen et al., 2003), has focused on types (Axis II categories) or prototypes (personality constellations assessed dimensionally, such as the extent to which the patient matches a prototype of histrionic personality disorder), this study focuses on traits: more specific, less molar constructs such as negative affect or emotional dysregulation. Whether personality pathology in adolescents (or adults) is best considered in terms of traits, types or prototypes, or some combination thereof, is an important question. Our aim is to identify the structure of pathological personality traits in a sample of adolescents described by their treating clinicians using a personality pathology instrument designed for adolescents, and to provide initial data on the validity of these constructs and their measurement.

METHOD

Sample and procedure

The study method has been described in detail elsewhere (Westen et al., 2003). Participants were randomly selected PhD and MD clinicians, who provided quantified data on a randomly selected adolescent patient (operationalised as ‘the last patient you saw last week before completing this form who meets study criteria’) currently in treatment for ‘enduring maladaptive patterns of thought, feeling, motivation, or behaviour – that is, personality’. We used this relatively generic definition of personality pathology to avoid limiting our study to patients meeting diagnostic criteria for personality disorder. We ascertained a sample stratified by gender and age (14–18 years) and asked clinicians to complete a packet of questionnaires for a modest honorarium (US$25).

Measures SWAP–200–A

The 200-item Shedler–Westen Assessment Procedure for Adolescents (SWAP–200–A) is a Q-sort instrument for assessing adolescent personality pathology designed for use by skilled clinical observers based on either longitudinal knowledge of the patient over the course of treatment or a systematic clinical interview of the patient and parents. A Q-sort is a set of statements that provides a ‘standard vocabulary’ for clinicians to use to describe their clinical observations. To describe a patient, the clinician sorts statements into categories based on their applicability to the patient, from those that are irrelevant or not descriptive to those that are highly descriptive. In this study clinicians used a semi-constrained rating scale version of the instrument (see Westen et al., 2003).

The SWAP–200–A was adapted from the SWAP–200 for adults; both measures have shown initial evidence of reliability and validity (Westen & Shedler, 1999a,b, 2000; Westen & Muderrisoglu, 2003). Preliminary research has shown high correlations between SWAP–200 descriptions made by treating clinicians and independent interviewers and between independent observers reviewing recorded interviews (Westen & Muderrisoglu, 2003). The SWAP–200–A correlates with a range of variables such as attachment status, and history of suicide attempts, psychiatric hospitalisations, arrests, and family and developmental history variables (Nakash-Eisikovits et al., 2003; Westen et al., 2003).

The items reflect constructs from a mixture of sources: Axis II criteria for DSM–III through DSM–IV; selected Axis I criteria associated with personality disturbance (e.g. depression and anxiety); clinical literature and research on personality disorders, normal personality traits and psychological health; a model of functional diagnosis (Westen, 1998); research on child and adolescent personality and psychopathology; videotaped clinical interviews; and feedback from over a thousand experienced clinicians. To develop the adolescent
version of the instrument, we deleted, revised and added items as appropriate based on the adolescent literature, the authors’ prior research and experience with adolescent personality pathology, and consultation with senior clinicians in adolescent psychiatry who used the instrument to describe patients and then provided feedback on items that were ambiguous, necessary for describing their patient but missing from the item set, and so on.

Two features of the SWAP–200–A are of particular relevance to assessment of adolescent personality pathology. First, the instrument is intended for use by clinically experienced observers, based on either all available data over the course of their work with a patient or a systematic clinical interview with the patient and parents, the Clinical Diagnostic Interview for Adolescents (CDI–A; further details available from the author upon request). The SWAP–200–A does not presume that patients, particularly adolescents, can self-report their maladaptive personality traits. Rather, it presumes that a skilled clinical interviewer can listen to patients’ narratives, observe their interactions with their parents and the interviewer, and integrate information across informants to make judgements about adolescents’ characteristic ways of thinking, feeling, regulating affect and so forth. Second, the instrument can be used to assess Axis II diagnoses in adolescents, by correlating patients’ 200-item profiles with diagnostic prototypes of each personality disorder derived from a normative adult sample (Westen & Shedler, 1999a). Alternatively, it can be used in taxonomic work, as in this study, to develop non-redundant (i.e. relatively non-comorbid) diagnostic categories or dimensions.

Clinical Data Form
Following basic demographic and diagnostic questions, the Clinical Data Form (CDF) (see Westen & Shedler, 1999a) asks clinicians to rate the patient’s adaptive functioning, including school functioning (1 severe conduct problems/suspensions, 7 working to potential); peer functioning (1 very poor, 7 very good); history of suicide attempts, arrests and hospitalisations; and social support (number of people in whom the patient feels comfortable confiding). Research has demonstrated that clinician ratings of adaptive functioning variables, including the variables assessed by the CDF, show strong correlations with the same variables obtained by interview (see Westen & Weinberger, 2004). The CDF also measures family and developmental history variables; however, because we do not analyse those data here, we shall not describe them further.

Axis II pathology
To maximise reliability of measurement, we assessed Axis II pathology as defined by DSM–IV (American Psychiatric Association, 1994) in more than one way. First, we listed the Axis II disorders and asked clinicians to rate the extent to which the patient met criteria for each disorder on a seven-point rating scale. Second, we provided clinicians with a checklist of all Axis II criteria, randomly ordered, and asked them to rate each criterion as present or absent, as in DSM–IV, and then to rate the extent to which each item applied using a seven-point scale. These checklist data generated two additional dimensional measures of Axis II pathology (number of diagnostic criteria met for each personality disorder and the mean of the ratings 1–7 for each criterion for each diagnosis), as well as categorical diagnoses derived by summing the number of criteria present and applying DSM–IV thresholds. To maximise reliability, we created a composite measure of personality pathology by transforming the three sets of dimensional scores (global ratings, number of Axis II criteria met and summed seven-point ratings across criteria for each disorder) into Z scores, which we then averaged to form composite personality disorder ratings.

Free Factor Model adjective checklist
The Five Factor Model (FFM; McCrae & Costa, 1997) is a model of personality derived by factor analysis. It isolates five general personality traits: neuroticism, extraversion, openness to experience, agreeableness and conscientiousness. According to the most widespread version of the model, embodied in the NEO Personality Inventory Revised (NEO–PI–R; McCrae & Costa, 1997), each factor includes six sub-factors or ‘facets’. For this study we developed a brief clinician-report FFM adjective checklist, consisting of 35 items rated on a seven-point scale, one for each of the NEO–PI–R factors and one for each of the six facets. Coefficient alphas for the five NEO–PI–R factors were largely acceptable, ranging from 0.64 to 0.92, with both mean and median greater than 0.80.

Child Behavior Checklist
The Child Behavior Checklist (CBCL; Achenbach, 1991) is a widely used questionnaire designed to assess the behavioural problems and social competencies of children aged 4–18 years; it includes 11 problem scales. The CBCL also yields two broadband, higher-order psychopathology scales, ‘internalising’ and ‘externalising’. We asked clinicians to complete the parent-report version of the CBCL, which they were able to do without difficulty. Clinician-reported data on the CBCL show similar psychometric properties to parent-reported data, including high internal consistency for the problem scale scores (median coefficient $\geq 0.80$), virtually identical factor structure and predictable correlates suggesting convergent and discriminant validity (Dutra et al, 2004).

Statistical analyses
We analysed the data as follows. First, we subjected the SWAP–200–A items to exploratory factor analysis (because of the absence of prior research on the factor structure of the instrument). As a preliminary test of the validity of the factors, we then performed a series of analyses. (For simplicity of presentation, in the tables that follow, we indicate criterion variables predicted a priori to be most strongly associated with each factor in bold. To minimise overinterpretation of findings, we focus only on findings that are relevant to our hypotheses, form a coherent pattern, or were not predicted but were significant at $P<0.01$.) In a first set of analyses, we correlated patients’ factor-based scores with dimensional measures of personality pathology, to locate them within a nomological net (Cronbach & Meehl, 1955) provided by the more familiar DSM–IV Axis I diagnoses. We then examined their relation to personality as measured by the FFM adjective checklist and the problem scales of the CBCL. Finally, we assessed the relation between SWAP–200–A factors and adaptive functioning variables selected a priori as likely to be associated with different forms of personality pathology, including ratings of school performance and quality of peer relationships; number of close friends or confidantes; and history of psychiatric hospitalisations, suicide attempts and arrests.
RESULTS

Sample characteristics

The sample consisted of 294 patients. Clinician respondents (61.4% psychiatrists, 50.2% male) were on average highly experienced (mean years of experience post training 13.4, s.d.=9.4). Clinicians varied in theoretical orientations, and most worked in more than one setting. Clinicians tended to know the patients well: the median length of treatment prior to completing the questionnaire was 20 sessions. Patients were evenly distributed by gender (52.9% female) and age. The majority (84.9%) were White, with most of the remaining patients Black or Hispanic. Clinicians rated the patients as 7.5% poor, 20.9% working class, 50.7% middle class and 20.9% upper class. The most prevalent Axis I diagnoses included major depressive disorder (25.3%), dysthymic disorder (24.3%), attention-deficit hyperactivity disorder (16.1%), oppositional defiant disorder (9.0%) and conduct disorder (6.1%).

Factor structure of the SWAP–200–A

As a first step, we subjected the SWAP–200–A items to a principal components analysis, examining the resulting eigenvalues, percentage of variance accounted for by each factor, and scree plot. The scree plot showed a gradual break between 12 and 15 factors. We obtained similar factor structures using 12–15 factors with both varimax (orthogonal) and promax (oblique) solutions and multiple extraction methods. We retained and report here the first 11 of the 12-factor promax (oblique) solution using principal axis factoring. These factors were readily interpretable, reproducible across several estimation procedures and algorithms, and well marked by multiple items. Commonalities were all greater than 0.70, with most between 0.80 and 0.90, suggesting that the items did in fact include substantial common components. The 11 factors cumulatively accounted for 52% of the variance. Reliability (coefficient α) was above 0.80 for all factors except factor 10 (α=0.72), with a median of 0.86. Table 1 lists the items that loaded most highly on each of the 11 factors.

The first factor, psychopathy/malignant narcissism, includes items associated with narcissism, disruptive behaviour disorders, hostility, antisocial personality disorder and psychopathy. The second factor, dysphoria/inhibition, includes depression, anhedonia, shame, guilt and a number of related cognitive and emotional processes. It also includes a tendency to be inhibited in a number of domains, including social, motivational and affective inhibition. The third factor, psychological health, reflects the presence of psychological strengths and inner resources, including the capacity to love, find meaning in life experiences and gain insight into the self. Factor 4, histrionic sexualisation, reflects sexuality typically seen in histrionic personality disorder. Patients scoring high in this dimension tend to be sexually provocative and promiscuous; they also have a tendency to fantasise about ideal love but become involved in emotionally charged, unhealthy romantic relationships. The fifth factor, schizotypy, describes patients with subclinical positive and negative symptoms. Patients scoring high on this dimension tend to have barren representations of themselves and others, impoverished thought more generally and emotional flatness or constriction. They also tend to have odd appearances, mannerisms, reasoning processes and/or perceptual experiences. Factor 6, emotional dysregulation, includes a deficiency in the capacity to modulate and regulate affect, so that in such individuals affect tends to spiral out of control, change rapidly, be expressed in intense and unmodified form, and overwhelm reasoning. This emotional dysregulation may lead to self-destructive attempts to regulate affects, such as suicidality and parasuicidality, self-injury and/or binging and purging. The construct of emotional dysregulation is central to contemporary clinical thought, especially with respect to borderline personality disorder (see Westen, 1991; Linehan, 1993; Westen et al, 1997). Of note is its statistical independence from our two negative affect factors tapping dysphoria and anxiety. Factor 7, anxious obsessiosity, reflects highly anxious individuals who may experience obsessions, compulsions, phobias and/or panic attacks; these patients tend to develop somatic problems in response to stress. Factor 8, delinquent behaviour, reflects a tendency to engage in criminal behaviour, misuse drugs and alcohol, run away from home, seek out thrills and adventure and surround oneself with delinquent peers. Factor 9, sexual conflict, is descriptive of patients who are confused about their sexual orientation and appear to be struggling with counter-normative sexual desires and mannerisms. Factor 10, attentional dysregulation, describes a constellation of personality processes associated with attentional deficits, including low tolerance of frustration and irresponsibility. Factor 11, peer rejection, describes adolescents who have poor social skills and tend to be neglected, avoided or bullied by their peers. Adolescents with high scores on this factor tend to lack close friendships and relationships.

Assessing validity

The factors that emerged are clinically and theoretically coherent, and most resemble factors that emerged from factor analysis of the SWAP–200 Q-sort for adults (Shedler & Westen, 2004). To assess the validity of these factors (their association with external criteria), we first examined their relation to dimensional personality disorder diagnoses. As can be seen from Table 2, the SWAP–200–A factors tended to be associated with theoretically relevant variables in predictable ways (predicted correlations are in bold). For example, schizotypy and peer rejection were both strongly associated with personality disorders involving social isolation and peculiarity. The emotionally dysregulated and histrionic sexualisation factors were both associated with borderline, histrionic and dependent personality disorders – three disorders that tend to demonstrate significant diagnostic overlap in adult samples. Also of note is that the psychological health factor was negatively associated with most of the personality disorders.

Our next set of analyses examined the association between the SWAP–200–A factors and the FFM checklist (Table 3) and the CBCL (Table 4). As predicted, SWAP–200–A factors involving negative emotions (notably dysphoria/inhibition, anxious obsessiosity and emotional dysregulation) were strongly associated with neuroticism, whereas factors involving externalising pathology (particularly malignant narcissism) tended to correlate negatively with agreeableness and conscientiousness. Of note was the negative correlation between schizotypy and openness to experience, which makes theoretical sense in light of the concreteness and affective detachment of patients rated high on this factor, but does not accord with predictions of FFM researchers who have tried to account for schizotypal thinking as extreme openness to experience (e.g. Widiger et al,
<table>
<thead>
<tr>
<th>Factor</th>
<th>Psychopathy/malignant narcissism</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tends to be critical of others</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Tends to be angry or hostile (whether consciously or unconsciously)</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td>Tends to hold grudges; may dwell on insults or slights for long periods</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Tends to get into power struggles with adults</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Appears to gain pleasure or satisfaction by being sadistic, aggressive or bullying (whether consciously or unconsciously)</td>
<td>0.75</td>
<td></td>
</tr>
<tr>
<td>Tends to react to criticism with feelings of rage or humiliation</td>
<td>0.73</td>
<td></td>
</tr>
<tr>
<td>Tends to blame others for own failures or shortcomings; tends to believe his/her problems are caused by external factors</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>Tends to see own unacceptable feelings or impulses in other people instead of in him/herself</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>Tends to be rebellious or defiant toward authority figures</td>
<td>0.72</td>
<td></td>
</tr>
<tr>
<td>Manages to elicit in others feelings similar to those he or she is experiencing (e.g. when angry, acts in such a way as to provoke anger in others; when anxious, acts in such a way as to induce anxiety in others)</td>
<td>0.70</td>
<td></td>
</tr>
<tr>
<td>Tends to express intense and inappropriate anger, out of proportion to the situation at hand</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Tends to be oppositional, contrary or quick to disagree</td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td>Tends to be controlling</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>Takes advantage of others; is out for number one; has minimal investment in moral values</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>Tends to show reckless disregard for the rights, property or safety of others</td>
<td>0.67</td>
<td></td>
</tr>
<tr>
<td>Draws pleasure or self-esteem from being, or being seen as, ‘bad’ or ‘tough’</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>Tends to see certain others as ‘all bad’, and loses the capacity to perceive any positive qualities the other person may have</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>Is quick to assume that others wish to harm or take advantage of him/her; tends to perceive megalomaniacal intentions in others' words and actions</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Tends to be arrogant, haughty or dismissive</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Tends to be competitive with others (whether consciously or unconsciously)</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Tends to seek power or influence with peers (whether in beneficial or destructive ways)</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Seeks to dominate an important other (e.g. sibling, parent, boyfriend, girlfriend) through violence or intimidation</td>
<td>0.64</td>
<td></td>
</tr>
<tr>
<td>Tends to feel misunderstood, mistreated or victimised</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Tends to break things or become physically assaultive when angry</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>Has an exaggerated sense of self-importance; tends to boast or brag</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>Tries to manipulate others' emotions to get what s/he wants</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>Appears to feel privileged and entitled; expects preferential treatment</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Has little empathy; seems unable to understand or respond to others' needs and feelings unless they coincide with his/her own</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>Tends to elicit extreme reactions or stir up strong feelings in others</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>Tends to elicit dislike or animosity in others</td>
<td>0.56</td>
<td></td>
</tr>
<tr>
<td>Appears to experience no remorse for harm or injury caused to others</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>Tends to be deceitful; tends to lie or mislead</td>
<td>0.55</td>
<td></td>
</tr>
<tr>
<td>Tends to be unconcerned with the consequences of his/her actions; appears to feel immune or invulnerable</td>
<td>0.52</td>
<td></td>
</tr>
</tbody>
</table>

Factor 2: Dysphoria/inhibition

<table>
<thead>
<tr>
<th>Factor</th>
<th>Dysphoria/inhibition</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appears to find little or no pleasure, satisfaction or enjoyment in life's activities</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td>Tends to feel s/he is inadequate, inferior or a failure</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>Tends to be insufficiently concerned with meeting own needs; appears not to feel entitled to get or ask for things s/he deserves</td>
<td>0.63</td>
<td></td>
</tr>
<tr>
<td>Tends to feel unhappy, depressed, or despondent</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>Tends to feel life has no meaning</td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td>Tends to be shy or reserved in social situations</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>Appears inhibited about pursuing goals or successes; aspirations or achievements tend to be below his/her potential</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Tends to be inhibited or constricted; has difficulty allowing self to acknowledge or express wishes and impulses</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Has difficulty allowing self to experience strong pleasurable emotions (e.g. excitement, joy, pride)</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Tends to feel listless, fatigued or lacking in energy</td>
<td>0.60</td>
<td></td>
</tr>
<tr>
<td>Tends to express aggression in passive and indirect ways (e.g. may make mistakes, procrastinate, forget, become sulky, etc.)</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Tends to be passive and unassertive</td>
<td>0.59</td>
<td></td>
</tr>
<tr>
<td>Has difficulty acknowledging or expressing anger</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>Tends to feel ashamed or embarrassed</td>
<td>0.50</td>
<td></td>
</tr>
</tbody>
</table>
### Table 1 (continued)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tends to feel helpless, powerless or at the mercy of forces outside his/her control</td>
<td>0.50</td>
</tr>
<tr>
<td>Tends to feel like an outcast or outsider; feels as if s/he does not truly belong</td>
<td>0.47</td>
</tr>
<tr>
<td>Tends to feel s/he is not his/her true self with others; tends to feel false or fraudulent</td>
<td>0.47</td>
</tr>
<tr>
<td>Tends to deny or disavow own needs for caring, comfort, closeness, etc., or to consider such needs unacceptable</td>
<td>0.47</td>
</tr>
<tr>
<td>Appears to want to ‘punish’ self; creates situations that lead to unhappiness, or actively avoids opportunities for pleasure and gratification</td>
<td>0.46</td>
</tr>
<tr>
<td>Tends to feel empty</td>
<td>0.44</td>
</tr>
<tr>
<td>Tends to avoid social situations because of fear of embarrassment or humiliation</td>
<td>0.44</td>
</tr>
<tr>
<td>Tends to be self-critical; sets unrealistically high standards for self and is intolerant of own human defects</td>
<td>0.43</td>
</tr>
<tr>
<td>Is simultaneously needy of, and rejecting toward, others (e.g. craves intimacy and caring, but tends to reject it when offered)</td>
<td>0.42</td>
</tr>
<tr>
<td>Has trouble making decisions; tends to be indecisive or to vacillate when faced with choices</td>
<td>0.42</td>
</tr>
<tr>
<td>Tends to feel guilty</td>
<td>0.42</td>
</tr>
<tr>
<td>Factor 3: Psychological health</td>
<td></td>
</tr>
<tr>
<td>Generally finds contentment and happiness in life’s activities</td>
<td>0.75</td>
</tr>
<tr>
<td>Is creative; is able to see things or approach problems in novel ways</td>
<td>0.72</td>
</tr>
<tr>
<td>Is able to find meaning and satisfaction in the pursuit of goals and ambitions</td>
<td>0.72</td>
</tr>
<tr>
<td>Is psychologically insightful; is able to understand self and others in subtle and sophisticated ways</td>
<td>0.72</td>
</tr>
<tr>
<td>Is able to form close and lasting friendships characterised by mutual support and sharing of experiences</td>
<td>0.71</td>
</tr>
<tr>
<td>Is resilient in the face of extreme stress; seems to be able to face loss, trauma or deeply troubling events with appropriate feeling and continue functioning effectively</td>
<td>0.71</td>
</tr>
<tr>
<td>Has the capacity to recognise alternative viewpoints, even in matters that stir up strong feelings</td>
<td>0.70</td>
</tr>
<tr>
<td>Is capable of hearing information that is emotionally threatening (i.e. that challenges cherished beliefs, perceptions and self-perceptions) and can use and benefit from it</td>
<td>0.70</td>
</tr>
<tr>
<td>Appreciates and responds to humour</td>
<td>0.70</td>
</tr>
<tr>
<td>Is empathic; is sensitive and responsive to other peoples’ needs and feelings</td>
<td>0.65</td>
</tr>
<tr>
<td>Is able to assert him/herself effectively and appropriately when necessary</td>
<td>0.65</td>
</tr>
<tr>
<td>Is able to use his/her talents, abilities and energy effectively and productively</td>
<td>0.64</td>
</tr>
<tr>
<td>Has moral and ethical standards and strives to live up to them</td>
<td>0.64</td>
</tr>
<tr>
<td>Enjoys challenges; takes pleasure in accomplishing things</td>
<td>0.63</td>
</tr>
<tr>
<td>Tends to express affect appropriate in quality and intensity to the situation at hand</td>
<td>0.62</td>
</tr>
<tr>
<td>Tends to be conscientious and responsible</td>
<td>0.60</td>
</tr>
<tr>
<td>Tends to be energetic and outgoing</td>
<td>0.57</td>
</tr>
<tr>
<td>Tends to elicit liking in others</td>
<td>0.53</td>
</tr>
<tr>
<td>Finds meaning in belonging and contributing to a larger community (e.g. volunteer organisations, church, neighbourhood)</td>
<td>0.51</td>
</tr>
<tr>
<td>Is able to find meaning and fulfillment in guiding, mentoring or nurturing others</td>
<td>0.49</td>
</tr>
<tr>
<td>Is articulate; can express self well in words</td>
<td>0.47</td>
</tr>
<tr>
<td>Tends to seek out or create interpersonal relationships in which s/he is in the role of caring for, rescuing or protecting the other</td>
<td>0.44</td>
</tr>
<tr>
<td>Appears comfortable and at ease in social situations</td>
<td>0.41</td>
</tr>
<tr>
<td>Factor 4: Histrionic sexualisation</td>
<td></td>
</tr>
<tr>
<td>Tends to be overly sexually seductive or provocative, whether consciously or unconsciously (e.g. may be inappropriately flirtatious)</td>
<td>0.88</td>
</tr>
<tr>
<td>Tends to have numerous sexual involvements; is promiscuous for a person of his/her age</td>
<td>0.81</td>
</tr>
<tr>
<td>Fantasises about finding ideal, perfect love</td>
<td>0.78</td>
</tr>
<tr>
<td>Tends to become attached quickly or intensely; develops feelings, expectations, etc. that are not warranted by the history or context of the relationship</td>
<td>0.75</td>
</tr>
<tr>
<td>Tends to choose sexual or romantic partners who seem inappropriate in terms of age, status (e.g. social, economic, intellectual), etc.</td>
<td>0.75</td>
</tr>
<tr>
<td>Tends to use his/her physical attractiveness to an excessive degree to gain attention or notice</td>
<td>0.75</td>
</tr>
<tr>
<td>Tends to become involved in romantic or sexual ‘triangles’ (e.g. is most interested in partners who are already attached or sought by someone else)</td>
<td>0.68</td>
</tr>
<tr>
<td>Tends to become attached to, or romantically interested in, people who are emotionally unavailable</td>
<td>0.64</td>
</tr>
<tr>
<td>Tends to be drawn into relationships outside the family in which s/he is emotionally or physically abused</td>
<td>0.58</td>
</tr>
<tr>
<td>Interpersonal relationships tend to be unstable, chaotic and rapidly changing</td>
<td>0.57</td>
</tr>
</tbody>
</table>

(continued overleaf)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Factor</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tends to idealise certain others in unrealistic ways; sees them as ‘all good’, to the exclusion of commonplace human defects</td>
<td>0.57</td>
</tr>
<tr>
<td>Is preoccupied with sex</td>
<td>0.56</td>
</tr>
<tr>
<td>Tends to be suggestible or easily influenced</td>
<td>0.53</td>
</tr>
<tr>
<td>Appears to fear being alone; may go to great lengths to avoid being alone</td>
<td>0.51</td>
</tr>
<tr>
<td>Tends to fear s/he will be rejected or abandoned by those who are emotionally significant</td>
<td>0.47</td>
</tr>
<tr>
<td>Tends to identify with admired others to an exaggerated degree; tends to become an admirer or ‘disciple’ (e.g. may take on the other’s attitudes, beliefs and mannerisms)</td>
<td>0.45</td>
</tr>
<tr>
<td>Beliefs and expectations seem clichéd or stereotypical, as if taken from books or films</td>
<td>0.43</td>
</tr>
<tr>
<td>Expresses emotion in exaggerated and theatrical ways</td>
<td>0.43</td>
</tr>
<tr>
<td>Tends to be overly needy or dependent (e.g. requires excessive reassurance or approval, is ‘clingy’ with friends or parents)</td>
<td>0.41</td>
</tr>
<tr>
<td>Seeks to be the centre of attention</td>
<td>0.41</td>
</tr>
<tr>
<td>Factor 5: Schizotypy</td>
<td></td>
</tr>
<tr>
<td>Is not verbally articulate; has limited ability to express self in words</td>
<td>0.70</td>
</tr>
<tr>
<td>Appearance or manner seems odd or peculiar (e.g. grooming, hygiene, posture, eye contact, speech rhythms, etc. seem somehow strange or ‘off’)</td>
<td>0.64</td>
</tr>
<tr>
<td>Reasoning processes or perceptual experiences seem odd and idiosyncratic (e.g. may make seemingly arbitrary inferences; may see hidden messages or special meanings in ordinary events)</td>
<td>0.61</td>
</tr>
<tr>
<td>Speech tends to be circumstantial, vague, rambling, digressive</td>
<td>0.60</td>
</tr>
<tr>
<td>Tends to elicit boredom in others (e.g. may talk incessantly, without feeling or about inconsequential matters)</td>
<td>0.59</td>
</tr>
<tr>
<td>Tends to think in concrete terms and interpret things in overly literal ways; has limited ability to appreciate metaphor, analogy or nuance</td>
<td>0.55</td>
</tr>
<tr>
<td>Seeks to know less about the ways of the world than might be expected, given his/her intelligence, background and age; appears naive or innocent</td>
<td>0.52</td>
</tr>
<tr>
<td>Appears to experience the past as a series of disjointed or disconnected events; has difficulty giving a coherent account of his/her life or actions</td>
<td>0.51</td>
</tr>
<tr>
<td>Appears to have little need for human company or contact; is genuinely indifferent to the presence of others</td>
<td>0.50</td>
</tr>
<tr>
<td>Appears to have a limited or constricted range of emotions</td>
<td>0.49</td>
</tr>
<tr>
<td>Tends to describe experiences in generalities; is unwilling or unable to offer specific details</td>
<td>0.42</td>
</tr>
<tr>
<td>Factor 6: Emotional dysregulation</td>
<td></td>
</tr>
<tr>
<td>Struggles with genuine wishes to kill him/herself</td>
<td>0.69</td>
</tr>
<tr>
<td>Tends to make repeated suicidal threats or gestures, either as a ‘cry for help’ or as an effort to manipulate others</td>
<td>0.65</td>
</tr>
<tr>
<td>Tends to engage in self-mutilating behaviour (e.g. cutting, burning)</td>
<td>0.64</td>
</tr>
<tr>
<td>Tends to be preoccupied with death and dying</td>
<td>0.64</td>
</tr>
<tr>
<td>Tends to be overly concerned with rules, procedures, order, organisation and schedules</td>
<td>0.39</td>
</tr>
<tr>
<td>Has uncontrolled eating binges followed by ‘purges’ (e.g. makes self vomit, abuses laxatives, fasts); has bulimic episodes</td>
<td>0.48</td>
</tr>
<tr>
<td>Tends to enter altered, dissociated state of consciousness when distressed (e.g. the self or the world feels strange, unfamiliar or unreal)</td>
<td>0.45</td>
</tr>
<tr>
<td>Repeatedly re-experiences or re-lives a past traumatic event (e.g. has intrusive memories or recurring dreams of the event; is startled or terrified by present events that resemble or symbolise the past event)</td>
<td>0.41</td>
</tr>
<tr>
<td>Appears to have a specific phobia (e.g. of snakes, spiders, dogs, aeroplanes, lifts)</td>
<td>0.67</td>
</tr>
<tr>
<td>Has unfounded fears of contracting medical illness; tends to interpret normal aches and pains as symptomatic illness; is hypochondriacal</td>
<td>0.66</td>
</tr>
<tr>
<td>Is troubled by recurrent obsessive thoughts that s/he experiences as senseless and intrusive</td>
<td>0.52</td>
</tr>
<tr>
<td>Tends to adhere rigidly to daily routines and become anxious or uncomfortable when they are altered</td>
<td>0.52</td>
</tr>
<tr>
<td>Tends to develop somatic symptoms in response to stress or conflict (e.g. headache, backache, abdominal pain, asthma)</td>
<td>0.48</td>
</tr>
<tr>
<td>Tends to be anxious</td>
<td>0.43</td>
</tr>
<tr>
<td>Tends to be overly concerned with rules, procedures, order, organisation and schedules</td>
<td>0.39</td>
</tr>
</tbody>
</table>

(continued)
Table 1 (continued)

<table>
<thead>
<tr>
<th>Factor 8: Delinquent behaviour</th>
<th>Loading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tends to misuse illicit drugs</td>
<td>0.72</td>
</tr>
<tr>
<td>Tends to misuse alcohol</td>
<td>0.69</td>
</tr>
<tr>
<td>Tends to engage in unlawful or criminal behaviour</td>
<td>0.64</td>
</tr>
<tr>
<td>Tends to surround him/herself with peers who are delinquent or deeply alienated</td>
<td>0.52</td>
</tr>
<tr>
<td>Tends to seek thrills, novelty and adventure</td>
<td>0.46</td>
</tr>
<tr>
<td>Tends to run away from home</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Factor 9: Sexual conflict
- Is unsure whether s/he is heterosexual, homosexual or bisexual 0.63
- Tends to express qualities, mannerisms or attitudes traditionally associated with the opposite gender (e.g. an effeminate boy, or a girl who disparages anything traditionally feminine) 0.58
- Has conscious homosexual interests (moderate placement implies bisexuality; high placement implies exclusive homosexuality) 0.56
- Shows evidence of unconscious homosexual wishes or interests (e.g. may be excessively homophbic, or may show signs of unacknowledged attraction to a person of the same gender) 0.40

Factor 10: Attentional dysregulation
- Tends to be unreliable and irresponsible (e.g. may fail to meet school or work obligations) 0.63
- Tends to feel bored 0.58
- Is easily frustrated (e.g. ‘gives up’ quickly, has trouble accepting appropriate limits) 0.54
- Is inattentive or easily distracted; has trouble concentrating 0.52
- Tends to use his/her psychological or medical problems to avoid school, work or responsibility (whether consciously or unconsciously) 0.51
- Behaviour at school or work is erratic, unpredictable or grossly inappropriate (e.g. is truant or severely disruptive in class) 0.43

Factor 11: Peer rejection
- Tends to be ignored, neglected or avoided by peers 0.66
- Tends to be bullied or teased by peers 0.62
- Lacks social skills; tends to be socially awkward or inappropriate 0.57
- Lacks close friendships and relationships 0.43

1. Included are items loading above 0.50 (factor 1) or 0.40 (factors 2–11). As common in factor-analytic research, we progressively relaxed criteria for inclusion (from 0.50 to 0.40) to maximize reliability of the smaller factors. Factor 7 includes an item loading 0.39 because this item was clearly conceptually related to the factor’s other items.

2002). With respect to CBCL variables, SWAP–200–A factors with item content suggesting negative emotionality were most highly associated with the internalising score and related sub-scales, whereas factors suggesting externalising pathology correlated most highly with the externalising scale and its component scales.

Finally, we examined the relationship between the SWAP–200–A factors and adaptive functioning variables (Table 5). The data provided additional preliminary support for validity. For example, psychopathy/malignant narcissism correlated with all six variables in the expected directions and also predicted history of trouble with the law. Although most of the SWAP–200–A factors predicted poor peer relationships, the correlations were particularly large for schizotypy and peer rejection. The psychological health factor also performed as expected. Of particular interest is the strong pattern of associations between SWAP–200–A factors and variables such as history of suicide attempts, arrests and psychiatric hospitalisation, which are relatively objective and require minimal clinical inference (and hence are not readily attributable to clinician biases).

DISCUSSION

The SWAP–200–A shows promise as an instrument for assessing personality pathology in adolescents. Its factors are theoretically and clinically coherent, internally consistent, and show convergent and discriminant validity in predicting a range of variables including Axis II diagnosis, FFM scores, CBCL scores and measures of adaptive functioning.

Several aspects of the factor solution are interesting from a conceptual and clinical point of view. The factor structure of the adolescent instrument strongly resembles the factor structure obtained using the adult version of the instrument (Shedler & Westen, 2004), with a few important exceptions. Whereas the adult instrument yields separate factors for narcissism, psychopathy and hostility, the SWAP–200–A yields a large first factor that blends these constructs, and includes a separate factor that assesses delinquent behaviour. These differences may reflect developmental differences in the expression of the underlying traits, or they may reflect instability of the factor solution reflecting sample size. A distinction between a core psychopathy factor and a delinquent behaviour factor, however, mirrors results of factor analysis of the Psychopathy Checklist – Revised, which similarly distinguishes a callous, hostile, externalising personality style from a tendency to become involved in criminal activity (Hare, 1998). Other differences between the adult and adolescent factor structures seem to reflect...
Table 2  Correlations between factor scores and composite personality disorder ratings

<table>
<thead>
<tr>
<th>SWAP–200–A factor</th>
<th>Composite personality disorder ratings (n = 285)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cluster A</td>
</tr>
<tr>
<td></td>
<td>Paranoid Schizoid Schizotypal Antisocial Borderline Histrionic Narcissistic Avoidant Dependent Obsessive</td>
</tr>
<tr>
<td>1 Psychopathy/malignant narcissism</td>
<td>0.50*** 0.33*** 0.27***</td>
</tr>
<tr>
<td>2 Dysphoria/inhibition</td>
<td>0.30*** 0.41*** 0.40***</td>
</tr>
<tr>
<td>3 Psychological health</td>
<td>−0.28*** −0.51*** −0.35***</td>
</tr>
<tr>
<td>4 Histrionic sexualisation</td>
<td>0.35*** 0.08 0.19***</td>
</tr>
<tr>
<td>5 Schizotypy</td>
<td>0.41*** 0.63*** 0.62***</td>
</tr>
<tr>
<td>6 Emotional dysregulation</td>
<td>0.24*** 0.13 0.21***</td>
</tr>
<tr>
<td>7 Anxious obsessuality</td>
<td>0.13*** 0.18** 0.29***</td>
</tr>
<tr>
<td>8 Delinquent behaviour</td>
<td>0.28*** 0.11 0.04</td>
</tr>
<tr>
<td>9 Sexual conflict</td>
<td>0.26*** 0.12 0.24***</td>
</tr>
<tr>
<td>10 Attentional dysregulation</td>
<td>0.33*** 0.35*** 0.32***</td>
</tr>
<tr>
<td>11 Peer rejection</td>
<td>0.37*** 0.58*** 0.62***</td>
</tr>
</tbody>
</table>

1. Criterion variables predicted to be most strongly associated (positively or negatively) with each factor are identified in bold.

*P < 0.05, **P < 0.01, ***P < 0.001.

developmental differences. For example, where the adult instrument yields a schizoid factor and a thought disorder (schizotypy) factor, the adolescent version yields a single schizotypy factor and a separate peer rejection factor. Further, although both versions of the instrument produced a sexual conflict factor, in adolescents the items focus primarily on conflicts regarding sexual orientation, probably reflecting the salience of this issue in adolescents struggling with homosexual feelings, and the lack of knowledge both teenagers and their clinicians are likely to have about other kinds of sexual conflict that might not become expressed until adulthood.

Also notable is the distinction between dysphoria/inhibition (the factor most closely related to negative affectivity or neuroticism), anxious obsessuality and emotional dysregulation, three variables that were only moderately intercorrelated. The distinction between negative affectivity on the one hand and emotional dysregulation on the other has emerged recently in other samples using both factor and Q-factor analysis with a variety of instruments (e.g. Livesley et al, 1998; Westen & Shedler, 1999b; Westen et al, 2003). Whereas most factor-analytically derived models of personality and mood distinguish positive and negative affectivity – or (in the FFM) their close cousins, extroversion and neuroticism – the emergence of an independent emotional dysregulation factor in clinical samples with instruments intended for clinical use may be significant, drawing attention to the distinction between stably anxious or dysphoric

Table 3  Correlations between factor scores and Five Factor Model Checklist ratings

<table>
<thead>
<tr>
<th>SWAP–200–A factor</th>
<th>Five Factor Model ratings (n = 284)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neuroticism  Extroversion  Agreeableness  Conscientiousness  Openness</td>
</tr>
<tr>
<td>1 Psychopathy/malignant narcissism</td>
<td>0.14 0.06 −0.58*** −0.49*** −0.16**</td>
</tr>
<tr>
<td>2 Dysphoria/inhibition</td>
<td>0.55*** −0.38*** 0.03 −0.04 0.00</td>
</tr>
<tr>
<td>3 Psychological health</td>
<td>−0.08 0.56*** 0.76*** 0.66*** 0.65***</td>
</tr>
<tr>
<td>4 Histrionic sexualisation</td>
<td>0.30*** 0.15** 0.18** −0.25*** 0.11</td>
</tr>
<tr>
<td>5 Schizotypy</td>
<td>0.15 −0.30*** −0.30*** −0.32** −0.22***</td>
</tr>
<tr>
<td>6 Emotional dysregulation</td>
<td>0.46*** −0.01 0.09 0.12 0.01</td>
</tr>
<tr>
<td>7 Anxious obsessuality</td>
<td>0.38*** −0.15** 0.14 0.20*** 0.06</td>
</tr>
<tr>
<td>8 Delinquent behaviour</td>
<td>0.04 0.24*** −0.28*** −0.43*** 0.04</td>
</tr>
<tr>
<td>9 Sexual conflict</td>
<td>0.21*** 0.01 −0.01 0.00 0.16**</td>
</tr>
<tr>
<td>10 Attentional dysregulation</td>
<td>0.35*** −0.10 −0.33*** −0.56*** −0.13</td>
</tr>
<tr>
<td>11 Peer rejection</td>
<td>0.23*** −0.43*** −0.28*** −0.25*** −0.13</td>
</tr>
</tbody>
</table>

1. Criterion variables predicted to be most strongly associated (positively or negatively) with each factor are identified in bold.

*P < 0.05, **P < 0.01, ***P < 0.001.
### Table 4  Correlations between factor scores and Child Behavior Checklist variables

<table>
<thead>
<tr>
<th>SWAP–200 A factor</th>
<th>Child Behavior Checklist (n=224–282)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Withdrawn</td>
</tr>
<tr>
<td></td>
<td>Somatic complaints</td>
</tr>
<tr>
<td></td>
<td>Anxious/depressed</td>
</tr>
<tr>
<td></td>
<td>Social problems</td>
</tr>
<tr>
<td></td>
<td>Thought problems</td>
</tr>
<tr>
<td></td>
<td>Attention problems</td>
</tr>
<tr>
<td></td>
<td>Delinquent behaviour</td>
</tr>
<tr>
<td></td>
<td>Aggressive behaviour</td>
</tr>
<tr>
<td></td>
<td>Other problems</td>
</tr>
<tr>
<td></td>
<td>Sex problems</td>
</tr>
<tr>
<td></td>
<td>Internalising</td>
</tr>
<tr>
<td></td>
<td>Externalising</td>
</tr>
<tr>
<td></td>
<td>Total problems</td>
</tr>
<tr>
<td>Psychopathy/malignant narcissism</td>
<td>0.22***</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>0.40***</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>0.41***</td>
</tr>
<tr>
<td></td>
<td>0.64***</td>
</tr>
<tr>
<td></td>
<td>0.82***</td>
</tr>
<tr>
<td></td>
<td>0.21***</td>
</tr>
<tr>
<td></td>
<td>0.23***</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td>Dysphoria/inhibition</td>
<td>0.59***</td>
</tr>
<tr>
<td></td>
<td>0.29***</td>
</tr>
<tr>
<td></td>
<td>0.64***</td>
</tr>
<tr>
<td></td>
<td>0.23***</td>
</tr>
<tr>
<td></td>
<td>0.39***</td>
</tr>
<tr>
<td></td>
<td>0.33***</td>
</tr>
<tr>
<td></td>
<td>–0.02</td>
</tr>
<tr>
<td></td>
<td>–0.08</td>
</tr>
<tr>
<td></td>
<td>0.44***</td>
</tr>
<tr>
<td></td>
<td>0.18**</td>
</tr>
<tr>
<td>Psychological health</td>
<td>–0.40***</td>
</tr>
<tr>
<td></td>
<td>–0.10</td>
</tr>
<tr>
<td></td>
<td>–0.02</td>
</tr>
<tr>
<td></td>
<td>–0.40***</td>
</tr>
<tr>
<td></td>
<td>–0.13</td>
</tr>
<tr>
<td></td>
<td>–0.29***</td>
</tr>
<tr>
<td></td>
<td>–0.27***</td>
</tr>
<tr>
<td></td>
<td>–0.36***</td>
</tr>
<tr>
<td></td>
<td>–0.06</td>
</tr>
<tr>
<td></td>
<td>–0.02</td>
</tr>
<tr>
<td>Histrionic sexualisation</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>0.37***</td>
</tr>
<tr>
<td></td>
<td>0.28***</td>
</tr>
<tr>
<td></td>
<td>0.23***</td>
</tr>
<tr>
<td></td>
<td>0.36***</td>
</tr>
<tr>
<td></td>
<td>0.32***</td>
</tr>
<tr>
<td></td>
<td>0.34***</td>
</tr>
<tr>
<td></td>
<td>0.43***</td>
</tr>
<tr>
<td></td>
<td>0.44**</td>
</tr>
<tr>
<td>Schizotypy</td>
<td>0.50***</td>
</tr>
<tr>
<td></td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>0.21***</td>
</tr>
<tr>
<td></td>
<td>0.51***</td>
</tr>
<tr>
<td></td>
<td>0.45***</td>
</tr>
<tr>
<td></td>
<td>0.48***</td>
</tr>
<tr>
<td></td>
<td>0.25***</td>
</tr>
<tr>
<td></td>
<td>0.26***</td>
</tr>
<tr>
<td></td>
<td>0.31***</td>
</tr>
<tr>
<td></td>
<td>0.26***</td>
</tr>
<tr>
<td></td>
<td>0.33***</td>
</tr>
<tr>
<td></td>
<td>0.27***</td>
</tr>
<tr>
<td>Emotional dysregulation</td>
<td>0.24***</td>
</tr>
<tr>
<td></td>
<td>0.18**</td>
</tr>
<tr>
<td></td>
<td>0.47***</td>
</tr>
<tr>
<td></td>
<td>0.22***</td>
</tr>
<tr>
<td></td>
<td>0.31***</td>
</tr>
<tr>
<td></td>
<td>0.19**</td>
</tr>
<tr>
<td></td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>0.60***</td>
</tr>
<tr>
<td></td>
<td>0.23**</td>
</tr>
<tr>
<td></td>
<td>0.41***</td>
</tr>
<tr>
<td></td>
<td>0.23**</td>
</tr>
<tr>
<td>Anxious obsessiornality</td>
<td>0.22***</td>
</tr>
<tr>
<td></td>
<td>0.43***</td>
</tr>
<tr>
<td></td>
<td>0.51***</td>
</tr>
<tr>
<td></td>
<td>0.21***</td>
</tr>
<tr>
<td></td>
<td>0.44***</td>
</tr>
<tr>
<td></td>
<td>0.21***</td>
</tr>
<tr>
<td></td>
<td>–0.07</td>
</tr>
<tr>
<td></td>
<td>–0.12</td>
</tr>
<tr>
<td></td>
<td>0.52***</td>
</tr>
<tr>
<td></td>
<td>0.26***</td>
</tr>
<tr>
<td></td>
<td>0.54***</td>
</tr>
<tr>
<td></td>
<td>–0.10</td>
</tr>
<tr>
<td>Delinquent behaviour</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>–0.04</td>
</tr>
<tr>
<td></td>
<td>–0.04</td>
</tr>
<tr>
<td></td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>0.38***</td>
</tr>
<tr>
<td></td>
<td>0.82***</td>
</tr>
<tr>
<td></td>
<td>0.54***</td>
</tr>
<tr>
<td></td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>0.21***</td>
</tr>
<tr>
<td></td>
<td>–0.20</td>
</tr>
<tr>
<td>Sexual conflict</td>
<td>0.16**</td>
</tr>
<tr>
<td></td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>0.22***</td>
</tr>
<tr>
<td></td>
<td>0.16**</td>
</tr>
<tr>
<td></td>
<td>0.20***</td>
</tr>
<tr>
<td></td>
<td>0.18**</td>
</tr>
<tr>
<td></td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>0.41***</td>
</tr>
<tr>
<td></td>
<td>0.57***</td>
</tr>
<tr>
<td></td>
<td>0.21***</td>
</tr>
<tr>
<td></td>
<td>0.09</td>
</tr>
<tr>
<td>Attentional dysregulation</td>
<td>0.41***</td>
</tr>
<tr>
<td></td>
<td>0.21***</td>
</tr>
<tr>
<td></td>
<td>0.22***</td>
</tr>
<tr>
<td></td>
<td>0.30***</td>
</tr>
<tr>
<td></td>
<td>0.18**</td>
</tr>
<tr>
<td></td>
<td>0.68***</td>
</tr>
<tr>
<td></td>
<td>0.42***</td>
</tr>
<tr>
<td></td>
<td>0.45***</td>
</tr>
<tr>
<td></td>
<td>0.26***</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
</tr>
<tr>
<td></td>
<td>0.32**</td>
</tr>
<tr>
<td></td>
<td>0.46***</td>
</tr>
<tr>
<td>Peer rejection</td>
<td>0.50***</td>
</tr>
<tr>
<td></td>
<td>0.16**</td>
</tr>
<tr>
<td></td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>0.77***</td>
</tr>
<tr>
<td></td>
<td>0.39***</td>
</tr>
<tr>
<td></td>
<td>0.45***</td>
</tr>
<tr>
<td></td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>0.22***</td>
</tr>
<tr>
<td></td>
<td>0.29***</td>
</tr>
<tr>
<td></td>
<td>0.22***</td>
</tr>
<tr>
<td></td>
<td>0.45***</td>
</tr>
<tr>
<td></td>
<td>0.18**</td>
</tr>
</tbody>
</table>

1. Range of numbers reflects missing data for particular variables.
2. Criterion variables predicted to be most strongly associated (positively or negatively) with each factor are identified.


*^P < 0.05, **^P < 0.01, ***^P < 0.001.
personality traits and a distinct form of affectivity in which emotions spiral out of control. Emotional dysregulation appears to be associated with history of traumatic experiences such as sexual abuse and early parental separations (e.g. Nakash-Eisikovits et al., 2003) and may be related to dysregulation of corticotrophin releasing factor, which has been linked both to childhood trauma and to later vulnerability to syndromes such as major depression and panic disorder (Heim & Nemeroff, 2001).

Differences between the factor solution obtained here and the personality prototypes obtained from the same sample using Q-factor analysis are also instructive. It is noteworthy that the first six factors resemble six of the seven obtained Q factors, which means that two very different ways of analysing the data – one identifying types (prototypes) and the other traits – yielded similar dimensions. However, the remaining scales identify traits or psychological functions that represent important aspects of personality pathology in adolescents but do not define a broader personality style: anxious obsessivity, delinquent behaviour, sexual conflict, attentional dysregulation and peer rejection.

**Potential objections and limitations**

This study has three primary limitations. The first is its exclusive reliance on a single informant (the treating clinician), which creates the possibility (like most studies of psychopathology, which rely exclusively on patient reports) of observer bias. Future research should attempt to replicate these findings using interview, informant and laboratory data as external criteria. Nevertheless, several factors limit this concern. First, SWAP–200 personality descriptions and ratings of adaptive functioning show high interrater reliability and validity and strongly predict relevant criterion variables as assessed by independent informants (Westen & Maderrisoglu, 2003; Westen & Weinberger, 2004). Second, clinicians varied in their training (psychiatrists and psychologists) and theoretical orientations, and were unfamiliar with the factor structure of the instrument, minimising the likelihood of systematic sources of error stemming from rater biases. Finally, whereas factor analysis of the DSM–IV Axis II checklist in this sample produced a factor structure that strongly resembled the DSM–IV classification (Durrett & Westen, 2005), factor analysis of the items of the SWAP–200–A, which include items assessing all of the Axis II criteria, did not. Thus, it is difficult to see how clinician biases could both lead to convergence with and divergence from the DSM–IV description of personality pathology in an adolescent sample.

A second potential objection is sample size. Clearly, the next step in this research requires a substantially larger sample, and such a study is now nearing completion (projected n=1000). Nevertheless, recent thinking about factor analysis, based on data from Monte Carlo simulations and other studies, suggests that factor solutions stabilise with far fewer cases than previously believed (often by 100 cases) as long as the factors are well marked by a sufficient number of items with loadings above 0.40 or 0.50, as they were here (see Fabregat et al., 1999; Russell, 2002).

A final potential objection regards the question of the durability of personality pathology in adolescents and the appropriateness of diagnosing personality pathology at all in teenagers, an issue we have addressed elsewhere in detail (Westen & Chang, 2000). The data presented here are cross-sectional, and future research should employ longitudinal designs. Nevertheless, recent research using different designs and measures suggests not only that personality can be assessed reliably in adolescents but that recognisable forms of personality pathology can be measured in adolescents and predict substantial variance in a range of outcomes, including outcomes measured longitudinally, above and beyond Axis I diagnosis (Ludolph et al., 1990; Johnson et al., 1999; Westen et al., 2003).

**Implications**

We note here two implications. First, research on adolescent psychopathology has often ignored personality variables because of the lack of appropriate constructs and

---

**Table 5 Correlations between factor scores and adaptive functioning variables**

<table>
<thead>
<tr>
<th>SWAP–200–A factor</th>
<th>Adaptive functioning variables (n=242–283)1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>School functioning</td>
</tr>
<tr>
<td>1 Psychopathy/malignant narcissism</td>
<td>-0.42***</td>
</tr>
<tr>
<td>2 Dysphoria/inhibition</td>
<td>-0.07</td>
</tr>
<tr>
<td>3 Psychological health</td>
<td>0.51***</td>
</tr>
<tr>
<td>4 Histrionic sexualisation</td>
<td>-0.06</td>
</tr>
<tr>
<td>5 Schizotypy</td>
<td>-0.24***</td>
</tr>
<tr>
<td>6 Emotional dysregulation</td>
<td>-0.06</td>
</tr>
<tr>
<td>7 Anxious obsessivity</td>
<td>0.15</td>
</tr>
<tr>
<td>8 Delinquent behaviour</td>
<td>-0.40***</td>
</tr>
<tr>
<td>9 Sexual conflict</td>
<td>0.01</td>
</tr>
<tr>
<td>10 Attentional dysregulation</td>
<td>-0.55***</td>
</tr>
<tr>
<td>11 Peer rejection</td>
<td>-0.22***</td>
</tr>
</tbody>
</table>

1. Range of numbers reflects missing data for particular variables.
2. Criterion variables predicted to be most strongly associated (positively or negatively) with each factor are identified in bold.

*P<0.05, **P<0.01, ***P<0.001.
measures. Availability of reliable and valid measures of adolescent personality pathology may prove useful in distinguishing clinical phenomena that may be quite distinct, such as different types of adolescents who share an Axis I mood disorder diagnosis.

Second, the data raise questions about whether we do better to characterise adolescent personality pathology in terms of the kinds of broad constellations connoted by the term ‘personality disorder’ or whether we might do better to focus on more specific traits that combine in various ways to form some of those constellations. Although we are far from a definitive answer to foundational questions such as these (with respect to either child and adolescent or adult psychopathology), one potential solution might be to combine both forms of classification. For example, an adolescent personality axis could include both a set of personality prototypes describing personality constellations as well as a small set of non-redundant traits such as attentional dysregulation or sexual conflict. From a clinical standpoint, aside from these descriptive (typological and trait approaches), another approach to diagnosis – explicit in the construction of the SWAP–200–A item set (which describes not only behaviours but also internal processes) – is functional assessment. A functional assessment focuses on understanding what is going right, what is going wrong and what conditions certain pathological processes manifest for a given patient (i.e. the conditions under which certain functions go awry or break down). One way to reduce the gulf between clinicians and researchers is to decrease the gap between descriptive nosological constructs, which tend to be the focus of research, and functional constructs, which are essential to everyday clinical practice.

ACKNOWLEDGEMENTS

Preparation of this paper was supported in part by National Institute of Mental Health grants MH59685 and MH60892. The authors gratefully acknowledge the assistance of more than 300 clinicians who helped us refine the SWAP–200–A assessment instrument, including the 294 who participated in this study.

REFERENCES


CLINICAL IMPLICATIONS

- Personality pathology can be assessed in adolescents using the Shedler–Westen Assessment Procedure for Adolescents, which shows promise as a clinically and empirically useful instrument for assessing and taxonomising adolescent personality pathology.

- Negative affectivity (e.g. anxiety and depression) appears to be distinct from emotional dysregulation (the tendency for emotions to spiral out of control and motivate desperate measures to regulate it) as a personality trait, and the two constructs are likely to have different aetiologies and implications for treatment.

- A clinically useful, comprehensive personality axis for adolescents may need to incorporate both a set of personality prototypes or constellations similar to those described in DSM–IV, as well as a small set of traits such as delinquent behaviour, attentional dysregulation and sexual conflict, which do not describe personality styles but do represent enduring problems of adolescence.

LIMITATIONS

- Reliance on data from a single informant requires replication using criterion data from other sources.

- Factor analysis with a large item set requires replication with a larger sample.

- Identification of personality constellations or traits in adolescents requires longitudinal research to study their course, change and development.

DREW WESTEN, PHD, Emory University, Atlanta, Georgia; LISSA DUTRA, EdM, Center for Anxiety and Related Disorders and Department of Psychology, Boston University, Boston, Massachusetts; JONATHAN SHEDLER, PHD, University of Denver, Denver, Colorado, USA

Correspondence: Dr Drew Westen, Department of Psychology and Department of Psychiatry and Behavioral Sciences, Emory University, 532 N. Kilgo Cir., Atlanta, Georgia 30322, USA. E-mail: dwesten@emory.edu

(First received 5 January 2004, final revision 9 September 2004, accepted 10 September 2004)


Suitability and utility of the CORE–OM and CORE–A for assessing severity of presenting problems in psychological therapy services based in primary and secondary care settings

MICHAEL BARKHAM, NAOMI GILBERT, JANICE CONNELL, CHRIS MARSHALL and ELSPETH TWIGG

Background There is a need for reliable assessment tools that are suitable for the counselling and the psychological therapy services in primary and secondary care settings.

Aims To test the suitability and utility of the Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE–OM) and CORE–Assessment (CORE–A) assessment tools.

Method Service intake data were analysed from counselling and psychological therapy services in 32 primary care settings and 17 secondary care settings.

Results Completion rates exceeded 98% in both of the settings sampled. Intake severity levels were similar but secondary care patients were more likely to score above the risk cut-off and the severe threshold and to have experienced their problems for a greater duration.

Conclusions The CORE–OM and CORE–A are suitable assessment tools that show small but logical differences between psychological therapy services in primary- and secondary-based care.

Declarations of interest MB received funding from the Mental Health Foundation and the Artemis Trust to support the development of the CORE–OM and CORE–A, respectively.

There is increasing pressure on mental health services to adopt assessment and outcome measures that can be used routinely in mental health settings (Department of Health, 2001). Measures need to be appropriate for specific patient populations but also be capable of ‘following the patient’ through the various tiers of mental health services. The Clinical Outcomes in Routine Evaluation – Outcome Measure (CORE–OM; Barkham et al, 1998, 2001; Evans et al, 2002) has become a widely used patient self-report measure across service settings delivering psychological treatments, together with a practitioner-completed component termed the CORE–Assessment (CORE–A; Mellor-Clark et al, 1999; Mellor-Clark & Barkham, 2000). However, there has been no test to compare the CORE–OM and CORE–A in assessing the severity of presenting problems in bona fide primary versus secondary care settings. Accordingly, first we investigate whether the CORE–OM and CORE–A are appropriate as assessment tools in both service settings, and then we identify whether they reflect differences between the two settings.

METHOD

The data
This study reports on data collected by 49 National Health Service (NHS) sites routinely using the CORE–OM to monitor patients at intake to their services. The data were anonymised and aggregated and are independent of data set out in a previous study reporting psychometric properties of the CORE–OM (Evans et al, 2002). In total, 32 sites were primary care based, providing counselling or psychology services within primary care groups or trusts. The remaining 17 sites were secondary care based and provided clinical psychology and psychotherapy services. The majority of referrals were from general practitioners, accounting for 93.3% of referrals to primary care sites and 64.5% to secondary care sites. Data (CORE–OM and/or CORE–A) were completed for 6610 primary care patients and 2311 secondary care patients in total.

Patient samples
Patients not completing the CORE–OM or missing more than three items from the 34-item measure were excluded from the mean score calculations. Using these criteria, 5733 primary care patients and 1918 secondary care patients were selected for inclusion. Table 1 presents demographic information for the two patient samples.

Measures
Patient-completed measure: CORE–OM
The CORE–OM comprises 34 items addressing domains of subjective well-being (4 items), symptoms (12 items), functioning (12 items) and risk (6 items; 4 ‘risk to self’ items and 2 ‘risk to others’ items). Within the symptom domain ‘item clusters’ address anxiety (4 items), depression (4 items), physical problems (2 items) and trauma (2 items). The functioning domain item clusters address general functioning (4 items), close relationships (4 items) and social relationships (4 items).

Items are scored on a five-point scale from 0 (‘not at all’) to 4 (‘all the time’). Half of the items focus on low-intensity problems (e.g. ‘I feel anxious/nervous’) and half focus on high-intensity problems (e.g. ‘I feel panic/terror’). Eight items are keyed positively.

All services in the study asked patients to complete the CORE–OM as a measure of distress at intake (i.e. before any intervention). In practice, the CORE–OM was completed during screening or assessment by 73.8% of primary care patients and 87.3% of secondary care patients, and completed at the first therapy session by the remaining 26.2% in primary care and 12.7% in secondary care.

Practitioner-completed measure: CORE–A
The CORE–A enables the collection of referral information, demographics, assessment, outcome, and data on presenting problem severity and duration. The CORE–A lists the following 14 problems: depression, anxiety, psychosis, personality problems, cognitive/learning difficulties, eating disorder, physical problems, addictions, trauma/abuse, bereavement,
Table 1  Patient sample demographics

<table>
<thead>
<tr>
<th></th>
<th>Primary care (n=5733)</th>
<th>Secondary care (n=1918)</th>
<th>( \chi^2 )</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1629</td>
<td>659</td>
<td>24.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Female</td>
<td>4104</td>
<td>1259</td>
<td>24.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 20</td>
<td>243</td>
<td>86</td>
<td>0.2</td>
<td>0.65</td>
</tr>
<tr>
<td>20–29</td>
<td>1283</td>
<td>470</td>
<td>3.7</td>
<td>0.06</td>
</tr>
<tr>
<td>30–39</td>
<td>1883</td>
<td>600</td>
<td>1.6</td>
<td>0.21</td>
</tr>
<tr>
<td>40–49</td>
<td>1225</td>
<td>410</td>
<td>0.0</td>
<td>0.99</td>
</tr>
<tr>
<td>50–59</td>
<td>727</td>
<td>260</td>
<td>1.0</td>
<td>0.32</td>
</tr>
<tr>
<td>&gt; 60</td>
<td>335</td>
<td>84</td>
<td>5.9</td>
<td>0.02</td>
</tr>
<tr>
<td>Not recorded</td>
<td>37</td>
<td>8</td>
<td>1.3</td>
<td>0.26</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>213</td>
<td>44</td>
<td>8.9</td>
<td>0.003</td>
</tr>
<tr>
<td>Black</td>
<td>105</td>
<td>23</td>
<td>3.5</td>
<td>0.06</td>
</tr>
<tr>
<td>White European</td>
<td>4526</td>
<td>1691</td>
<td>80.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Mixed race</td>
<td>23</td>
<td>10</td>
<td>0.5</td>
<td>0.49</td>
</tr>
<tr>
<td>Other</td>
<td>79</td>
<td>29</td>
<td>0.2</td>
<td>0.67</td>
</tr>
<tr>
<td>Not recorded</td>
<td>787</td>
<td>121</td>
<td>75.6</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Data analysis

All data were scanned optically using FORMIC software (Formic Design and Automatic Data Capture, 1996). Statistical analyses were carried out using the Statistical Package for the Social Sciences for Windows (version 11). The CORE–OM overall mean scores and non-risk scores were calculated using ‘pro-rating’, where up to three items were missed (i.e. if two items were not completed, the total score would be divided by 32 rather than 34). Domain mean scores were not ‘pro-rated’ if more than one item was missing from that domain.

Completion rates (n clients completing the CORE–OM) and missing items were analysed using the full data-set (n=6610 primary care and n=2311 secondary care).

All subsequent analyses were carried out on the samples of patients completing the CORE–OM and fulfilling the criteria for pro-rating (n=5733 primary care and n=1918 secondary care).

Internal consistency of the CORE–OM was calculated using Cronbach’s coefficient \( \alpha \) (Cronbach, 1951). Statistical power was high due to the large sample sizes, therefore differences in mean scores between samples are reported using confidence intervals (Gardner & Altman, 1986) and effect sizes (Cohen, 1988) rather than significance tests. An ‘effect size’ represents a standard deviation unit and is calculated as the difference between means divided by the pooled standard deviation. The standard guide to the effect size differences denotes three bands: 0.2 (small) 0.5 (medium) and 0.8 (large). On the basis of Cohen (1988), noting that a 0.2 effect size involved an 85% overlap between distributions, it has been suggested that an effect size of 0.4 (involving a 73% overlap) be used as the criterion for clinically meaningful differences (Elliott et al, 1993). Chi-squared analysis was used to test proportional differences between samples (e.g. demographic characteristics).

To facilitate comparisons regarding the range of severity, we applied two cut-offs to the CORE–OM data that reflected differing levels of severity (for details, see Jacobson & Truax, 1991). The first cut-off on the CORE–OM, termed ‘clinical’, was defined as a score of 1.19 for men and 1.29 for women and was derived from calculating the CORE–OM score that would best demarcate membership of the general population (i.e. a lower score) or a clinical population (i.e. a higher score) using the following formula (see Evans et al, 2002):

\[
\text{mean}_{\text{norm}} \pm 1 \times s_{\text{norm}} \pm \text{mean}_{\text{clin}} \pm 1 \times s_{\text{clin}}
\]

The second cut-off, termed ‘severe’, was a CORE–OM score of 2.50 (both men and women) that approximated to a score of 1 s.d. above the mean for a clinical population and differentiated a mild/moderate clinical population from a severe clinical population (see Barkham et al, 2001). Odds ratio analysis was applied to estimate the caseness rate ratio using clinical cut-off points for the CORE–OM. Effect sizes and confidence intervals for proportions (Wilson, 1927) were calculated using Microsoft Excel 2000.

Table 2  The CORE–Outcome Measure items above the 95% CI omission threshold

<table>
<thead>
<tr>
<th>Item</th>
<th>Primary care</th>
<th>Secondary care</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n missed</td>
<td>%</td>
</tr>
<tr>
<td>16 I made plans to end my life</td>
<td>74</td>
<td>1.3</td>
</tr>
<tr>
<td>31 I have felt optimistic about my future</td>
<td>75</td>
<td>1.3</td>
</tr>
<tr>
<td>32 I have achieved the things I wanted to</td>
<td>85</td>
<td>1.5</td>
</tr>
<tr>
<td>17 I have felt overwhelmed by my problems</td>
<td>92</td>
<td>1.6</td>
</tr>
<tr>
<td>19 I have felt warmth or affection for someone</td>
<td>129</td>
<td>2.2</td>
</tr>
<tr>
<td>3 I have felt I have someone to turn to for</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

1. Values not given are below the threshold.
RESULTS

Acceptability
In order to assess whether the CORE–OM was acceptable to clients in both primary and secondary care settings, we examined completion rates (i.e. number of clients completing the CORE–OM) and missed items at intake assessment. Of the total, 5833 (88.3%) primary care patients and 1940 (84.0%) secondary care patients completed the CORE–OM. The completion rate was significantly higher for the primary care sample ($\chi^2=28.2, P<0.001$, 95% CI difference 2.7–6.0%). However, the proportion of completed measures with fewer than three items missing (i.e. within the criteria for pro-rating) was similar in both settings: 5733 (98.3%) in primary care and 1918 (98.9%) in secondary care ($\chi^2=3.2, P=0.08$, 95% CI −1.0 to 0.1%).

The most commonly missed item in both primary and secondary settings was no. 19 (‘I have felt warmth and affection for someone’). The overall item omission rates were 0.9% (95% CI 0.7–1.2%) for primary care and 0.8% (95% CI 0.5–1.3%) for secondary care. In the primary care sample, five items had missing cases above the upper threshold (1.2%) of the 95% confidence interval. In the secondary care sample, two items had missing cases above the threshold (1.3%). Table 2 summarises the items above the threshold in each sample.

Table 3 Internal consistency of CORE–Outcome Measure by service setting (Cronbach’s $\alpha$)

<table>
<thead>
<tr>
<th>Domain</th>
<th>Primary (n=5733)</th>
<th>Secondary (n=1918)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$n$ items</td>
<td>$\alpha$</td>
</tr>
<tr>
<td>Well-being</td>
<td>4</td>
<td>0.70</td>
</tr>
<tr>
<td>Symptoms</td>
<td>12</td>
<td>0.87</td>
</tr>
<tr>
<td>Anxiety</td>
<td>4</td>
<td>0.78</td>
</tr>
<tr>
<td>Depression</td>
<td>4</td>
<td>0.74</td>
</tr>
<tr>
<td>Physical</td>
<td>2</td>
<td>0.40</td>
</tr>
<tr>
<td>Trauma</td>
<td>2</td>
<td>0.72</td>
</tr>
<tr>
<td>Functioning</td>
<td>12</td>
<td>0.85</td>
</tr>
<tr>
<td>General</td>
<td>4</td>
<td>0.77</td>
</tr>
<tr>
<td>Close relationships</td>
<td>4</td>
<td>0.65</td>
</tr>
<tr>
<td>Social relationships</td>
<td>4</td>
<td>0.70</td>
</tr>
<tr>
<td>Risk</td>
<td>6</td>
<td>0.77</td>
</tr>
<tr>
<td>Risk to self</td>
<td>4</td>
<td>0.81</td>
</tr>
<tr>
<td>Risk to others</td>
<td>2</td>
<td>0.59</td>
</tr>
<tr>
<td>Non-risk items</td>
<td>28</td>
<td>0.93</td>
</tr>
</tbody>
</table>

All items | 34     | 0.93    | 0.93–0.94 | 0.95    | 0.94–0.95 |

Internal consistency
We used Cronbach’s coefficient $\alpha$ to calculate the internal reliability of the CORE–OM domains and item clusters within domains for both primary and secondary care settings. Although the item clusters were originally selected to represent the range of patient experience and not intended to be used as sub-scales, we calculated $\alpha$ values for them in order to test the robustness of the components within each domain. The $\alpha$ value indicates the proportion of covariance between items. Table 3 illustrates that all domains showed good internal reliability, with $\alpha >0.70$ and <0.97 in each setting. In both primary and secondary care, the well-being domain had the lowest internal consistency. Values of $\alpha$ exceeded 0.70 for six of the nine item clusters – anxiety, depression, trauma, general functioning, social relationships, and risk to self – whereas for close relationships $\alpha$ was 0.65–0.70. Only for physical problems and risk to others (both of which comprised just two items) was $\alpha < 0.60$.

The CORE–OM profile of severity of problems

Overall scores
To compare the overall CORE–OM scores in primary and secondary care settings, we generated notched boxplots and histograms presenting the distribution of CORE–OM mean scores for all items (see Figs 1 and 2). In terms of overall mean scores, the two settings showed a strikingly similar distribution. Figure 1 shows that there were

---

241
four outliers in the primary care sample scoring above the maximum secondary care score of 3.65 and no patient in either setting scored 4. As illustrated in Fig. 2, the distributions are near symmetrical although different in total frequency as a result of the different $n$ in each sample.

### Domain scores

We calculated mean scores for each domain to determine whether patients in primary and secondary care settings showed a different profile of scores. Table 4 presents CORE-OM scores by domain for the two service settings, together with effect sizes indicating the degree of difference between populations. Although all effect size differences were ‘small’ (i.e. appreciably below 0.20), secondary patients did report higher levels of risk (effect size $-0.15$). The well-being domain showed the opposite trend, with primary care patients reporting poorer subjective well-being than secondary care patients (effect size 0.08).

### Item scores

We analysed the mean scores for each of the 34 CORE-OM items across the two service settings to establish whether any items appeared to function differently in these patient groups. Comparison of the mean item scores using Cohen’s effect size methodology indicated that secondary care patients scored higher than primary care patients on all four ‘risk to self’ items: item 9 ‘I have thought of hurting myself’ (effect size $-0.14$), item 16 ‘I have made plans to end my life’ (effect size $-0.12$), item 24 ‘I have thought it would be better if I were dead’ (effect size $-0.14$) and item 34 ‘I have hurt myself physically or taken dangerous risks with my health’ (effect size $-0.15$). There was no difference between primary and secondary care patients on the two ‘risk to others’ items: item 6 ‘I have been physically violent to others’ (effect size 0.00) and item 22 ‘I have threatened or intimidated another person’ (effect size $-0.03$). Primary care patients scored higher than secondary patients on item 14 ‘I have felt like crying’ (effect size 0.22) and item 18 ‘I have had difficulty getting to sleep or staying asleep’ (effect size 0.13).

### Application of clinical cut offs

We applied the two cut-off thresholds to the data and Table 5 presents the proportion of patients in each setting above or equal to the cut-off thresholds. Chi-squared tests showed that a significantly higher proportion of primary care patients than secondary care patients were above the clinical cut-off for the well-being domain and non-risk items ($P < 0.01$). However, as noted in the methodology, the statistical power of the data-set was high due to the large $n$, increasing the likelihood of statistical significance for small differences. Odds ratio (OR) analysis showed that secondary care patients were only marginally less likely to be above these cut-offs (OR $= 0.84$ for well-being; OR $= 0.85$ for non-risk items). Secondary care patients were more likely than primary care patients to be above the risk cut-off (OR $= 1.23$, CI 1.10–1.36) and more likely to be above the ‘severe’ threshold (OR $= 1.34$, CI 1.17–1.53).

### Patient-rated CORE-OM severity and presenting problems

We used the practitioner rating provided on the CORE-A form to determine patients’ presenting problems. We classified each problem as present if given any rating by the practitioner from 1 (‘minimal’) to 4 (‘severe’) and absent if no rating was given. Table 6 presents the mean CORE-OM scores for patients grouped by presenting problem. Groups were not independent because many patients were rated as presenting with more than one problem.
Table 6  The CORE–Outcome Measure risk and non-risk scores by presenting problem

| Presenting problem | Items | Primary care | | | Secondary care | | | 95% CI | Effect size |
|--------------------|-------|--------------|| | | | | | | |
|                     | n    | Mean | s.d. | n | Mean | s.d. | | | |
| Depression          | Risk | 3704 | 0.54 | 0.66 | 1360 | 0.66 | 0.74 | -0.16 to -0.08 | -0.18 |
|                     | Non-risk | 3714 | 2.22 | 0.70 | 1364 | 2.23 | 0.77 | -0.05 to 0.04 | 0.00 |
| Anxiety             | Risk | 4010 | 0.47 | 0.63 | 1410 | 0.56 | 0.70 | -0.13 to -0.05 | -0.14 |
|                     | Non-risk | 4016 | 2.13 | 0.73 | 1415 | 2.12 | 0.80 | -0.03 to 0.06 | 0.02 |
| Psychosis           | Risk | 38 | 0.51 | 0.69 | 37 | 0.90 | 0.90 | -0.75 to -0.02 | -0.48 |
|                     | Non-risk | 38 | 2.23 | 0.76 | 37 | 2.40 | 0.67 | -0.50 to 0.15 | -0.25 |
| Personality problems | Risk | 255 | 0.75 | 0.79 | 252 | 1.04 | 0.89 | -0.44 to -0.14 | -0.34 |
|                     | Non-risk | 256 | 2.44 | 0.74 | 252 | 2.45 | 0.74 | -0.13 to 0.12 | -0.01 |
| Cognitive problems  | Risk | 85 | 0.70 | 0.76 | 34 | 0.83 | 0.86 | -0.45 to 0.19 | -0.16 |
|                     | Non-risk | 85 | 2.33 | 0.60 | 34 | 2.26 | 0.65 | -0.17 to 0.32 | 0.12 |
| Eating disorder     | Risk | 147 | 0.59 | 0.68 | 137 | 0.80 | 0.72 | -0.38 to -0.05 | -0.30 |
|                     | Non-risk | 147 | 2.16 | 0.64 | 137 | 2.41 | 0.81 | -0.42 to -0.08 | -0.34 |
| Physical problems   | Risk | 1060 | 0.51 | 0.67 | 352 | 0.61 | 0.74 | -0.19 to -0.02 | -0.15 |
|                     | Non-risk | 1062 | 2.22 | 0.76 | 353 | 2.17 | 0.80 | -0.05 to 0.14 | 0.06 |
| Addictions          | Risk | 275 | 0.89 | 0.84 | 158 | 0.90 | 0.79 | -0.17 to 0.16 | -0.01 |
|                     | Non-risk | 276 | 2.29 | 0.70 | 158 | 2.32 | 0.74 | -0.17 to 0.11 | -0.05 |
| Trauma/abuse        | Risk | 999 | 0.66 | 0.75 | 406 | 0.84 | 0.82 | -0.28 to -0.10 | -0.24 |
|                     | Non-risk | 1002 | 2.33 | 0.72 | 407 | 2.39 | 0.75 | -0.14 to 0.03 | -0.08 |
| Bereavement/loss    | Risk | 1690 | 0.47 | 0.62 | 363 | 0.59 | 0.72 | -0.19 to -0.05 | -0.19 |
|                     | Non-risk | 1694 | 2.15 | 0.72 | 364 | 2.18 | 0.76 | -0.11 to 0.06 | -0.04 |
| Self-esteem         | Risk | 2617 | 0.55 | 0.67 | 895 | 0.67 | 0.75 | -0.17 to -0.06 | -0.17 |
|                     | Non-risk | 2622 | 2.25 | 0.70 | 896 | 2.23 | 0.76 | -0.04 to 0.07 | 0.03 |
| Interpersonal problems | Risk | 2892 | 0.53 | 0.66 | 932 | 0.71 | 0.77 | -0.23 to -0.13 | -0.26 |
|                     | Non-risk | 2898 | 2.19 | 0.71 | 934 | 2.24 | 0.75 | -0.10 to 0.01 | -0.06 |
| Living/welfare      | Risk | 724 | 0.64 | 0.70 | 196 | 0.96 | 0.87 | -0.44 to -0.21 | -0.44 |
|                     | Non-risk | 727 | 2.34 | 0.68 | 196 | 2.44 | 0.76 | -0.20 to 0.02 | -0.13 |
| Work/academic       | Risk | 1090 | 0.48 | 0.63 | 380 | 0.62 | 0.73 | -0.22 to -0.06 | -0.21 |
|                     | Non-risk | 1093 | 2.17 | 0.72 | 383 | 2.14 | 0.80 | -0.06 to 0.11 | 0.04 |

Table 7  Practitioner-rated CORE–Assessment profile of severity* of presenting problems

| Presenting problem | Primary care | | | | Secondary care | | | 95% CI | Effect size |
|--------------------|--------------|| | | | | | | |
|                     | n | Mean | s.d. | n | Mean | s.d. | | | |
| Depression          | 3714 | 2.73 | 0.77 | 1364 | 2.59 | 0.81 | 0.09 to 0.19 | 0.18 |
| Anxiety/stress      | 4016 | 2.84 | 0.77 | 1415 | 2.69 | 0.78 | 0.11 to 0.20 | 0.20 |
| Psychosis           | 38 | 1.79 | 0.96 | 37 | 2.41 | 0.96 | -1.06 to -0.17 | -0.64 |
| Personality problems | 256 | 2.57 | 0.85 | 252 | 2.85 | 0.84 | -0.42 to -0.13 | -0.33 |
| Cognitive/learning  | 85 | 2.16 | 0.88 | 34 | 2.41 | 0.96 | -0.61 to 0.12 | -0.27 |
| Eating disorder     | 147 | 2.23 | 0.94 | 137 | 2.50 | 0.97 | -0.50 to -0.05 | -0.28 |
| Physical problems   | 1062 | 2.66 | 0.86 | 353 | 2.56 | 0.91 | -0.01 to 0.20 | 0.11 |
| Addictions          | 276 | 2.54 | 0.97 | 158 | 2.43 | 1.04 | -0.09 to 0.30 | 0.11 |
| Trauma/abuse        | 1002 | 2.90 | 0.87 | 407 | 2.89 | 0.80 | -0.09 to 0.11 | 0.01 |
| Bereavement/loss    | 1694 | 2.84 | 0.85 | 364 | 2.60 | 0.85 | 0.15 to 0.34 | 0.29 |
| Self-esteem         | 2622 | 2.84 | 0.77 | 896 | 2.74 | 0.80 | 0.05 to 0.17 | 0.14 |
| Interpersonal       | 2898 | 2.82 | 0.79 | 934 | 2.76 | 0.80 | 0.00 to 0.12 | 0.08 |
| Living/welfare      | 727 | 2.60 | 0.83 | 196 | 2.55 | 0.88 | -0.08 to 0.18 | 0.06 |
| Work/academic       | 1093 | 2.74 | 0.83 | 383 | 2.58 | 0.87 | 0.06 to 0.26 | 0.19 |

1. Practitioners rated severity on a scale from 1 (minimal) to 4 (severe).
The effect size analysis in Table 6 shows that CORE-OM risk scores were a key factor in differentiating secondary care patients from primary care patients across the presenting problems. Secondary care patients had higher risk scores than primary care patients for all presenting problems, except addictions where both primary and secondary patients had relatively high mean risk scores. For patients with psychosis, personality problems and eating disorders (problems traditionally seen in specialist services), risk scores were substantially higher in secondary than in primary care (effect size > 0.3). In addition, patients with psychosis, eating disorders and living/welfare problems also showed higher non-risk scores in secondary care than in primary care (i.e. higher levels of overall distress; effect size > 0.1).

**Practitioner-rated CORE–A profile of severity of presenting problems**

We used the CORE–A data to compare the practitioner-rated severity and duration of problems experienced in primary and secondary care settings. Table 7 presents the mean practitioner rating of the severity of the presenting problems. Effect size analysis in Table 7 shows that practitioners rated the severity of anxiety and bereavement higher in primary care than in secondary care settings (effect size > 0.2), but the severity of personality problems, cognitive difficulties, eating disorder and physical problems was rated as higher in secondary care than in primary care settings (effect size >0.2). We were mindful that such differences could reflect differential anchor points in terms of perceptions of problems between practitioners from primary and secondary care settings. Accordingly, we sampled two ranges of CORE–OM scores – a lower range (CORE–OM range 1.00–1.60) and a higher range (CORE–OM range 2.20–2.80) – to check that there were no meaningful differences between primary and secondary care practitioners’ ratings within these ranges. The mean effect size (low and high range) between primary and secondary care practitioners’ ratings for each presenting problem fell below the 0.4 effect size criterion. Table 8 presents the mean rating of duration of the presenting problems in primary and secondary care settings. For all the presenting problems, secondary care patients were rated as having experienced the problem for a greater duration than primary care patients (all effect sizes > −0.2). The greatest difference in problem duration was for psychosis (effect size −0.7).

**DISCUSSION**

The purpose of this article was to investigate the suitability and utility of the CORE–OM and CORE–A for assessing the severity of the presenting problems in primary and secondary care-based psychological therapy services.

**Suitability**

In relation to the appropriateness of these tools in different service settings, the findings show that CORE–OM is acceptable to clients in both settings (as evidenced by high completion rates) and is robust in its structure across different settings (as evidenced by high internal reliabilities), even to the extent of most of the item clusters. However, it is acknowledged that this evidence pertains to counselling and psychological therapy services and could differ in other service settings. In addition, a minority of patients completed their measures at their first session rather than at screening or intake assessment. However, the realities of routine practice settings probably demand reasonable flexibility in the pursuit of maximising compliance in completing the assessment measures.

In administering the same measure in both primary and secondary settings, it might be presumed that the CORE–OM would generate a ceiling effect in secondary care services. We found no evidence of this in the data that we examined. However, we distinguish clearly between patients seen in out-patient settings within secondary care services (as reported here) and patients deemed to be within a category that has been referred to as ‘serious and enduring mental illness’. For such patients, the
process of understanding and completing a
self-report measure might yield results that
are not necessarily continuous with those
reported here (e.g. they might underscore
rather than produce logically higher
scores). However, Whewell & Bonanno
(2000) reported that the risk sub-scale
was ‘clinically valid’ in the CORE–A and
CORE–OM scores matched for patients
with borderline personality disorder.
Where CORE–OM scores might not be
considered safe, the CORE–A form
completed by the practitioner would be
the sole source of information.

Utility
Although we found general heterogeneity
between primary and secondary care
settings in self-rating on the CORE–OM,
there was clear evidence that the CORE–
OM discriminated between patients in
secondary and primary care by showing
them to be more likely to score higher on
risk and be above the severe threshold.
These two components support the ability
of the CORE–OM to discriminate appro-
priately between service settings, a finding
supported by the practitioners’ consistent
reporting of greater duration of patients’
presenting problems in secondary care.
These findings may provide an additional
tool in the recognition by healthcare pro-
fessionals of those patients potentially at
risk of suicide (e.g. Gunnell & Harbord,
2003).

Our data showed primary care patients
to be characterised by more acute problems
(i.e. problems that received a lower dura-
tion rating). The self-severity rating may
be related to the acute nature of the prob-
lems. Item analysis showed this with higher
ratings on the item ‘felt like crying’, which
is likely to reflect the immediacy of the
problems experienced. In contrast, sec-
ondary care patients were characterised by
more chronic problems (i.e. of higher dura-
tion) and higher risk scores on the CORE–
OM. This agrees with the therapist-rated
chronicity of problems in practice settings
of counselling and clinical psychology
(Cape & Parham, 2001). This profile of
patients in secondary services appears to
be a logical consequence of referral pro-
cedures and waiting times. However, we are
mindful that practitioners in primary and
secondary care settings may have differ-
ential anchor points in their evaluation of
the severity of the presenting problems.
When we controlled for patient-rated
severity, we still found at least a 75% over-
lap in the distributions of primary- and
secondary-based practitioners’ ratings.
Notwithstanding this overlap, our view on
this is that practitioner ratings will be
influenced by a myriad of professional
and contextual factors that will require
further research to ensure standard use in
routine settings.

The use of both patient- and
practitioner-completed assessment forms
marks a step forward from reliance on
either patient perception alone or estab-
lished assessment packages using practi-
tioner ratings alone (e.g. Health of the
Nation Outcome Scales; Wing et al,
1998). The use of such data provides a logi-
cal base for benchmarking service delivery
systems (e.g. Barkham et al, 2001) and
adds to a developing literature (e.g. Slade et al,
2001) providing low-cost but reliable
measures that can be adopted routinely in
mental health settings.

ACKNOWLEDGEMENTS
Authors affiliated to the Psychological Therapies
Research Centre were funded by the Priorities and
Needs R&D Levy via Leeds Community Mental
Health Trust.

REFERENCES
The rationale for developing and implementing core
outcome batteries for routine use in service settings
and psychotherapy outcome research. Journal of Mental
Health, 7, 35–47.

Service profiling and outcome benchmarking using
the CORE–OM: toward practice-based evidence in the
CORE–OM. Journal of Consulting & Clinical
Psychology, 69, 184–196.


Psychiatric advance directives: qualitative study of informed deliberations by mental health service users

MICHAELA AMERING, PETER STASTNY and KIM HOPPER

Background Established legal mandates and high expectations for psychiatric advance directives are not matched by empirical evidence documenting their actual implementation.

Aims To explore the interests, concerns and planning activities of informed mental health service users contemplating such directives.

Method Standard qualitative research techniques were used: field observations, interviews, focus groups, archival research and key informant interviews; 33 persons participated in the interviews and focus groups. Transcripts were coded and analysed for thematic content, and results were member-checked.

Results Training set in motion labour-intensive projects: conceptualising how a psychiatric advance directive would work in one’s life, mobilising resources, reviewing past experiences and assessing risks. Especially meaningful was the prospect of being treated as a responsible agent in future interactions with the mental health system.

Conclusions Advance directives are best thought of as complex planning tools for future psychiatric crisis management, rather than focal interventions to enhance compliance. Research is needed to explore the institutional response to this prospective decision-sharing initiative.

Declaration of interest None.

Psychiatric advance directives are legal instruments that allow competent individuals to appoint proxies and specify how treatment decisions should be made in the event they become incompetent. These directives have been hailed as instruments for honouring patient autonomy, improving communication and ensuring participation in treatment decisions, and providing pragmatic alternatives to coercive measures (Appelbaum, 1991; Swanson et al., 2000; Srebnik et al., 2003). Progress towards implementing advance directives as routine planning options has not kept pace with such enthusiasm.

The only randomised effectiveness study of psychiatric advance directives showed no impact on outcome of care with regard to compulsory admissions, hospital days or patient satisfaction (Papageorgiou et al., 2002). However, as with other complex community-based interventions (Schoenwald & Hoagwood, 2001; Ablon & Jones, 2002; Hohmann & Shear, 2002), questions have been raised about actual operational specifics and how they might have influenced the impact of local efforts to implement such directives (Thomas, 2003).

No scientific consensus exists on what the ‘intervention’ is. It is unclear what it means for people with severe and persistent mental illness to undertake the anticipatory planning culminating in an advance directive. This study empirically explores the deliberations of informed mental health service users contemplating this option.

METHOD

Long accepted as a productive partner in public health and evaluation research, qualitative research methods have begun to proliferate in health services research (DeVets et al., 1999), clinical studies (Miller & Crabtree, 1994), health technology assessment (Murphy et al., 1998) and community-based intervention research (Hohmann & Shear, 2002). Such methods are preferred in field settings where the scope of work has yet to be determined, the relevant questions have still to be precisely formulated, local understandings are in flux and institutional arrangements unsettled (Buston et al., 1998). Such conditions aptly describe the state of patients’ training in psychiatric advance directives in New York State late in 2001, when this study commenced.

Formal training in psychiatric advance directives, by the user-run Resource Center, had begun in May 2000. Up to five educators crisscrossed the state, working with official healthcare proxy forms, to recipients of public mental health services. By January 2002 (when recruitment for this study ended), the programme had logged 279 training sessions statewide, involving nearly 6000 participants.

Data collection

Standard qualitative techniques were employed. Field observations were made over a period of 6 weeks by one of the authors (M.A.), who attended eight training sessions. These sessions also served as recruiting sites for potential interviewees and focus group participants. Archival research was undertaken to locate and review relevant legislation and memoranda, annual reports of the training programme and interim evaluations of its work. Key informant interviews were conducted with the training programme director, three trainers, state officials, legislative aides, advocates and peer specialists. Extensive open-ended interviews were conducted by one or two researchers (M.A. and K.H.) with 20 individuals who had received training in the preparation of advance directives at one of three different psychiatric rehabilitation programmes. The aim of our study was to interview ten people who had decided to compile a psychiatric advance directive after participating in the training and ten people who had declined to do so. The latter were recruited from among 88 of 270 trainees who had expressed interest in participating in the study, as were three of those who had decided to draft an advance directive; the other seven people in the advance directive group were recruited through snowball sampling. Another 13 people were recruited for two focus groups. All 33 participants (Table 1)
had extensive experience with mental health services and crisis interventions; most had been hospitalised on several occasions for severe mental illness. Interviews lasted 1–6 h and could take two sessions to complete; the focus groups ran for 1 ½ h. Participants were invited to ‘think out loud’ as they reflected on the training and recalled their subsequent reasoning and actions vis-à-vis psychiatric advance directives. All sessions were tape-recorded and transcribed. Finally, preliminary results underwent provisional member-checking in subsequent discussions with the training project director and six study participants.

After the study had been fully explained, all participants gave written informed consent to the taping of all communications, analyses of transcripts and publication of results and anonymous quotes. Informed consent forms had been approved with the study plan by the New York State Office of Mental Health’s Institutional Review Board.

**Analysis**

Transcripts of interviews and focus group meetings were coded for content following prescribed conventions (Miles & Huberman, 1994; Emerson et al., 1995; Ryan & Bernard, 2000). Data collection (interviewing) and initial analysis (coding) proceeded in a linked fashion. All three investigators closely read transcripts of the initial interviews, identifying topics of interest (open coding) either because they pertained to issues already flagged in the literature (e.g. comprehension, illness and identity, coercion) or because they emerged in the interview as issues of vital concern to the participant (ensuring respect, patching system gaps in record-keeping, customising treatment). These separate readings were then compared and extensively discussed, and a synthetic set of codes was constructed to guide both the conduct of the next interviews and analysis of their transcripts. This process was repeated, modifying the working codebook and recoding earlier sets of transcripts in a directed fashion (focused coding), then testing the most recent codes on fresh transcripts, until a final set of 10 codes (see Appendix) and 24 sub-codes was arrived at; in this way, ‘saturation’ was reached, at which point new data could be accommodated by existing categories. This coding protocol was systematically applied to the entire corpus of transcribed text using software (QSR–N6) to facilitate marking and retrieval of segments. The substantive groupings of text (e.g. knowledge of advance directives; feasibility worries; family and friends) were examined and ‘interrogated’ so that higher-order themes – concepts (Becker, 1998), schemata (Agar & Hobbis, 1985) and low-level ‘theories’ (Strauss & Corbin, 1990) – might be analytically derived. This process was aided by working memoranda that set down provisional interpretations or hypotheses about what we had witnessed. In this way, we attempted to map the contours (or cultural domains) of a still-evolving practice: the implicit rules, organising anxieties, defined tasks and unresolved tensions at stake in these attempts to reckon with the prospects of psychiatric advance directives (Spradley, 1980).

**RESULTS**

Where questions of meaning are vital, varied and contested, qualitative methods serve as exploratory tools. Even commonsensical notions are best approached sceptically. We soon learned, for example, that it would be necessary to rethink the ‘decision’ to execute an advance directive. Interest, understanding, readiness, follow-through, problem-solving and persistence all took shape as complex embodiments of history, present circumstance and informal risk assessment. Executing an advance directive, it became clear, is a project: a sustained act of self-examination and situation appraisal.

After describing the training process, we report results under two headings: the work involved and the meanings construed. To anticipate somewhat: it was the deliberating process itself – a process that can all too conveniently be ‘black-boxed’ in simple outcome studies – that emerged as the defining feature of this inquiry.

**Training**

Eight workshops, involving one or two trainers and 10–100 participants, were observed. All took place as a part of the regular rehabilitation curriculum at participating sites. Written material and actual forms were distributed, a detailed run-through of the process was conducted and questions actively solicited. The last typically concerned the legal and administrative aspects of psychiatric advance directives, their scope and limitations, procedures for revising or revoking them and details regarding appointing an agent. Most of those attending the workshops clearly understood what the directives were about and endorsed their intent. Discussion of probable logistical complications (e.g. lack of a central registry), however, raised doubts about the feasibility of these directives.

**Work**

*Grasping the concept*

Interviewees showed no difficulty in understanding the concept of psychiatric advance directives. The syntax might have been irregular at times – ‘It means you have to go back to the word itself, it means advance is thinking forward, in a moment of center and lucidity and being directive’ – but, as subsequent discussion confirmed, the meaning was clear. Several interviewees and focus group members were already familiar with healthcare proxies through prior experience with kin or friends. Others had learned about them before the training from consumer advocates or presentations in mental health settings.

*Imagining directives in one’s own life*

Notwithstanding the intuitive appeal of psychiatric advance directives, committing to the process that would culminate in a fully executed directive proved difficult
because it meant individualising what one would entail. Participants were drawn to the concept because their wishes would be more likely to be honoured in situations in which they had felt powerless in the past. A designated agent and documented preferences would improve record-keeping and communication in a system widely seen as inadequate in both.

Participants imagined trusted agents undertaking interpretation and negotiation on their behalf: articulating their wishes, translating what clinicians were saying and responding in accordance with the participant’s preferences. Directives held the promise that participants would not have to explain everything yet again, that they would be believed without having to persuade strangers, that confrontations could be avoided through judicious intervention by their proxy and that appropriate treatment would be expedited. Instead of fruitless and increasingly furious exchanges with hospital personnel, one participant imagined producing a card with her proxy’s name and stepping back to let negotiations proceed in a calm and orderly fashion.

No participant viewed advance directives as a blanket means to refuse treatment. None seized upon it to make a point or lodge a grievance. Without exception, the participants were concerned chiefly with improving their treatment, ensuring that their own experience would become part of the ‘evidence base’ on which future decisions would be predicated. The directives were designed to inform clinical teams, and a great deal of thought went into enhancing their instructive value. Interviewees were hopeful, too, that having a directive would alleviate the burden on others by assuming responsibility for treatment in advance of crises.

Imagining advance directives in action could also produce reasons for avoiding them. One interviewee feared that executing a directive could actually invite the situation it was designed to manage. For others, pragmatic issues dominated: ‘more paperwork, more hassles’; concerns that the directive might not ‘stand up in court’ or would be overridden in practice. For many, past experience counselled scepticism, if not outright pessimism: in their or would be overridden in practice. For others, pragmatic issues dominated: ‘more could also produce reasons for avoiding responsibility for treatment in advance of alleviate the burden on others by assuming future decisions would be proceeding in a calm and orderly fashion.

Reviewing past experience
An essential part of the labour was a systematic review of past experiences with psychiatric crises, hospital admissions and medication history. Rehearsing the past, we were repeatedly told, is an exercise fraught with trepidation and pain – but necessary to ‘make sense of your own history’. Contemplating an advance directive can reframe a psychiatric history as a resource from which valuable guides to future system involvement might be gleaned. It can also be risky: memories of traumatic experiences (psychotic crises or difficult treatment situations) can be both upsetting and feared as possible triggers of a new episode. This reinforces the impression that the drafting of advance directives may need to proceed according to the individual’s own timetable, as service users work through their personal risk–benefit ratios.

Mobilising resources
Executing an advance directive means negotiating a host of practical difficulties: securing the technical assistance to complete the documentation, filing the requisite copies at appropriate sites, notifying key collaborators and setting up procedures for periodic review. Appointing the agent proved the decisive issue. Family members were typically the first option, with siblings and aunts or uncles often preferred to parents (seen as too close to serve as advocate for the individual’s wishes). Some candidates were eliminated because the responsibility ‘would be too hard on them’ or because they might be unable to act with the requisite discretion. Peers were considered, together with concerns that they might not be seen as competent by professionals. Some believed that appointing a mental health professional would give them a distinctive edge. A number were disappointed to learn that their own therapists refused because of possible conflict of interest. Even when appropriate candidates were identified, some were reluctant to approach them, fearing rebuff.

Completing the process
Respondents varied greatly in the duration of decision-making. Again, few executed a directive in the immediate aftermath of training. Others took years to mobilise the resources and muster the courage to see the process through. For one interviewee, it had taken a decade-long, wholesale reconfiguration of her personal life before the trusted others were in place. It was not unusual for several versions to be discarded before a satisfactory directive was completed. In all but one executed, the process included appointing an agent.

Although not a specific goal of this study, we did learn much about the content of psychiatric advance directives. Apart from preferences for certain hospitals or specific professionals, requests to allow favoured coping strategies (such as being left alone at times or listening to music) and boundary rules (such as not being touched by staff without being asked) or to proscribe certain drugs or treatments (such as electroconvulsive therapy) were among the detailed instructions. Favoured medications were also named and dosages specified. Reasons for choices, drawn from past experience with ineffectiveness or unwanted side-effects, were documented. Some listed people whose company they preferred and others whose presence they could do without. Much thought was given to ensuring that the advance directives were feasible and that preferences fell reasonably within the range of options of the mental health system. Participants were intent on not compromising mental health professionals with either the content or the style of the directive. It was widely assumed that an ill-advised set of instructions for the appointment of an overly aggressive advocate could backfire, provoking resentment.
and non-compliance on the part of health-care staff.

**Meaning**

**Taking on responsibility**

Committing to the work of drawing up an advance directive was demanding and consequential. For many, this meant accepting a new order of responsibility for managing their illness, containing the trouble crises might cause, securing the allegiance of essential others and thinking through what would constitute a productive response from mental health professionals. It meant taking oneself seriously as an actor in the orchestration of one’s treatment, not merely being its object or victim.

**The ordeal of looking ahead — and back**

Anticipatory planning for psychiatric emergencies is both taxing and unnerving. Undertaking a searching inventory of past experience and present-day resources, facing up to ‘the possibility of losing capacity’ and the humiliation it so often entails, taking pre-emptive action to forestall gratuitous harm and ensure interventions of proven benefit, and asking someone to assume the role of agent: this takes courage and perseverance. Effectively taking the process through to completion was a source of great satisfaction and pride for the participants.

**A doubly instructive effort**

Psychiatric advance directives provide a means of coming to terms with the exigencies of living with severe mental illness and understanding how others figure in that story. Their benefits additionally extend to critically positioned future readers of these directives – mental health professionals in crisis settings – whose knowledge of the directive’s author would be substantially enhanced.

**Holding on to one’s identity**

Making formal arrangements for one’s true self to be present, consulted and represented by a surrogate was a much-prized feature of psychiatric advance directives. The force of the document (‘that they would have to honour it’) resided in its power to transmit the intentions of a self whose capacity calmly to articulate its interests could become degraded at times. In the hectic circumstances of psychiatric crises, the directive would contain and project a corrective competence, stemming the inertial process of becoming ‘just a psychiatric patient’. Here was written evidence that people had ‘thought enough of themselves before coming to the hospital that they put some thought into this’. That made it a ‘thing of respect’.

**A new form of cooperation**

Advance directives serve not only a protective purpose (as a ‘shield’ and ‘buffer’) but a collaborative purpose (as a ‘contract’) as well. In exchange for stipulating their preferred terms and conditions for receiving care, individuals pledge to work with the system, negotiating problems through the intervention of an appointed intermediary.

**DISCUSSION**

Reframing our research concern as deliberation rather than decision with respect to psychiatric advance directives reflected the actual reality of a continuum of interest and action. With few exceptions, those who already had executed an advance directive had undertaken the difficult work that the still-undecided participants were actively engaged in.

**Discordant cases**

Even on their own terms, qualitative methods are incomplete if they fail to account for discordant cases in the textual evidence amassed. In one departure from the pattern, a long-term veteran of the mental health system saw advance directives as completely irrelevant in his own case, but did endorse them for others as protection against abuse he had witnessed. A second exception, occasionally voiced in training sessions, should also be mentioned: the stance of people with so little trust in the system that the prospect of undertaking the work of preparing a directive seemed no more than an elaborate joke.

**Limitations of the study**

Although the small sample size and interview time frame are limitations, the exploratory nature of this study should be stressed. Within the limits of experience to date, this group of informants proved willing and able to ‘think out loud’ about the whole process with researchers in attendance, rehearsing anxieties, misgivings, hopes and their own peculiar brands of risk–benefit accounting. Sufficient resonance across the group was reached that we think it safe to say that the themes emerging in these labour-intensive planning projects are worth pursuing with more rigorous and protracted methods of investigation.

**Interest v. follow-through**

Our findings are congruent with other studies, cited earlier, that documented high levels of interest in psychiatric advance directives among people with a diagnosis of severe mental illness, their families and providers (Amering et al., 1999; Backlar et al., 2001; Swanson et al., 2003; Srebnik et al., 2003). We, too, found that most people easily understood the concept of advance directives and were uninterested in using them to refuse all treatment (Sherman, 1998; Backlar et al., 2001; Valletto et al., 2002). Given the well-documented obstacles to implementing advance directives – lack of information and support for staff, legal uncertainties, want of ongoing assistance and administrative indifference (Howe, 2000; Srebnik & Brodoff, 2003) – the low rates of completion of such directives are not surprising. Discordance between interest and follow-through is not uncommon in studies of preparatory planning: for example, only 16% of healthy, motivated relatives of people with Alzheimer’s disease returned a self-addressed stamped envelope to ratify an earlier willingness to participate in research (Wendler et al., 2002).

**Advance directives as projects**

We found the space between expressed interest and formal execution to be a patchwork of industry. Queried about their thinking and activity after the training sessions, study participants talked about work: sometimes inefficient, occasionally stalled, usually protracted and operating at variable pace. Although a few were flummoxed by the scale of the project, most were not only ready to engage in constructive thinking about their directives, but also willing to take on more responsibility for treatment preparedness than ever before. This meant substantial activity and personal risk. What sustained the work and made the risks worth taking was the promise of a new relationship with the mental health system, one built on collaboration and respect for their own experience-based expertise. Accordingly, people were careful
to structure their directives so that professionals could respond constructively.

**System receptivity**

As psychiatric advance directives slowly make their way into the clinical landscape, the mental health system faces a formidable challenge. If it is not met, the offer of advance directives risks becoming a charade, instead of the good-faith invitation to dialogue that most of our informants interpreted it to mean. Implementing new procedures, even ones acclaimed in principle, is invariably experienced as cumbersome and artificial at the outset. A good deal of trial and error is needed before workable, culturally congruent practices are arrived at. Our findings suggest that sufficient deliberative time, and the resources to support and inform decision-making, will be crucial to the implementation of advance directives. ‘Pushed timing’ undermines voluntarism, degrades capacity to make a considered choice tailored to one’s life circumstances and values, and nullifies the ethic of respect that motivates shared decision-making (Roberts, 2002). Peer advocates may well be the ideal training mentors (Backlar et al. 2001), but the protracted and unpredictable nature of deliberation makes a standing offer of technical support advisable.

Psychiatric advance directives are best conceived as ‘complex interventions that require lengthy development work if they are to stand a chance of success’ (Thomas, 2003). The development work includes not only patients, family members and clinicians, but also the system itself. A host of legal, procedural and logistical problems, well described in the literature on healthcare proxies (Slewchuk, 1998), remain to be addressed. Resolving them will require substantial investment of resources to effect the necessary changes in administrative infrastructure and professional culture, at a time when the moral economy of care appears to be heading in a number of contrary directions (Hopper, 2001). Advance directives hold great potential to engage consumers in a serious and balanced collaborative process, but without a serious systemic commitment this bold experiment may prove a cautionary tale instead.

**ACKNOWLEDGEMENTS**

This research was funded by the Center for the Study of Issues in Public Mental Health, based at the York State Office of Mental Health. The authors thank Melanie Shaw for extensive discussions of earlier drafts.

Nathan S. Kline Institute for Psychiatric Research, Orangeburg, New York, and by ongoing support of two of the investigators (P.S. and K.H) by the New York State Office of Mental Health. The authors MICHAELA AMERING, MD, Department of Psychiatry, University of Vienna, Austria; PETER STASTNY, MD, Bronx Psychiatric Center, New York; KIM HOPPER, PhD, Nathan S. Kline Institute for Psychiatric Research, New York, USA; KIM HOPPER, PhD, Nathan S. Kline Institute for Psychiatric Research, New York, USA; e-mail: hopper@nki.rfmh.org

(First received 16 December 2003, final revision 16 September 2004, accepted 29 September 2004)

**CLINICAL IMPLICATIONS**

- Understanding how (and how well) psychiatric advance directives work will require close documentation of the processes involved in implementing them.
- Intensive training and attentive follow-through would greatly aid the deliberating activities of people contemplating executing psychiatric advance directives.
- Substantial administrative commitment on the part of the mental health system is the necessary complement to expressed patient interest and preparatory action.

**LIMITATIONS**

- Findings of our study may not be generalisable because the sample size was small and participants were not chosen randomly.
- Full documentation of deliberations over advance directives would have required interviews and observations over a longer period, as the directives were tested and changes made.
- Not enough people who had decided against executing an advance directive were interviewed to determine how they differed from those actively considering one.

**APPENDIX**

**Text coding categories**

**Knowledge**

Informant’s knowledge about advance directives, whether gained through training or experience.

**Feasibility**

Concerns about the logistics of advance directives, including legal status, mechanics of use and the system’s capacity to respond.

**Experience**

Experience with the mental health system — positive and negative — and especially with coercion.

**Anxiety**

Fears raised with respect to drafting an advance directive thinking about past, future, oneself as agent.

**Intent**

Changes the informant hopes to make through the vehicle of an advance directive.

**Unintended changes**

Changes the informant fears might ensue because of an advance directive.

**Content**

Actual content of an advance directive.

**Action**

Comments relating to deciding on or taking action, or barriers to action, in drafting an advance directive.
Editorial
Expressed opinions, ideological statements.

Uncoded
Material not otherwise coded.

REFERENCES


Revised self-poisoning: increasing severity of self-harm as a predictor of subsequent suicide

GREG CARTER, DAVID M. REITH, IAN M. WHYTE and MICHELLE McPHERSON

Background Prediction of suicide risk is difficult in clinical practice.

Aims To identify changes in clinical presentation predictive of suicide in patients treated for repeated episodes of self-poisoning.

Method A nested case–control study used the Hunter Area Toxicology Service database to identify exposure variables and the National Death Index to identify suicide. Cases were patients who had hospital treatment on more than one occasion between 15 January 1987 and 31 December 2000.

Results There were 31 cases, for which 93 controls were selected. Study variables associated with an increased risk of subsequent suicide were an increase in the number of drugs ingested (odds ratio 2.59, 95% CI 1.48–4.51), an increase in the dose ingested (OR 1.33, 95% CI 1.01–1.76), an increase in coma score (OR 1.71, 95% CI 1.11–2.66), a decrease in Glasgow Coma Score (OR 1.21, 95% CI 1.03–1.43) and an increase in drug or alcohol misuse (OR 2.33, 95% CI 1.06–5.10).

Conclusions Patients who have escalating severity of self-poisoning episodes are at high risk of completed suicide.

Declaration of interest None.
95% confidence intervals (CIs) were calculated using conditional logistic regression for matched case-control groups using Stata (Stata, 2003). The grouping variable was the age and gender strata for the cases and the matched controls.

The independent variables studied were similar to those used as indicators of medical severity in previous studies of outcome and comparative toxicity in self-poisoning (Reith et al., 2004):

(a) indicators of medical seriousness: intensive care admission, length of stay in intensive care, overall length of stay, presence of seizures, Glasgow Coma Scale (GCS; Teasdale & Jennet, 1974) score on presentation and coma scale (Plum & Posner, 1972) on presentation;

(b) indicators of serious intent: number of tablets ingested, total dose ingested in defined daily doses (Capella, 1997), number of different medications ingested, time from overdose to presentation and choice of poisoning method (carbon monoxide v. medications);

(c) changes in drug and alcohol status (medical staff ratings in the emergency department);

(d) type of poisoning, psychiatric diagnosis (new diagnosis on subsequent presentation) and new occurrence of involuntary psychiatric admission or absconding.

The variables that were associated with subsequent suicide were then tested for their clinical usefulness as predictors of suicide (Sacket, 1992). Continuous variables were assessed using receiver operating characteristic (ROC) plots, and the cut-off points that resulted in the greatest proportion of correct classifications (i.e. patients correctly classified as ‘suicide’ or ‘not suicide’ by the test) were used to generate dichotomous variables. These variables were then assessed for their suitability as predictors by calculating sensitivity, specificity and the respective 95% confidence intervals using Stata. Positive predictive values and negative predictive values were not calculated because these variables would have been biased by the pre-test probabilities of the sample being affected by the case-control design.

RESULTS

There were 34 patients who presented on two or more occasions and subsequently died by suicide. For three of these patients death occurred during the last admission (all from medicinal poisoning) and these cases were excluded from the analysis. This resulted in 31 cases, for which 93 controls were selected (Table 1). For the cases, the median time from last admission to suicide was 305 days (range 4–2636). Nine of the controls died during the study period: two from cardiovascular causes, one from respiratory causes, one from endocrine causes, one from hypotensive dependence, one from opioid dependence, one from malignancy and one from accidental poisoning; for one patient the cause of death was unknown.

The indicators of medical seriousness associated with subsequent suicide were an increase in coma score and a decrease in GSC score (both indicating greater degrees of coma) in the cases compared with the controls (Table 2). As indicators of intent, the number of medications, number of tablets and total dose ingested increased from first to last visit in the cases, but remained stable or decreased in the controls. This indicated a significantly increased poison exposure from first to last presentation in the case group relative to the control group. There was also a worsening in drug and alcohol status in the case group compared with the control group. There was no significant change in time to presentation, nor in intensive care unit admission or length of stay. There was no significant change in the patterns of poison exposure. There was no significant difference in change in length of stay, psychiatric diagnosis or discharge destination between the groups.

When the significant variables were examined for their sensitivity and specificity as tests for patients who would subsequently kill themselves, none was sufficiently useful to be used alone as a screening test for subsequent completed suicide (Table 3). The most promising predictor variable was the change in the number of tablets ingested, with an area under the ROC curve of 0.73 (95% CI 0.59–0.88) (Fig. 1). An increase of 70 or more in the number of tablets ingested had a high specificity, and the best sensitivity of any individual test (Table 3). Combining this with deterioration in drug and alcohol misuse status increased the sensitivity to 47%. However, combining an increase of 70 or more in the number of tablets ingested with a decrease in GCS score of 2 or more resulted in the best combined test, with a sensitivity of 53% and a specificity of 87%. When tested for their association with subsequent suicide, as a post hoc analysis, the odds ratio for a 70 or more increase in tablets ingested was 3.59 (95% CI 0.98–3.15), for a two or more increase in number of drugs ingested was 3.60 (95% CI 1.03–12.53) and for a decrease in GCS score of 2 or more was 5.36 (95% CI 1.34–21.53).

DISCUSSION

Methodological issues

Some of the limitations of our study include the number of deaths in the control group, the validity of the resident assessment of drug and/or alcohol misuse and the difference in follow-up time between the cases and controls. Some of the deaths in the control group might have been misclassified suicides; the resultant bias would be in the direction of a negative result (type 2 error) and hence would not affect the positive findings of the study. However, other potential risk factors such as length of stay in hospital and discharge to an involuntary psychiatric admission or absconding might have been incorrectly found not to be associated with suicide. The longer mean follow-up time in the control group is expected, because the deaths of those in the case group would have limited the follow-up period. The longer follow-up time in the control group would also be expected to lead to a negative result (type 2 error), and would not have affected the positive

### Table 1 Characteristics of patients in the case and control groups at first presentation

<table>
<thead>
<tr>
<th></th>
<th>Cases (n=31)</th>
<th>Controls (n=93)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years: median (range)</td>
<td>26 (14–70)</td>
<td>29 (12–81)</td>
</tr>
<tr>
<td>Gender male/female, n/n</td>
<td>22/9</td>
<td>66/27</td>
</tr>
<tr>
<td>Deliberate self-poisoning, n (%)</td>
<td>29 (94)</td>
<td>84 (90)</td>
</tr>
<tr>
<td>Died during study, n</td>
<td>31</td>
<td>9</td>
</tr>
<tr>
<td>Follow-up, days: median (range)</td>
<td>970 (170–4409)</td>
<td>2579 (18–5102)</td>
</tr>
</tbody>
</table>
Table 2  Characteristics at first and last hospital-treated episodes and odds ratios for change from first to last for subsequent suicide

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>First admission</th>
<th>Last admission</th>
<th>OR (95% CI)</th>
<th>Change from first to last</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Case</td>
<td>Control</td>
<td>Case</td>
<td>Control</td>
</tr>
</tbody>
</table>

Indicators of medical seriousness

- ICU admission, n (%)  
  - Case: 5 (16.13)  
  - Control: 16 (17.2)  
  - OR: 1.75 (0.55–5.56)

- Length of ICU stay, h: mean (range)  
  - Case: 43 (22.5–75.22)  
  - Control: 23.4 (12.67–45.95)  
  - OR: 4.07 (0.34–48.23)

- Length of hospital stay, h: mean (range)  
  - Case: 41 (6–166)  
  - Control: 50 (1–1,222)  
  - OR: 1.00 (1.0–1.01)

- Seizure, n (%)  
  - Case: 1 (3)  
  - Control: 2 (2)  
  - OR: 4.07 (0.34–48.23)

- Decrease in GCS, score: mean (range)  
  - Case: 14.3 (9–15)  
  - Control: 13.5 (3–15)  
  - OR: 1.21 (1.03–1.43)

- Increase in coma score: mean (range)  
  - Case: 0.56 (0–2)  
  - Control: 0.96 (0–6)  
  - OR: 1.71 (1.11–2.66)

Indicators of intent

- Number of defined daily doses  
  - Case: 21 (12^30)  
  - Control: 100 (100^100)  
  - OR: 1.00 (1.0^1.01)

- Length of ICU stay, h: mean (range)  
  - Case: 43 (22.5–75.22)  
  - Control: 23.4 (12.67–45.95)  
  - OR: 1.00 (1.0–1.01)

- Seizure, n (%)  
  - Case: 1 (3)  
  - Control: 2 (2)  
  - OR: 4.07 (0.34–48.23)

- Decrease in GCS, score: mean (range)  
  - Case: 14.3 (9–15)  
  - Control: 13.5 (3–15)  
  - OR: 1.21 (1.03–1.43)

- Increase in coma score: mean (range)  
  - Case: 0.56 (0–2)  
  - Control: 0.96 (0–6)  
  - OR: 1.71 (1.11–2.66)

- Time since previous admission, days: mean (range)  
  - Case: 761 (3–4000)  
  - Control: 1.00 (1.0^1.0)

- New rating of lifetime drug or alcohol misuse  
  - Case: 29 (21^37)  
  - Control: 86 (80^92)

Medical officer ratings in the emergency department

- Poisoning type: deliberate self-poisoning, n (%)  
  - Case: 29 (93)  
  - Control: 86 (80^92)

- Ingestion of antidepressants, n (%)  
  - Case: 9 (29)  
  - Control: 40 (43.5)

- Ingestion of sedatives, n (%)  
  - Case: 1 (3.2)  
  - Control: 4 (4.3)

- Ingestion of cardiac drug, n (%)  
  - Case: 0  
  - Control: Not performed

- Time since previous admission, days: mean (range)  
  - Case: 452 (21–2003)  
  - Control: 76 (3–4000)  
  - OR: 1.00 (1.0–1.0)

Psychiatric diagnosis and discharge status

- Involuntary psychiatric admission or abscording, n (%)  
  - Case: 5 (16)  
  - Control: 20 (21)

- Diagnosis of mood disorder, n (%)  
  - Case: 5 (16)  
  - Control: 20 (21)

- Diagnosis of personality disorder, n (%)  
  - Case: 4 (13)  
  - Control: 14 (15)

- Diagnosis of substance disorder, n (%)  
  - Case: 2 (6)  
  - Control: 26 (28)

- Diagnosis of psychosis, n (%)  
  - Case: 3 (10)  
  - Control: 6 (6)

GCS, Glasgow Coma Scale; ICU, intensive care unit; NA, not available.
1. Number of defined daily doses.
2. OR, odds ratio; 95% CI, 95% confidence interval.

Table 3  Sensitivity and specificity (at optimal cut-off points) of changes in clinical characteristics for predicting suicide

<table>
<thead>
<tr>
<th>Clinical characteristic</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
</tr>
</thead>
</table>

Indicators of medical seriousness

- Decrease in GCS score of 2 or more  
  - Sensitivity: 24 (16–32)  
  - Specificity: 93 (88–98)

- Involuntary psychiatric admission or abscording, n (%)  
  - Sensitivity: 5 (16)  
  - Specificity: 20 (12–28)

- Diagnosis of mood disorder, n (%)  
  - Sensitivity: 5 (16)  
  - Specificity: 20 (12–28)

- Diagnosis of personality disorder, n (%)  
  - Sensitivity: 4 (13)  
  - Specificity: 14 (15)

- Diagnosis of substance disorder, n (%)  
  - Sensitivity: 2 (6)  
  - Specificity: 26 (28)

- Diagnosis of psychosis, n (%)  
  - Sensitivity: 3 (10)  
  - Specificity: 6 (6)

Psychiatric diagnosis and discharge status

- New rating of lifetime drug or alcohol misuse  
  - Sensitivity: 29 (21–37)  
  - Specificity: 86 (80–92)

Compared to

- Two or more increase in number of drugs ingested  
  - Sensitivity: 40 (30–50)  
  - Specificity: 91 (85–96)

- 70 or more increase in the number of tablets ingested  
  - Sensitivity: 47 (36–58)  
  - Specificity: 76 (66–86)

DA, drug and alcohol; DDDs, defined daily doses; GCS, Glasgow Coma Scale.
1. 95% confidence intervals.
findings of the present study. The medical officer’s rating of lifetime drug or alcohol misuse, when previously compared with the gold standard of DSM–IV (substance abuse) as rated by clinical psychiatric assessment, had a sensitivity of 91% and a specificity of 60% (Dawson, 2000). The medical officer’s rating of lifetime drug or alcohol misuse is probably a broader measure of substance exposure or misuse than the DSM diagnosis of substance-related disorder, which was also used in this study. Hence, the medical officer’s assessment of lifetime substance misuse, although readily performed, does not correspond to the DSM–IV criteria. The nested case-control design used in our study, unlike some other case-control designs, was not affected by ascertainment or recall bias because the Hunter Area Toxicology Service treats all self-poisoning patients from the region and all of the exposure variables were collected prospectively.

Risk factors for suicide after self-harm

People who deliberately harm themselves have an increased risk of suicide (Owens et al, 2002). Previously identified risk factors for subsequent suicide following deliberate self-harm include previous self-harm, male gender, older age, psychiatric illness (particularly schizophrenia, depression, bipolar disorder and substance-related disorders), medical illness and substance misuse (Suokas et al, 2001; Beautrais, 2003). Specifically following deliberate self-poisoning, identified additional risk factors for completed suicide include psychiatric disorders of childhood, male gender, increasing age, more than one previous suicide attempt, living alone, migrant status and being widowed or separated (Reith et al, 2004).

Our study showed that an increase in some markers of the severity of self-poisoning episodes was associated with subsequent death by suicide. The indicators of increased severity were indicators of potential physical harm (such as coma score) and increased severity of poison exposure. An increase in ingested dose of 70 or more tablets or capsules, an increase of two or more in the number of different agents ingested and an increase of 50 or more in the number of defined daily doses ingested were highly specific for subsequent suicide. These indicators had much greater specificity than previously identified indicators of future suicide such as Beck’s Hopelessness Scale: 51% for hospitalised patients with suicidal ideation and 41% for psychiatric out-patients (Beck et al, 1985, 1990). However, the sensitivity of Beck’s Hopelessness Scale was greater: 91% for hospitalised patients with suicidal ideation and 94% for psychiatric out-patients (Beck et al, 1985, 1990). Hence, although instruments such as Beck’s Hopelessness Scale may correctly identify those patients who subsequently die by suicide, but at the expense of also incorrectly identifying many who will not (Beck et al, 1985, 1990), our approach would not identify a large proportion of subsequent suicides.

Generalisability

The demographic characteristics and long-term outcomes of the patients treated by the Hunter Area Toxicology Services are similar to some populations in the UK (Hawton et al, 2003). The characteristics of the patients treated for repeated self-harm and their subsequent long-term outcomes are also similar (Zahl & Hawton, 2004). In addition, the factors found to be predictive of suicide in our study (number of tablets or different drugs ingested, and Glasgow Coma Scale score) can easily be measured and recorded by non-psychiatrists. The GCS is widely used internationally and would be expected to be part of the routine management of a patient presenting with self-poisoning. Hence the results of the study can be applied even to units where the clinicians have no special interest in self-harm. If the rate of suicide is known in the population, the individual’s risk of subsequent suicide can be estimated using likelihood ratios derived from this study (Sacket, 1992).

Interventions to prevent subsequent suicide

People admitted to hospital for treatment of self-poisoning constitute a population with a greatly increased risk of completed suicide (Owens et al, 2002), and within this group patients presenting on subsequent occasions with escalating severity of poisoning may be at higher risk. Hence interventions designed to prevent suicide or to reduce the repetition of self-poisoning could be tested in this population. A few interventions have demonstrated a decrease in rates of suicide – a letter-writing intervention (Motto & Bostrom, 2001) – or in repetition of self-harm: dialectical behaviour therapy, for chronically parasuicidal women meeting criteria for borderline personality disorder (Linehan et al, 1991); psychoanalytically informed partial hospitalisation, for people with borderline personality disorder who harm themselves (Bateman & Fonagy, 1999); a brief interpersonal therapy intervention for hospital patients admitted for deliberate self-harm (Guthrie et al, 2001); and depot flupentixol (Montgomery et al, 1979).

Although low cost interventions may be applied in all cases of deliberate self-poisoning, high-cost interventions may need to be restricted to high-risk subgroups. In order to apply such interventions it is first necessary to be able to identify patients...
at increased risk, without identifying an unnecessarily large number of people who are not at risk. Using the predictor variables identified from this study for those presenting with at least two episodes of self-poisoning would identify around half of those at risk (long-term) of death by suicide. To do this would require accurate data collection and the ability to compare sequential admissions. This process can be achieved using an electronic database, which could automatically identify patients at high risk of completed suicide. A model for such a system currently exists, where a clinical database could be used to inform psychiatry services of high risk patients (Whyte et al., 1997).

REFERENCES


CLINICAL IMPLICATIONS

- Patients whose repeated self-poisoning episodes are of escalating severity are at increased risk of completed suicide.

- These patients may warrant additional short-term and long-term attention from clinical services in order to reduce their long-term risk of subsequent suicide.

- Any long-term system of monitoring and treating self-poisoning patients might include these markers of increased risk.

LIMITATIONS

- Some of the deaths in the control group might have been misclassified suicides.

- The resident assessment of drug and/or alcohol misuse differs from the DSM–IV substance abuse criteria.

- There was a shorter follow-up time for cases than for controls.
Risk of puerperal and non-puerperal recurrence of illness following bipolar affective puerperal (post-partum) psychosis

EMMA ROBERTSON, IAN JONES, SAYEED HAQUE, ROGER HOLDER and NICK CRADDOCK

Summary  The clinical value of information on the risk of future psychiatric illness in women who have experienced puerperal (post-partum) psychosis has been limited by inconsistencies in terminology and nosology. Here we report rates of subsequent puerperal and non-puerperal episodes, in a well-characterised sample of women diagnosed with clearly defined bipolar affective puerperal psychosis (n=103). Out of 54 women having further children, 31 (57%; 95% CI 44–69) experienced an additional puerperal psychotic episode, and 64 of 103 women (62%; 95% CI 52–71) experienced a non-puerperal affective episode during the follow-up period (mean duration 9 years). A history of bipolar episodes prior to the puerperal psychosis did not predict risk following subsequent pregnancies, but positive family history of mental illness predicted shorter time to non-puerperal relapse.

Declaration of interest  None. Funding detailed in Acknowledgements.

Puerperal psychosis is an abrupt onset of severe psychiatric disturbance that occurs shortly following parturition in approximately 1–2 per 1000 deliveries. Despite wide variations in details of definition, it is known that most cases represent triggering by childbirth of episodes of bipolar disorder (Chaudron & Pies, 2003). Strikingly, up to a half of parous women with a lifetime diagnosis of bipolar disorder develop an episode of puerperal psychosis in the period immediately following childbirth (Brockington, 1996; Jones & Craddock, 2001). Such episodes usually require hospitalisation and are associated with substantial functional impairment and risk both to the woman herself and, in rare but tragic cases, to her newborn child.

Unfortunately, inconsistencies in terminology and nosology often result in a failure to provide patients with the information they need to make important decisions about family planning and illness management (Robertson & Lyons, 2003). In this short report we quantify the rates of puerperal and non-puerperal recurrences in a large sample of women diagnosed with clearly defined bipolar affective puerperal psychosis, and provide evidence that a simple clinical variable – family history – may be prognostically useful.

METHOD

The study group comprised 103 women, all UK residents, who had experienced at least one episode of puerperal psychosis. Their mean age at interview was 40 years (s.d.=8). After giving written informed consent, all participants were interviewed by a trained investigator using the Schedules for Assessment in Neuropsychiatry (SCAN; Wing et al, 1990) and case-note information was obtained. Best-estimate episode and lifetime diagnoses were made on the basis of all available clinical information by two independent investigators. The family history interview of the Research Diagnostic Criteria (Spitzer et al, 1978) was used to elicit family histories of mental illness. Each participant’s illness history from the first episode of puerperal psychosis (mean age 28 years, s.d.=4.8) was studied in depth and the number and timing of episodes of bipolar illness recorded. Full details of the clinical method are provided by Robertson et al (2000).

RESULTS

All participants had experienced at least one episode of bipolar affective puerperal psychosis, defined as onset of a manic or psychotic episode within 4 weeks of childbirth, and all had a lifetime best-estimate DSM-IV (American Psychiatric Association, 1994) diagnosis of bipolar disorder (n=90) or schizoaffective disorder, bipolar type (n=13). Median follow-up time from recovery from the puerperal episode to interview was 9 years (range 6 months to 33 years). Thirty women (29%) had a previous psychiatric history (defined as DSM-IV major depression, mania or hypomania before the puerperal episode) and 59 (57%) had a positive family history of psychiatric illness, defined as the participant reporting at least one first- or second-degree relative diagnosed with or treated for a psychiatric disorder.

Recurrence rates of puerperal psychotic episodes

Fifty-four participants had a subsequent delivery, of whom 31 (57%; 95% CI 44–69) experienced another episode of puerperal psychosis, and an additional 5 (9%; 95% CI 4–20) experienced an episode of mania, depression or psychosis during pregnancy or within 6 months (but not 6 weeks) of delivery. Using contingency table analysis, neither family history nor personal history of psychiatric illness was a significant predictor of puerperal recurrence in this sample. Of the 39 women for whom the index episode of puerperal psychosis was their first episode, 22 (56%) experienced a further episode following their subsequent delivery, compared with 8 of 15 women (53%) who had experienced other episodes of illness prior to the initial puerperal psychosis (χ²=0.04, d.f.=1, P=0.84).

Risk of further non-puerperal episodes of illness

Following the index episode of puerperal psychosis, 64 participants (62%; 95% CI 52–71) experienced at least one non-puerperal affective episode (DSM-IV mania, depression or hypomania) during the period of observation. Because of differing duration of follow-up, Kaplan–Meier survival curves were used to examine the influences of personal history and family history of psychiatric illness on time to non-puerperal relapse. A shorter time to non-puerperal recurrence was associated significantly with a positive family history of mental illness (mean survival 4 years ν. 7 years; log-rank statistic 6.53, d.f.=1, P<0.01; Fig. 1) and non-significantly with
previous personal history of illness (mean survival 4 years × 6 years; log-rank statistic 1.48, d.f. = 1, \( P = 0.22 \)).

**DISCUSSION**

Our findings are consistent with – and extend – previous research that used a wider phenotypic definition of post-partum psychosis, in finding high rates of recurrence of both puerperal and non-puerperal episodes of major mood disorder (Kirpinar et al., 1999; Terp et al., 1999; Robling et al., 2000). We have quantified these risks in a sample of women with clearly defined bipolar affective puerperal psychosis. We found the rates of recurrence following further deliveries were considerably higher than the rates we had reported for women with bipolar disorder in general (26% of deliveries in familial bipolar disorder; Jones & Craddock, 2001). However, we found no evidence that women whose puerperal psychosis is the first episode of illness have a different risk following subsequent deliveries than women who had previously experienced non-puerperal episodes.

We also provide data regarding the time course of risk for non-puerperal recurrences and evidence that family history may be a useful predictor regarding the timing of risk. The latter finding requires replication in independent samples before it can be regarded as a robust prognostic predictor.

**Clinical relevance**

Our findings have clinical relevance for the management of women who have experienced or are at risk of an episode of bipolar affective puerperal psychosis.

**Family planning**

It is vital to be aware of the high risk of puerperal recurrence, but avoiding further pregnancy (as has often been advised in the past) is not a guarantee of avoiding further illness. Many women in our sample reported that they were not made aware of the substantial risks of non-puerperal episodes of illness and made ill-informed reproductive decisions as a consequence. Moreover, we found no evidence to suggest that women who have only experienced a puerperal episode should be considered at higher risk of further post-partum episodes than women who had also had non-puerperal episodes.

**Prophylaxis**

Although lithium is an effective prophylactic medication in bipolar disorder for many patients, it must be taken regularly, has a narrow therapeutic window, several undesirable adverse effects and is teratogenic to the foetus. Other agents used in prophylaxis – such as sodium valproate or carbamazepine – have similar properties. Decisions regarding prophylaxis of bipolar disorder in women of childbearing age require very careful weighing up of risks and benefits, need to be based on robust evidence, and should be made jointly with the patient. Our data will inform this situation and suggest that a simple clinical predictor (family history) may help to individualise the risk assessment.

**ACKNOWLEDGEMENTS**

We thank all the women who participated in this study. This work was supported by grants from the Welcome Trust, the West Midlands Regional Health Authority Research & Development Directorate, South Birmingham Mental Health National Health Service Trust and the Women’s Mental Health Trust. E.R. was a New Blood Research Fellow and I.J. was a Welcome Trust Training Fellow in Mental Health at the time of the study.

**REFERENCES**


Suicide by hanging: multicentre study based on coroners’ records in England

OLIVE BENNEWITH, DAVID GUNNELL, NAVNEET KAPUR, PAULINE TURNBULL, SUE SIMKIN, LESLEY SUTTON and KEITH HAWTON

Summary We studied 162 cases of hanging by suicide occurring in 24 coroners’ jurisdictions in England within a 6-month period in 2001. Prison and psychiatric ward suicides accounted for only 6% of these. The most frequently used ligatures (ropes, belts and cable) and ligature points (beams, girders, lofts and trees) are commonly available in community settings, limiting opportunities for prevention. In only half the cases (52%) were victims fully suspended with both feet off the ground. Four per cent had also taken an overdose.

Declaration of interest None.

Hanging is the most common method of suicide in England and Wales, accounting for about 2000 deaths each year (Department of Health, 2002; Brock & Griffiths, 2003). Previous studies have been confined to specific geographical locations and have not comprehensively investigated potentially preventable aspects of these deaths (Bowen, 1982; Davison & Marshall, 1986; James & Silcocks, 1992).

In this short report we describe a large case series of hanging suicides from a wide geographical area focusing on the possibilities for restricting access to means, which is one of the objectives of the National Suicide Prevention Strategy for England (Department of Health, 2002).

METHOD

We studied hanging suicides occurring within the jurisdictions of 24 coroners between 1 July 2001 and 31 December 2001. We initially approached the coroners serving the 3 research centres (Oxford, Bristol and Manchester). A further 21 coroners, 7 from each centre, were randomly selected from a list of coroners within 50 miles (or 1.5 hours’ travelling time) of each centre. Three coroners, from 2 centres, did not agree to participate; they were replaced with the next randomly selected coroners in the relevant centres. Five of the jurisdictions were urban, 15 mixed urban/rural and 4 rural. All deaths by hanging or self-strangulation given a suicide or open verdict were examined.

We collected demographic data, information on the timing and location of the act, who discovered the person, whether they were alive when found, whether alcohol or drugs had been consumed (including details of toxicology reports), contact with psychiatric services, whether the person was suffering from a psychiatric disorder at the time of death and their history of self-harm. Psychiatric diagnoses were based on both general practitioner and psychiatric reports. Where the individual had not had contact with a psychiatrist, diagnoses were formulated on the basis of witness accounts. Information on the ligature, ligature point and degree of suspension was recorded.

RESULTS

One hundred and sixty-two cases of hanging (85.8% males) were identified across the 24 districts. A verdict of death by suicide was returned in nearly all cases (155 out of 162, 95.7%). The remainder (4.3%) were recorded as open verdicts. Seven individuals (4.3%) had engaged in simultaneous self-poisoning.

Demographic characteristics
Mean ages were similar in males (40.6 years) and females (42.2 years). Eleven individuals (6.8%) were psychiatric inpatients and 5 (3.1%) were prisoners. Only 5 (3.1%) of the psychiatric in-patients died on the ward.

Contact with psychiatric services and psychiatric diagnosis
Information on contact with psychiatric services was available for 117 individuals (72.2%). Forty-eight of these (41.0%) were in contact with a psychiatric service at the time of death.

Among the 128 cases where an assessment of psychiatric disorder was possible, more than half (71 out of 128, 55.5%) had a primary diagnosis of affective disorder and 13.3% had schizophrenia (17 out of 128). Information on past self-harm was recorded for most individuals (152 out of 162, 93.8%). Nearly half of these had previously self-harmed (68 out of 152, 44.7%).

Location of death and discovery
In two-thirds of cases (106 out of 162, 65.4%) the person hanged themselves at their home, either indoors or in their garden, shed or garage. In 27 cases the death occurred outside in a public area. Sixty-eight individuals (42%) were found by a family member or partner. In 7 of the 162 cases (4.3%) the individual was found alive and was taken to hospital.

Ligatures used for hanging
Where this was recorded (98.8% of cases) the main ligatures used were rope or cord (49.4%), belt (13.1%), electric cable (11.9%), and dog lead (6.3%). Items of clothing other than a belt (scarf, tie, dressing gown cord, shoe-lace) were used in one-tenth of hangings (16 out of 160, 10.0%). Information on the source of the ligature was available for only 73 cases (45.1%). In most (63 out of 73, 86.3%) this had been in the household at the time.

Suspension points and type of suspension
Rooftops and ceilings (beam/girder or loft) were used as ligature points in about one-third of hangings (58 out of 162, 35.8%) (see Table 1). Of the outdoor ligature points, a tree was most commonly used (25 out of 162, 15.4%). Information on degree of suspension was available for 149 cases. Seventy-eight of these (52.4%) were found totally suspended (both feet above the ground). In about one-quarter (35 out of 149, 23.5%) the subjects were suspended with their feet touching the ground, in 11 cases (7.4%) they were kneeling, in 13 lying (8.7%)
and in 7 seated (4.7%). The precise position was unclear in a further 5 cases involving partial suspension (3.4%).

**Hanging suicides in prison and psychiatric wards**

All prison hanging suicides (n=5) took place in the prison cell. In 4 cases the ligature point used was the cell bars/window and in one case the toilet door. Sheeting (usually torn) was the ligature in 4 cases.

Three of the 5 hanging suicides which occurred on psychiatric wards used a belt and in one case a dressing gown cord. Ligature points used were a radiator fitting, a pipe (in the bathroom), part of the bed (metal bedhead), a wardrobe handle and the bed curtain rail.

**DISCUSSION**

The most common ligatures used in this sample (rope, belts and flex) are similar to those found in previous community-based studies (see, for example, Davison & Marshall, 1986). Although information regarding from where these ligatures had been obtained was rarely noted in coroners' records, they are easily available. Similarly the main ligature points used (beams, girder, roofs and trees) are commonly available. It would appear, therefore, that restricting access to ligatures outside institutional settings is not possible.

Suicides in institutions (prisons and psychiatric hospital wards) made up only a small proportion of the total hanging suicides (6.2%). Nevertheless such suicides are potentially preventable through environmental modification (Appleby et al., 2001; Burrows et al., 2003; Shaw et al., 2003). The ligature point used in one of the psychiatric ward suicides – the bed curtain rail – should no longer be available in psychiatric units as trusts were required by law to remove non-collapsible bed curtain rails by March 2002 (National Institute for Mental Health in England, 2003). Environmental audits, planned as part of the National Suicide Prevention Strategy to minimise the risk of hanging (National Institute for Mental Health in England, 2003), need to take into account, when assessing potential ligature points, the finding that nearly half of all suicides do not involve full suspension. The prison hangings identified in our study could not have been carried out in a full specification (ligature-free ‘safer cell’ (Burrows et al., 2003).

In 7 cases of hanging the person also self-poisoned. If those who hang themselves are found alive, and treatment focuses on the hanging alone without investigation of possible additional suicide methods, the episode may still be fatal.

As it may not be possible to prevent hanging suicides through the restriction of access to ligatures and ligature points outside institutional settings, the focus needs to be on understanding the reasons for the use of this method and the prevention of factors leading to suicide generally.

**ACKNOWLEDGEMENTS**

We thank the coroners who gave us access to their records and their staff who also gave us assistance.

This research was funded by the Department of Health. Views expressed in this paper are those of the authors and not necessarily those of the Department of Health.

**REFERENCES**


Bristol: Department of Health.


Correspondence
EDITED BY KHALIDA ISMAIL

Contents  ■  Cognitive processing in schizophrenia  ■  Early interventions for psychosis  ■  Integration of psychiatric and physical health  ■  Quality of life and ECT

Cognitive processing in schizophrenia

I read the short report by Hall et al (2004) with interest. The authors reported a marked impairment in the ability of people with schizophrenia to make social judgements from facial expressions. Their findings complement and extend earlier studies by us and others (Hellewell et al, 1994; Edelstyn et al, 1996, 2003) that have reported the presence of impairments in facial recognition memory. However, these abnormalities in facial and emotion recognition do not appear to lead to obvious difficulties in day-to-day life; for example, individuals do not appear to exhibit problems with the recognition of familiar people. This apparent inconsistency between experimental findings and real-life situations raises issues about the role played by these cognitive abnormalities in schizophrenia. It is likely that these impairments are stable abnormalities rather than being transient indicators of dysfunction. This would be consistent with structural or functional abnormalities in schizophrenia, which only become evident when the processing systems are placed under high levels of stress, for example, during the prodromal or psychotic phases of a functional illness. This line of reasoning is supported by Hall et al’s finding that individuals with positive symptoms are unable to identify even basic facial emotions. These inherent weaknesses within the processing system may remain hidden during quiescent periods, but may be artificially exposed in the laboratory by challenging the processing system with particularly difficult tasks. Such deficits in visual processing, when combined with other factors such as changes in mental state and impaired cognitive reasoning, operate in a complex interaction to produce psychotic episodes.

In an attempt to understand the basis of their findings, Hall et al draw attention to the roles of the frontal and temporal cortices as well as the amygdala. In addition to these, we believe that abnormalities in the non-intentional, automatic acquisition of knowledge about the structural relations between objects or events may contribute to impairments in social cognition. Lewicki (1988) and others have suggested that intuitive knowledge can influence how people form impressions, draw inferences and react to situations and people. Interestingly, a number of recent studies have reported the presence of implicit learning abnormalities in people with schizophrenia (e.g. procedural learning, word-stem completion, lexical and semantic priming) (Schwartz et al, 2003). Future research might examine how those with schizophrenia acquire implicit knowledge of regularities in social contexts and how this knowledge relates to adaptive functioning in schizophrenia.


F. Oyebode  Department of Psychiatry, University of Birmingham, Queen Elizabeth Psychiatric Hospital, Edgbaston, Birmingham B15 2QZ, UK.  E-mail: Feri.Oyebode@bsmht.nhs.uk

Authors’ reply: Professor Oyebode draws attention to a number of interesting issues in response to our study of social cognition and face processing in schizophrenia. A key question raised by our study is why deficits in emotion recognition were state-dependent, being limited to individuals experiencing positive symptoms, while impairments in social cognition were stable. One possibility, as discussed by Professor Oyebode, is that those who are free of positive symptoms are able to use alternative cognitive strategies to identify basic facial emotions. This view is supported by a functional magnetic resonance imaging study in which individuals with schizophrenia, none of whom was experiencing positive symptoms, were able to identify facial emotions correctly but nevertheless showed deficits in amygdala activation when processing facial affect (Gur et al, 2002). These findings suggest that other brain regions compensate for the normal functions of the amygdala in facial affect processing when individuals with schizophrenia are free of positive symptoms. More difficult tests, such as our social cognition task, may prevent such compensation and thus reveal an underlying stable deficit.

Professor Oyebode also points out the apparent discrepancy between the finding that people with schizophrenia have impairments in facial recognition memory on formal testing, but are able to recognise familiar people in day-to-day life. In our study we found no deficit in the ability of those with schizophrenia to recognise the identity of novel faces presented concurrently, suggesting that the deficits seen in previous studies resulted from the mnemonic and attentional demands of the tasks used, which may be lower for familiar people.


J. Hall, S. M. Lawrie  Division of Psychiatry, School of Molecular and Clinical Medicine, University of Edinburgh, Kennedy Tower, Royal Edinburgh Hospital, Morningside Park, Edinburgh EH10 5HF, UK.  E-mail: jhall@doctors.org.uk

Early interventions for psychosis

The last Cochrane systematic review of early intervention for those with psychosis included cognitive–behavioural therapy (CBT), family therapy and medication, and reported no significant decrease in the development of psychosis at 12-month follow-up (Marshall & Lockwood, 2004).
The implications of the recent study of CBT for the prevention of psychosis (Morrison et al., 2004) need to be realistically interpreted with this background.

First, two people were excluded from the cognitive therapy arm after the trial had begun, which would have led to a non-significant result. This should have been acknowledged in the abstract, as an abstract has the most impact with service planners.

Second, after 6 months of cognitive therapy, there was a decrease in the development of psychosis compared with the control arm; however, there was similar distress for both groups. Cognitive therapy for psychosis has an aim of decreasing the distress of psychosis as well as the formulation of an explanatory model for that psychosis. It may be that a reframed and normalised explanatory language was taught to the individuals at high risk, and this led to the decreased identification of symptoms at 12 months and the masking of a psychotic episode. This would not ultimately lead to a decrease in distressing psychosis, but to a later identification of psychosis and a possible delay in pharmacological treatment.

The possible risk of harm or hazard was ignored, with a clear bias against the use of medication expressed by the authors in the discussion. Furthermore, the editorial comment alluded to the possibility of premature publication (Tyrer, 2004), but it is the implication of harm which needs to be explicitly stated.


K. H. Marlowe Early Psychosis Intervention Team (EPIT), Counties Manukau DHB, Hartford House, Papatoetoe, Auckland, New Zealand. E-mail: KarlMarlowe@middlemore.co.nz

Author’s reply: We welcome Dr Marlowe’s comments on our paper and would like to respond to the issues that he identified. The Cochrane review to which he refers examined more traditional approaches to early intervention (i.e. from first episode onwards) rather than a preventive approach in people at high risk, so we are unsure of the relevance of this. Within the manuscript we clearly acknowledge that there were several methodological limitations, including the exclusion of two participants, but we were unable to incorporate these in the abstract as he suggests because of limitations of abstract length imposed by the Journal (indeed, we were asked to further reduce the abstract at proof stage).

We agree that cognitive therapy for psychosis (and the prevention of psychosis) has an aim of decreasing the distress of psychotic experiences as well as the formulation of an explanatory model for a person’s difficulties. We also agree that a reframed and normalised explanatory language may be developed by the service users; however, it is unlikely that this would lead to a masking of a psychotic episode. Rather, it is intended to reduce the potential for catastrophic appraisals of psychotic experiences, which are very clearly implicated in the experience of distress (Chadwick & Birchwood, 1994), and the development of normalising appraisals is at the heart of cognitive therapy for established psychosis (Morrison et al, 2003) and the prevention of psychosis alike (French & Morrison, 2004). Even if such a masking were to occur, the assumption that this could cause harm clearly demonstrates a bias against the use of psychosocial interventions, as it suggests that only pharmacological treatments can reduce the potential harm that may result from an untreated psychotic episode, when there is evidence that psychological treatment is also important in this respect (de Haan et al, 2003).

We are accused of being biased against using antipsychotic medication; we certainly are against medication in a population who are yet to develop a psychotic disorder, for the ethical reasons outlined within our paper and elsewhere (Bentall & Morrison, 2002). Finally, it is suggested that we avoid explicitly stating the possibility of harm arising from such an intervention; however, we clearly highlight the possibility of harm resulting from stigmatisation.


A. P. Morrison Department of Psychology, University of Manchester, Copley Street, Manchester M13 9PL, UK. E-mail: tony.morrison@psy.man.ac.uk

Integration of psychiatric and physical health

In The Netherlands the British Journal of Psychiatry is distributed among Dutch psychiatrists by courtesy of the pharmaceutical industry. For the October issue of the Dutch edition I was asked to write the editorial comment, to be circulated with the Journal as an accompanying letter. My focus is integrated psychiatry in medicine.

Reading the October issue I was struck by the lack of an integrated perspective. Current epidemiological findings underscore how the organisation of our healthcare system is epidemiologically unfair and does not take into account the frequent co-occurrence of psychiatric disturbances and physical illness (Kendell, 2001; Royal College of Physicians & Royal College of Psychiatrists, 2003). The fragmentation of care is seen as one of the major problems of current healthcare (Institute of Medicine, 2001); this applies with regard to treatment of physical disorders in mental healthcare and vice versa.

The editorial by Kingdon et al (2004) on the recommendations of the Council of Europe lacks such an integrated perspective. Among the recommendations the quality of physical care is not mentioned by the Council other than in relation to restraint, and this omission is not mentioned by Kingdon et al.

Similarly, the review by Thornicroft & Tansella (2004) opens with the fact that depression leads to more disability-adjusted life-years than cardiovascular disease and cancer, but it does not report their meaningful interrelation, for instance through compliance (DiMatteo et al, 2000). In the section ‘Acute in-patient care’ it is mentioned that patients with physical comorbidity should preferentially be seen in such facilities and not in community...
correspondence

Declaration of interest
F.J.H. has received a fee for writing the editorial comment circulated with the Dutch edition of the British Journal of Psychiatry.


F. J. Huysen Free University Medical Centre, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands. E-mail: fj.huysen@vumc.nl

Quality of life and ECT

The first author of this study (McCall et al., 2004) has an apparent career, if not financial, conflict of interest in the treatment being reviewed. He is the President of the Association for Convulsive Therapy, the industry trade organisation, as well as the editor of its journal which he calls ‘the voice of ECT’ (McCall, 2004). This ought to have been revealed to readers directly; as it is, it reveals itself in the many flaws of research design which bias the study towards minimising the risks of electroconvulsive therapy (ECT).

The study included those who had had ECT as recently as 4 months previously, thus building into the research design the assumption that the adverse effects of ECT resolve within that time period; but there is evidence that this is not so. If it is not, then the study is simply comparing those who are still suffering the after-effects of ECT with those suffering more severe after-effects, a comparison which tells us nothing about the effects of ECT per se. The fact that those at baseline averaged a score of only 18 on the Mini-Mental State Examination suggests some type of cognitive dysfunction, perhaps due to ECT, even at that point.

The measures chosen by McCall et al. in all areas – cognition, amnesia and, most importantly, what he calls quality of life and functioning – are the grossest possible, and cannot register the deficits known to be associated with ECT because they are simply not designed to do so. The authors must be aware of the work of the Service User Research Enterprise (SURE) group (Rose et al., 2003) in which patients describe a highly specific pattern of permanent memory and cognitive deficits post ECT. This was a rigorous systematic review of the literature on ECT’s effects, and encompasses what most people would call quality of life and functioning. It revealed that for at least one-third of individuals ECT had deleterious, often devastating, effects on these areas which lasted more than 6 months and appeared to be permanent.

Individuals lost the ability to perform their jobs. They lost memory of up to 20 years of their lives. They were unable to handle schoolwork because of impaired memory function and concentration. They did not recognise persons previously well known to them. They waited anxiously for the promised ‘return of memory’ which never came. None of this is consistent with improvement in quality of life.

Why then are McCall et al.’s results so seemingly contradictory? Because he did not ask about these things. Instead, participants were asked, quite literally, whether they could wipe their own backsides. If they were simply able to get out of bed, feed and dress themselves, and use a bus or a telephone they were graded as functioning at the highest possible level. No one has ever reported that ECT affected their ability to use a toilet.

Finally, 4 weeks after ECT is too soon for individuals, who are unlikely to have tried to go back to work or school yet, to be able reliably to assess their altered memories and abilities.


L. Andre PO Box 1214, New York, NY 10003, USA. E-mail: ctip@erols.com

Author’s reply: We are grateful for Ms Andre’s interest in our paper. She is the director of the Committee for Truth in Psychiatry (CTIP), which is a vocal anti-ECT group in the USA (see http://www.harborside.com/~equinox/ect.htm). As such, we feel that our work must be on target and of some importance to attract their criticism. Ms Andre has some specific complaints with our work, which we address as follows.

First, Ms Andre suggests that I have an apparent ‘career, if not financial, conflict of interest’ that invalidates the paper, especially as pertains to my role as President of the Association of Convulsive Therapy (ACT). I receive no financial or material support for serving as president of ACT; ACT is self-supporting through the dues of its members. The idea of a ‘career conflict of interest’ is not a concept endorsed by the American Medical Association Code of Ethics, per section 8.031 (Council on Ethical and Judicial Affairs, 1997). It is just as likely that she has a conflict of interest as director of CTIP in writing her letter — any information that supports the use of ECT threatens the position of CTIP. We would
welcome Ms Andre’s full disclosure of her financial support from CTIP, and disclosure of the source of funding for CTIP since its website states that dues are not a requirement for membership.

Second, she claims that those in our study had an average Mini-Mental State Examination (MMSE) score of 18 at baseline. In fact, the mean baseline MMSE score was 27.4, as shown in Table 2 (McCall et al., 2004: p. 407). The minimum MMSE score for inclusion was 18.

Third, Ms Andre takes us to task for not citing Rose et al. (2003). The Rose et al. paper has merit, but has no direct bearing on our work. Those authors ‘aimed to . . . assess the debated distinction between efficacy, effectiveness, and satisfaction;’ the focus of our paper is quality of life (QOL) and function, not ‘satisfaction’. As reviewed by Asadi-Lari et al. (2004) satisfaction and QOL are discrete, non-overlapping ideas.

Fourth, Ms Andre asserts that memory effects of ECT must necessarily affect QOL. Ms Andre is changing the definition of terms to suit her purposes, or else remains unfamiliar with the field. QOL research is ‘. . . widely regarded as a robust measure of outcome assessment. . . ’ and is defined as ‘. . . the patient’s perspective of their own health status’ (Asadi-Lari et al., 2004). It is a violation of the concept for anyone, including Ms Andre, to define a patient’s QOL for them.

Fifth, Ms Andre belittles our work for showing that ECT is associated with significant improvement in activities of daily living and instrumental activities of daily living. She does not recognise that impairment of instrumental activity of daily living may be the deciding factor in referring patients for ECT (McCall et al., 1999) and that ECT is superior to medication in improving instrumental activities of daily living over 1 year of follow-up (McCall et al., 2001).

We do share one goal with Ms Andre – a desire for truth in psychiatry. We choose to reveal truth through the scientific method as opposed to rhetoric.


W. V. McCall Department of Psychiatry and Behavioral Medicine, Wake Forest University Health Sciences, Medical Center Boulevard, Winston-Salem, NC 27157-1071, USA. E-mail: vmccall@wfu.bmc.edu

Author’s reply: Dr McCall responds to my letter but does not answer it. I get very tired of explaining to ECT proponents that the Committee for Truth in Psychiatry is not an ‘anti-ECT’ group, but no matter how many times and in how many contexts I do so, that false statement continues to be made. More about CTIP later, since I cannot leave Dr McCall’s claims unrefuted. But much more important are the still unaddressed concerns about the methodology and validity of the McCall et al. study.

My point about building assumptions about the longevity of ECT’s adverse effects into the research design by including persons who had recently had ECT was not addressed.

Nor was any evidence presented to show that the rating scales chosen by McCall et al are relevant to the types of deficits reported by former ECT patients and illustrated so well in the SURE report. (Nor has there been evidence, which I requested privately from the author, to show that the study participants, who for some reason scored so poorly on both the MMSE and the IADL prior to this course of treatment, are representative of ECT patients as a whole.)

McCall’s point that ex-patients and only ex-patients define what quality of life is and by what standard it should be measured is exactly my own: no ECT survivor or ex-patient ever has or ever would define ‘quality of life’ or ‘functioning’ in the terms Dr McCall uses. He says, ‘It is a violation of the concept for anyone to define a patient’s QOL for them’, yet that’s exactly what he has done. Had he asked patients themselves, an approach taken by the Rose et al group, he would have set off in a productive direction instead of down a blind alley.

His attempt to selectively redefine the work of Rose et al as research on ‘satisfaction’, not relevant to work on quality of life, is without foundation, as a reading of the actual study will show. It was he who brought up the work ongoing in Britain as relevant, by his reference in his first sentence to the National Institute for Clinical Excellence guidelines which came out concurrently with, and used some of the same evidence base as, the report of the Rose group at the SURE.

There is a wide literature on non-financial conflicts of interest, best described as ‘an individual occupying dual roles which should not be performed simultaneously’ (Fava, 2001). Those include treatment researcher and editor of a journal promoting the treatment under study.

If you yourself read what CTIP says, and not what others say about us, you will begin to wonder where the ‘anti-ECT’ claim comes from. We are an international organisation made up entirely of persons who have received ECT. We represent the spectrum of outcomes, from persons who feel ECT is beneficial and have had it more than once, to persons whose lives were ruined by it. None of us was truthfully informed of the risks of ECT before consenting to it, and no one liked being lied to. Our organisation exists for one purpose only: to advocate truthful informed consent for prospective ECT patients. Thus, it makes no sense to say that ‘any information that supports the use of ECT threatens the position of CTIP’.

Whether you are of the opinion that being in favour of truthful and informed consent somehow makes you anti-ECT depends on whether you believe that patients have the right to full disclosure of ECT’s risks – and the right to make a decision for themselves based on that information – or whether you believe that ECT’s risks are such that full disclosure would result in patients in bloc deciding to forego the treatment. That Dr McCall and colleagues are in the latter camp speaks much more eloquently than their article as to what they really believe about ECT’s effects on quality of life.

CTIP, founded in 1984, has never received funding of any kind.


L. Andre PO Box 1214, New York, NY 10003, USA. E-mail: ctip@erols.com

Editor’s note: This correspondence is now closed.
One hundred years ago

Insanity and murder

SIR,—So far from agreeing with Dr. Mercier’s letter which appeared in the British Medical Journal, January 28th, may I be allowed to “demur, contest, renounce, repudiate and deny” his astonisihing doctrine that a person may be at once insane and responsible.

The proposition that insanity and irresponsibility are not convertible terms may commend itself to Dr. Mercier, but I shall be surprised if the majority of his brother alienists do not emphatically repudiate it, while it is one which is hardly likely to commend itself to the humanitarianism of the present day. As is well known, some of the most enlightened jurists have propounded the precisely opposite doctrine, namely, “that no man is responsible for an act which is the product of mental disease”, and it is hardly too much to say that this is in fact the principle adopted (in capital cases) by our own Home Office authorities, though not, unfortunately, by our Courts of Law. The fact that this divergence exists is, I venture to think, abundant justification for your recent article.

That it should be possible to deal with a person purely as a criminal, when competent medical evidence has been adduced at the trial to the effect that such person was insane at the time of committing the act and that the verdict and its consequences has subsequently to be disregarded by the Home Office in order that substantial justice should be done, is a proceeding that indicates that our present system is far from satisfactory.

A case occurred during the past year at the Central Criminal Court, in which the prisoner, though declared insane by competent medical evidence, was sentenced to death. Dr. Mercier may regard this as perfectly satisfactory (especially as the Home Office fortunately interfered and sent the case to Broadmoor); I do not.

It is clearly of the utmost importance that criminal proceedings should coincide with that sense of justice which is innate in mankind. If an insane person is dealt with by the Court as a criminal, that sense of justice is shocked, the law is not respected, and the main object of punishment in securing obedience to that law is not attained.

My own view remains now what it has been for some time past, that the question of sanity or insanity is purely a question for competent medical evidence, and that the question of the allocation of responsibility is one for the Court unhampered by the dicta of 1843; but whatever may be thought of this proposition, few, I venture to think, will agree with Dr. Mercier that those who are afflicted with the saddest of all afflictions that can befall a suffering humanity should be treated as felons, still less that they should be consigned to a shameful death. — I am, etc.,

A. DOUGLAS COWBURN.

Of the Middle Temple and Central Criminal Court, London, S.W., Feb. 6th.

Barrister-at-Law

REFERENCE


Researched by Henry Rollin, Emeritus Consultant Psychiatrist, Horton Hospital, Epsom, Surrey

---

Corrigendum

Needs for mental health treatment among general practice attenders. BJJ, 185, 318–327. Figure 1 (p. 320) was printed incorrectly. The correct figure appears right.

---

Fig. 1 Study design, sampling fractions and attrition rates.
**Book reviews**

EDITED BY SIDNEY CROWN, FEMI OYEBODE and ROSALIND RAMSAY

**Autism and Creativity: Is There a Link between Autism in Men and Exceptional Ability?**


In this book pathologising creativity and genius, tidbits from the lives of Ludwig Wittgenstein, William Butler Yeats, Lewis Carroll and others are proffered as ‘proof’ of Fitzgerald’s conclusion that high-functioning autism and Asperger’s syndrome are more common than we think and critical to genius and creativity. The problem with this is that Fitzgerald writes as if he were their psychiatrist. He isn’t. Nor is he really their biographer. The ‘facts’ he uses to support his case have been cobbled together from secondary sources, by his own admission, ‘using biographies that have received favourable reviews in professional journals and recognised publications such as the *Times Literary Supplement*’.

Fitzgerald finds what he’s looking for, trawling life stories for nuggets to fit his theory: Hitler’s autistic psychopathy, Wittgenstein’s autistic superego, Yeats’s classic ‘Asperger pose’ and ‘autistic aggression’. One might be forgiven for thinking that this sort of fudged pseudoscience comes with the genre. But retrospective psychobiography can be done without succumbing to this book’s shortcomings. Kay Redfield Jamison’s brilliant and captivating book *Touched with Fire* examines the relationship between bipolar disorder and creativity by presenting extracts of psychohistory as recorded by writers and artists themselves and consulting widely with colleagues working in the humanities (Jamison, 1991).

Fitzgerald’s conclusion is touted on the back cover as ‘spirited and controversial’. I think it’s shaky. Statements that he makes, such as ‘another important point emerging from this book is that the autistic spectrum is very wide and this book widens it still further’, seem as absurd as arbitrarily altering the definition of fever to fit a hypothesis that there is a link between pyrexia and genius.


**Psychotherapy for Borderline Personality Disorder, Mentalization-based Treatment**


While reading this book I also read reviews of Giełgud’s *Letters* (Mangan, 2004) and a biography of Michael Redgrave (Strachan, 2004). These gave accounts of the lives of these two actors that left little out from a diagnosis of borderline personality disorder. Narcissism, impulsivity, self-destructive behaviour, identity defusion; if they had turned up for an out-patient assessment, there would have been little difficulty in ascribing an Axis II diagnosis. Most psychiatrists are left uneasy about pathologising personality, as it would appear that the only difference between ‘them’ and ‘us’ is chance, circumstance or maybe the talent to get away with it.

On the other hand, psychiatry without a theory of personality development and how it can go wrong not only becomes an arid and dehumanising symptom checklist, but also has little to offer the many individuals who seek help for their chaotic lives and unhappy relationships with others.

Bateman and Fonagy have previously published the outcome of a randomised control trial of a day-hospital treatment for borderline personality disorder. This psychodynamically based treatment was shown to be highly effective on a number of measures to reduce morbidity. In the past psychodynamic treatment has lacked evidence of efficacy and its theory has not generally been backed by developmental studies. In addition the practice of psychodynamic therapy has been so unsystematic as to leave the majority of mental health workers totally bewildered about how therapy is conducted.

This current book addresses both these issues. The theoretical first half considers the evidence around the authors’ central
Early Intervention for Trauma and Traumatic Loss

It used to be said that post-traumatic stress disorder (PTSD) was a normal reaction to an abnormally traumatic situation, and that anyone was susceptible to its development. In fact, the vast majority of people exposed to natural or man-made trauma do not develop PTSD, and factors such as individual vulnerability, peri-traumatic influences and social support play a critical role in defining risk. However, for those who develop significant psychological distress, when should appropriate intervention take place? In the aftermath of the earthquake and tsunami in South-East Asia, the role of early psychological intervention for survivors has again come into sharp focus. News reports reassure viewers and listeners that counsellors are on hand to help relieve stress and prevent the development of more chronic psychological syndromes. These interventions continue unabated despite research showing that techniques such as psychological debriefing have limited, if any, benefit and may actually increase the risk of developing stress reactions in the future.

So, is there any role for early intervention in relieving distress and preventing future morbidity? Hot on the heels of a similar book with a more European perspective (Omer & Schnyder, 2003), this text, covering early-intervention initiatives across the lifespan and in various clinical settings, provides an authoritative summary of existing knowledge, as well as offering evidence-based recommendations and directions for future research.

It is divided into three sections. Part I focuses on the psychological impact of trauma and traumatic loss, centring on acute stress disorder, PTSD and traumatic bereavement, emphasising the importance of risk and resilience in promoting recovery in a selected at-risk population. Part II, the core of the book, highlights the importance of secondary prevention using evidence-based practice that is tailored to the population and context. Although ‘psychological first aid’ is appropriate following trauma, it should not be seen as a therapeutic or preventative intervention. The criticisms of secondary prevention strategies such as critical incident stress debriefing are clearly elucidated in the book and balanced with a summary of cognitive-based models of proven efficacy in early intervention. This is the section of the book that I enjoyed the most and is likely to be the most useful to clinicians. Part III covers experience gained in special populations, including the response to the terrorist attacks on New York and Washington in September 2001 and the challenge of developing effective early-intervention strategies in the US military, where the use of psychological debriefing is deeply ingrained.

Despite some repetition between chapters, this is a useful book, which goes some way towards rehabilitating the whole concept of early intervention and secondary prevention in trauma work. There is increasing evidence that such strategies work, but they must be selective and focused, appropriately timed and promote natural resources and personal resilience.


Duncan McLean Consultant Psychiatrist in Psychotherapy, Maudsley Hospital, Denmark Hill, London SE5 8AZ, UK

Ronan J. McIvor Consultant Psychiatrist, Department of Psychological Medicine, Maudsley Hospital, 103 Denmark Hill, London SE5 8AZ, UK. E-mail: r.mcivor@op.kcl.ac.uk

Practical Psychiatric Epidemiology

Enter the world of Practical Psychiatric Epidemiology. The book opens with a quotation, probably seldom surpassed in a discipline rarely afforded great literary expression, that epidemiology ‘... is like a cocktail blended from equal parts of science, art, and craft, laced with liberal applications of intellectual rigour and scepticism’. From this opening, one is confident of an enjoyable, critical introduction to the ‘young’ discipline of psychiatric epidemiology.

This text is primarily intended for students of epidemiology with an interest in psychiatry. Epidemiological concepts and techniques are introduced and discussed with exceptional clarity, and there is good application to psychiatric research throughout. Each well-referenced chapter...
Practical Psychiatric Epidemiology closes with a series of practical exercises that serve to reinforce the concepts previously introduced. For these reasons the book will appeal to anyone involved in the lecturing and teaching of (psychiatric) epidemiology. It should also be considered essential for more experienced clinicians in psychiatry or psychology who require a grounding in epidemiology.

The book consists of 21 chapters from leading contributors in psychiatric epidemiology, arranged into four distinct sections: basic principles, study design, interpretation, and special issues. Each section, indeed each chapter, provides an excellent stand-alone introduction to an important epidemiological theme, be it case-control studies, multi-level modelling or the emerging role of genetic epidemiology within psychiatry. What really makes this text stand out, however, is the skill of the editors in integrating a diverse range of themes, both succinctly and comprehensively.

The first section, ‘Basic principles’, introduces the historical development of psychiatric epidemiology, which provides a useful contextual perspective and relates older issues, such as Durkheim’s (1951) ideas on suicide with more recent concepts, such as social capital (Putnam, 1996). Good introductions to measurement and ethics in psychiatry are included as well as an informative chapter giving a balanced critique of the applicability of psychiatric paradigms across different cultures.

The heart of this book lies in the two sections introducing epidemiological methodology and analysis, ‘Study design’ and ‘Interpretation’. Each major epidemiological study design is introduced, and balanced arguments on the advantages and disadvantages are outlined. Informative examples from psychiatric epidemiology are included. The section on interpretation discusses the epidemiological mantra – chance, bias, confounding – in detail, but it is the inclusion of chapters on causality and further statistical methods that will appeal most to people with a prior background in epidemiology.

The book closes by addressing emerging issues in psychiatric epidemiology, including the role of genetic studies, qualitative data and health economics, finishing with the editors’ excellent appraisal of future challenges faced by the discipline.

For those who already have a qualification in epidemiology much of the content will be familiar, although as a stand-alone quick reference it remains useful. For those, however, with an interest in psychiatric research – from the new epidemiological researcher to the experienced clinical psychiatrist – this publication will be an invaluable companion for undertaking research in psychiatric epidemiology.


James Kirkbride  Research Student, Department of Psychiatry, University of Cambridge, Box 189, Level E4, Addenbrooke’s Hospital, Hills Road, Cambridge CB2 2QQ, UK
From the Editor’s desk

PETER TYRER

THIS MONTH’S ISSUE: TO GENERALISE IS TO BE AN IDIOT

No, these are not the words of a social scientist, but of William Blake, an individual thinker if ever there was one. The trouble is that science has to have a universal message and the best science of all, like Einstein’s theory of relativity, is the most universal. So when we fail to generalise we drown our message but we also drown it if we generalise into banality or empty words. This issue illustrates all parts of the generalisation problem. The NICE guidelines reviewed by Whitty & Gilbody (pp. 177–178) are like guidance on healthy eating; they are right, sensible and worthy, but publicity and audit is not going to be enough to get the message over to the patients who need it unless we improve the rapidly fraying link between primary and secondary care. In a similar vein, the relevance of gene–environment interaction to the manifestation of affective disorders is now universally appreciated (Farmer et al, pp. 179–181) but the key need now is to identify the ‘environomes’ to go with the ‘genomes’. The failure of our nicely defined diagnostic descriptions of common mental disorders to generalise to their identification and use in practice is illustrated by Maj (pp. 182–184) in his critical review of comorbidity, and by Khan et al (pp. 190–196), who show that the personality dimension of neuroticism, despite it being banned from polite nosological circles, is an aggressive gatecrasher and will not be kept away.

The failure to show at least some common interventions that are generalisable is illustrated by Cure and her colleagues (pp. 185–189), who deplore the diversity of treatments for aggressive behaviour as so ‘few studies focus on similar interventions for similar participants’. It might help if there were similar outcome measures too and Barkham et al (pp. 239–246) make a bold case for CORE measures being used in psychotherapy studies, but they will have to fight here in a very crowded marketplace. With all this uncertainty it is nice to stick with one generalisable message, ‘the past tends to predict the future’, and this is shown by both Carter et al (pp. 253–257) and Robertson et al (pp. 258–259). And to come back to Blake, what phrase could show greater generalisation than ‘to see a world in a grain of sand’? ‘Ah,’ he might have replied, ‘I said “a world”, not “the world”’.

THE JOURNAL IN 2003 AND 2004 – SOME BARE STATISTICS

In 2004 the Journal received 703 articles for intended publication, and in 2004, 798, an increase of 13.5%. This unfortunately is not encouraging news for authors, as we are not able to increase the size of the journal commensurately and the introduction of short reports only has a small effect on the total number of articles published. We currently have to reject approximately four out of every five articles submitted that are not commissioned (and the latter applies to very few). I hope authors appreciate the difficulties we have in turning down so many competent papers that in the past would have had no difficulty in finding space in the journal. Long gone are the days when the Journal could publish articles exceeding 100 pages (e.g. Aubrey Lewis’s classic 1934 paper on melancholia (Journal of Mental Science, 80, 277–378) (I have known some who reference this as 277–278 as they perceive it as a misprint). Nevertheless, we are trying to be flexible and recognise that sometimes a longer article is justified, as for example with Harris & Barraclough (1998) on the excess mortality of mental disorder (British Journal of Psychiatry, 173, 11–53), which had such a major impact on the thinking of a fellow editor, Povl Munk-Jørgensen (Psychiatric Bulletin, 28, 472). So if you feel a long article is absolutely necessary for the subject, please send it in – it will be looked at very closely.

The most cited paper in 2004 was that by Louise Arseneault and her colleagues from the Institute of Psychiatry (British Journal of Psychiatry, 184, 110–117), with 13 citations, and that in 2003 was that by Lakshmi Yatham and colleagues on the benefits of risperidone as an adjuvant treatment in mania (British Journal of Psychiatry, 182, 141–147), with 29 citations. This is more evidence of Canadians having the yen to stimulate rapid response (referred to in this column in the February issue, p. 175), as Professor Yatham is from Vancouver and Louise has an attachment to the Institute from the Canadian Institute of Health Research. How do these Canadians do it? I blame those long winters.