Psychiatry in pictures:

Psychiatry in pictures

Highlights of this issue:

Highlights of this issue
Sukhwinder S. Shergill

EDITORIALS:

What is the heartland of psychiatry?
Guy M. Goodwin and John R. Geddes

Affective neuroscience and psychiatry
Neil A. Harrison and Hugo D. Critchley

REAPPRAISAL:

Management of insomnia: treatments and mechanisms
Sue Wilson and David Nutt

REVIEW ARTICLES:

Antipsychotic treatment of primary delusional parasitosis: Systematic review
Peter Lepping, Ian Russell, and Roland W. Freudenmann

Guided self-help in primary care mental health: Meta-synthesis of qualitative studies of patient experience
Nagina Khan, Peter Bower, and Anne Rogers
PAPERS:

White matter changes and late-life depressive symptoms: Longitudinal study

Genetic risk of depression and stress-induced negative affect in daily life
Marieke Wichers, Inez Myin-Germeys, Nele Jacobs, Frenk Peeters, Gunter Kenis, Catherine Derom, Robert Vlietinck, Philippe Delespaul, and Jim Van Os

Women with autistic-spectrum disorder: magnetic resonance imaging study of brain anatomy
Michael C. Craig, Shahid H. Zaman, Eileen M. Daly, William J. Cutter, Dene M. W. Robertson, Brian Hallahan, Fiona Toal, Suzie Reed, Anita Ambikapathy, Mick Brammer, Clodagh M. Murphy, and Deean G. M. Murphy

Sulcal thickness as a vulnerability indicator for schizophrenia
Vina M. Goghari, Kelly Rehm, Cameron S. Carter, and Angus W. Macdonald

Insight in psychosis: influence of cognitive ability and self-esteem
Michael A. Cooke, Emmanuelle R. Peters, Kathryn E. Greenwood, Peter L. Fisher, Veena Kumari, and Elizabeth Kuipers

Evaluating the cost-effectiveness of reduced tardive dyskinesia with second-generation antipsychotics
R. A. Rosenheck

Treatment of social phobia through pure self-help and therapist-augmented self-help
Ronald M. Rapee, Maree J. Abbott, Andrew J. Baillie, and Jonathan E. Gaston

Firearm legislation reform in the European Union: impact on firearm availability, firearm suicide and homicide rates in Austria
Nestor D. Kapusta, Elmar Etzersdorfer, Christoph Krall, and Gernot Sonneck

SHORT REPORTS:

Caudate volume in offspring of patients with schizophrenia
Rajaprabhakaran Rajarethnam, Ameet Upadhyaya, Pon Tsou, Margie Upadhyaya, and Matcheri S. Keshavan
Familial liability to schizophrenia and premorbid adjustment
Muriel Walshe, Mark Taylor, Katja Schulze, Elvira Bramon, Sophia Frangou, Daniel Stahl, Eugenia Kravariti, Eileen Daly, Paul Fearon, Robin M. Murray, and Colm McDonald

Insight, psychopathology and global functioning in schizophrenia in urban Malawi
Niall Crumlish, Prince Samalani, Anthony Sefasi, Anthony Kinsella, Eadbhard O'Callaghan, and Harris Chilale

Correspondence:

Omega-3 fatty acid for recurrent self-harm: unanswered questions
D. Basu and P. K. Barnwal

Authors' reply
M. Garland, B. Hallahan, J. R. Hibbeln, and J. M. Davis

Callous–unemotional traits and autistic psychopathy
M. F. Fitzgerald

Authors' reply
E. M. Viding

Diagnostic stability: clinical v. research
P. P. Das and S. Grover

Authors' reply
E. Baca-Garcia, M. M. Perez-Rodriguez, and M. A. Oquendo

Limitations of cognitive–behavioural therapy for sleep disorders in older adults
O. Prakash

Authors' reply
B. Siervset and I. H. Nordbus

One hundred years ago:

Asylum reports – Royal Edinburgh Asylum
Henry Rollin
Ten books:

Chosen by Michael King
Michael King

Book reviews:

The Metaphor of Mental Illness
Derek Bolton

Mood and Anxiety Disorders in Women
Jessica Heron

Understanding Nicotine and Tobacco Addiction
Paul Aveyard

Mental Health Issues in the Media: An Introduction for Health Professionals
Peter Byrne

Psychiatry in the Scientific Image
Derek Bolton

Theory of Addiction
Amanda Redvers

Aviation Mental Health: Psychological Implications for Air Transportation
Roger Bloor

Eating Disorders in Children and Adolescents
Bryan Lask

From the Editor's desk:

From the Editor's desk
Peter Tyrer
Highlights of this issue

BY SUKHWINDE R SHERGILL

HEART- OR HINTER LAND IN PSYCHIATRY

A fundamental question about the development of psychiatry in the UK is raised in an editorial in this issue of the Journal. Goodwin & Geddes (pp. 189–191) argue that the use of schizophrenia as the defining disorder of psychiatry has had a detrimental effect on psychiatry as a medical specialty. They illustrate the ways in which the focus on schizophrenia has distorted both the provision of care and the research agenda in psychiatry, and served to enhance negative views of psychiatry; they contend that other types of illness, such as bipolar disorder, may usefully replace schizophrenia as the model of a prototypical psychiatric disorder.

ANTIPSYCHOTICS: TARDIVE DYSKINESIA AND DELUSIONAL PARASITOSIS

The cost-effectiveness of second-generation antipsychotic medication has been called into question by more recent treatment studies suggesting that there are no substantial differences in health outcomes or side-effects compared with some first-generation drugs. Rosenheck (pp. 238–245) examines the cost-effectiveness of using second-generation antipsychotic medication in reducing the risk of tardive dyskinesia, which may be left as the main perceived advantage of this class of drug. The reduction in tardive dyskinesia with the use of second-generation antipsychotic treatment is shown not to be cost-effective, unless the costs of the first- and second-generation drugs are made comparable. A review by Lepping et al (pp. 198–205) of the utility of antipsychotic treatment in primary delusional parasitosis, a relatively rare but extremely challenging disorder, found limited evidence that antipsychotic treatment was effective in this condition and that remission rates did not differ between first- and second-generation antipsychotics.

NEUROSCIENCE, AFFECTIVE DISORDERS AND AUTISM

There has been increasing interest in the brain mechanisms underlying emotional behaviour – both in mood disorders and its relevance to normal and dysfunctional social intercourse. In their editorial, Harrison & Critchley (pp. 192–194) provide an elegant contemporary review of this developing area of affective neuroscience, and provocatively conclude that the recognition that emotion influences a broad spectrum of human functioning could result in a rebranding of biological psychiatry as clinical affective neuroscience. Stress is a fairly endemic part of contemporary lifestyles and its relationship with mood states remains difficult to disentangle. Wichers and colleagues (pp. 218–223) used twin data and daily life measures to demonstrate that there is a significant genetic component in the manner that people evaluated stressful events during their day-to-day existence. Moderately stressful events resulted in little negative affect in healthy participants but induced significant negative affect in those with a high familial loading of depression, suggesting that this may be an endophenotype for depression. Depressive illness in older people has been linked to changes in cerebral white matter, but the temporal relationship has not been clear. Teodorczuk et al (pp. 212–217) report that white matter changes pre-date the development of depressive illness in a longitudinal pan-European study of older people. They clarify that while white matter changes were correlated with subsequent depressive symptoms at a 1-year follow-up, these were not predictive of the onset of depressive episodes; this suggests that white matter changes on magnetic resonance imaging (MRI) should be included with other independent predictors of illness, such as previous illness and quality of life, in assessing risk of depression in this group. Craig et al (pp. 224–228) report an MRI study of women with a diagnosis of autistic-spectrum disorder, demonstrating changes in their brain structure in multiple cortical regions, but with decrements in the right limbic cortex associated with dysfunctional social interaction. They suggest that this is compatible with the role of this region in affective processing, noted in the lesion literature in both monkeys and humans.

INSOMNIA, AVAILABILITY OF FIREARMS AND SOCIAL PHOBIA

Insomnia is a common clinical complaint and Wilson & Nutt (pp. 195–197) provide a concise reappraisal of the routine treatments in current use, and their mechanisms of action. They note that it may be reassuring that longer-term data on the newer hypnotic medications suggest that they maintain their efficacy over 12 months, but caution that withdrawal reactions may occur in some patients. Reducing the availability of higher-risk means of suicide is accepted as a useful strategy for reduction of suicide. The availability of firearms at home has been shown to be a risk factor for both suicide and homicide. The implementation of a new firearms law in Austria provided the opportunity to examine the effects of this change on firearm-related suicides and homicides. Kapusta et al (pp. 253–257) found that both firearm suicide and homicide rates were reduced following the implementation of a new, more stringent, firearms law. They suggest that this may be a useful strategy for other countries to follow. Social phobia is amenable to psychological treatment, but it is not clear whether the therapist is necessary for treatment. Rapee and colleagues (pp. 246–252) compared a ‘pure’ self-help treatment with self-help augmented by a therapist group session, and the more standard therapist-led group therapy. They found that the augmented option was superior to the pure self-help condition, and equivalent to the more intensive standard group therapy. This may offer an effective treatment, but requires less therapist time.
Psychiatry in pictures
EDITED BY ALLAN BEVERIDGE


*Malinconia*, by Venezuelan artist Javier Rodriguez, deals with one of the most prominent mental health issues of recent times: depression. *Malinconia* is the Italian word for melancholy. Depression and melancholia can be viewed as one and the same thing.

The piece is a collage based on the Jan van Eyck painting *Portrait of a Man with a Turban*, a famous work of the Northern Renaissance. From the mouth down the picture is untouched retaining the classic Renaissance pose and stoic expression. But from the nose up the picture is a distortion: a confused jumble of images. This image of the Renaissance and all it entails – discovery, progress and a greater understanding of the universe – is juxtaposed with images of confusion and doubt, perfectly encapsulating the milieu that so often leads to melancholy: that greater knowledge leads not to greater understanding and certainty, but instead to more questions and uncertainty.

The most powerful part of the piece is the man’s eyes. Slightly off-centre, they are where the face first becomes distorted. His eyes are the windows to his melancholia, their downward slant conveying his angst.

Peering from within the folds of the turban are four eyes, indicating a strong influence of the surrealist movement, who themselves were influenced by psychoanalysis.

Rodriguez’s works are intricate compositions made from antique books. This mixture of old materials and modern technique perfectly encapsulates his sensibility of classic ideas in a contemporary context. Please visit http://www.javierrodriguez.co.uk to see his work.
What is the heartland of psychiatry?

GUY M. GOODWIN and JOHN R. GEDDIES

Summary  Psychiatry has long identified schizophrenia as its defining disorder, its heartland as it has been called. In the past 20 years, this has had a number of negative consequences for psychiatry as a medical specialty, which result from the uncertainty of diagnosis and an increasing emphasis on demedicalising services in an attempt to provide social care outside hospital. These changes have probably increased the stigma attached to psychiatric practice and threaten to deskil doctors. They have also meant that services for other disorders do not meet the needs of patients. To continue to allow schizophrenia to be the paradigm condition is against the interests of psychiatrists and their patients.

Declaration of interest  Both authors have received financial support from pharmaceutical companies. G.M.G. is a non-executive director of the Oxfordshire and Buckinghamshire NHS Mental Health Partnership Trust.

About 20 years ago, it was a commonplace to refer to schizophrenia as the heartland of psychiatry (see Bebbington & McGuffin, 1988). The reason for this curious use of an emotive reference to territory seemed rather odd, but not terribly important, because it was in many ways an exciting time to do research in psychiatry or have a scientific interest in its progress and understanding. Unfortunately, the identification of schizophrenia in this way has had largely negative consequences for the practice of psychiatry as a distinctive medical specialty, the full effects of which are only fully being felt now.

WHY SCHIZOPHRENIA?

Why was schizophrenia so important to the generation that looked to Aubrey Lewis, John Wing, Bob Kendell and others for leadership? It had a lot to do with diagnosis, and perhaps with beds. Schizophrenia’s rich phenomenology, arcane language and foreign literature made its diagnosis seem difficult and interesting. Moreover, making the diagnosis of schizophrenia came to seem very important because psychotherapists had indulged in an almost unlimited extension of the ‘schizophrenia’ concept in the 1950s. America, to include most neurotic problems. The US–UK collaborative diagnostic project probably marked the high-point of confidence and international influence of British psychiatry (Kendell et al., 1971); its focus effectively defined what psychiatry was. It established that psychiatrists could reliably play by operational rules in making a diagnosis and led directly to the DSM and ICD classifications.

Furthermore, schizophrenia was restored as a core disorder that trumped other diagnoses in a notional hierarchy of importance. This was fundamental. The prevailing view was that schizophrenia could be reliably recognised on the basis of symptoms such as thought insertion. Thought insertion was taken to have the property that Jaspers required of a true delusion: it can be traced back to an ‘irreducible and non-understandable experience’ (Wing, 1978). It is not an extreme experience of a normal kind. When dominated by such phenomena, a mental state is thus qualitatively different from the normal. Finally, despite appearances to the contrary (comorbidity, anxiety, depression, mood elevation and cognitive impairment were commonly present), schizophrenia was held to be a unitary diagnosis. Just how unreasonable this was, and remains, seems still to be poorly appreciated.

The first and most obvious problem was that an emphasis on diagnosis delivered psychiatrists into very uncomfortable arguments about the status of schizophrenia. To suppose a qualitatively abnormal mental state is, first, foremost and inevitably, stigmatising. It echoes the sane/insane legal distinction and to this day many psychiatrists are reluctant to tell their patients that their diagnosis is schizophrenia. Second, it throws up a boundary problem: however confident one may be about its more severe forms, schizophrenia has rather friable edges, and the diagnosis at its most friable too often hinges on what people say they do and do not believe, and might or might not do on the basis of such belief. A clinical diagnosis of schizophrenia could, when based on partial delusions alone, be a social construct and the antipsychiatry argument, in all its pompous certainty, proceeds from that. Finally, patients frequently disagree with the psychiatrist’s interpretation of their mental state, and may have to be detained against their will. Historically, English psychiatrists have probably been too enmeshed in the workings of the Mental Health Act and their diagnosis has been too central in deciding detention of patients with schizophrenia. This is rarely the best basis for a positive therapeutic alliance. So, by clinging to schizophrenia as a heartland, psychiatrists have helped define and strengthen the negative view others have, both of psychiatry and of themselves.

Then there was the matter of beds. The system of large asylums may have originally been a humane innovation, but by the 1960s it had come to be an increasing cause for scandal. The institutions mirrored the chronicity of schizophrenia and seemed to amplify rather than correct the disabilities in everyday living that so many sufferers experience. Far too many patients languished in long-stay beds with minimal dignity and very little medical attention. However, the reaffirmation of the status of schizophrenia seemed to impede rather than facilitate the creation of really new services, as radicals of the time such as William Sargant had argued were needed in general hospitals to treat affective disorder (see Sargant, 1967). Possession of beds was also a perverse measure of a doctor’s individual power within the existing administrative structure, and the mentality of many psychiatrists was undeniably too hospital orientated. So, although the possession of some in-patient beds for acute treatment or respite remained and remains essential for good care, there was a failure.
MODERNITY: SCHIZOPHRENIA PLUS SOCIOLOGY

The shape of the ‘modern’ psychiatric service has, therefore, been defined as much by what it was against, as by what it was for. As a corrective to the hospital-based treatment of schizophrenia, there had to be a transfer of resources away from in-patient services. Although this was reasonable in the age of the physically remote asylum, it has continued well after long-stay beds have disappeared, with predictably dire consequences for the quality of acute in-patient care. Moreover, avoiding hospital admission irrespective of illness outcome has, almost unhinkingly, become an objective for psychiatric services – and highly inappropriate if the patient actually needs hospital care.

Just as beds defined what was to be avoided, so the emphasis on diagnosis has acquired an unwanted flavour. The problems that patients with schizophrenia face were reformulated as ‘social’. So what the patients need is ‘social care’; this remains the Department of Health’s ‘big idea’ for the future of psychiatry and the sociologically correct answers were formulated in the National Service Framework for Mental Health (Department of Health, 1999). This formalised and extended in a surprisingly concrete way the services required for severe mental illness in England. Psychiatrists were notable only by their exclusion from the process whereby the Framework was developed. Bipolar disorder was not mentioned at all, and the National Service Framework, largely unmodified, remains the dogmatic top-down blueprint against which targets managers continue to measure themselves today.

The idea that there might be specific conditions that require specific effective treatments obviously echoes in a ghostly way the US–UK collaborative diagnostic project. What if it had been the dominant paradigm – the heartland condition? Bipolar disorder is no less debilitating, on comparable measures of morbidity and mortality, than schizophrenia (Clement et al., 2003) and is much more common. Moreover, in almost every respect it would have afforded psychiatry a model within which the medical role is easier to define. The development of this model could have informed psychiatric services in general with a greater balance between medical, psychological and social care.

First, the diagnosis of bipolar I disorder characterised by mania is largely uncontroversial because it is based on observable and obvious changes in behaviour. At least, we have yet to hear anyone claim that mania does not exist. Second, bipolar disorder is, exactly like schizophrenia, a complex phenotype that can include virtually all the key phenomenological entities we recognise in psychiatry – depression, mania, psychosis, anxiety, substance misuse, cognitive impairment, neuroendocrine abnormality, sleep disturbance and distinctively variable illness course. The difference is that we do not pretend otherwise, and the fact that these apparently independent dimensions cluster within the single diagnosis of bipolar disorder is accepted as very challenging. Does it mean that the dimensions are themselves related to each other and severity in one will entail severity in the other – perhaps because of common developmental variations in biology or the cumulative effects of illness? Or are cases of bipolar disorder simply represented by those people who sit on the wrong end of these multiple domains, all of which can behave relatively independently? Can diagnosis usefully continue to be categorical without measuring the dimensions that characterise the disease? These are interesting questions that could also reasonably be asked of schizophrenia, but seldom are. Finally, we are not embarrassed to tell patients they have bipolar disorder. They are often grateful to have a diagnosis that explains more than it obscures.

There are other contrasts with the schizophrenia model which are equally important to clinical practice. The course of illness in bipolar disorder allows a much more meaningful distinction between the needs of patients for symptomatic in-patient respite care when acutely ill and for outpatient-based interventions when comparatively well. The treatment of bipolar disorder is also much less amenable to one-size-fits-all social care, which, like most such provision for dependent groups, tends with time and inattention more to reflect the needs of staff than patients: staff become rather more willing to assess patients’ needs than to try and satisfy them. Bipolar disorder is more likely to challenge clinicians to understand the illness and its treatment in relation to individual and autonomous patients.

Finally, treatment of bipolar disorder demonstrably requires the medical expertise which we should take a pride in. The medications that we have available seem often to require use in combination, which probably reflects the complexity of the phenotype. Therefore, prescribing for patients with bipolar disorder requires knowledge, skill and experience. We make a distinction between acute and long-term medication and seek active involvement by patients in managing acute exacerbations of symptoms. Psychological treatments complement the medical approach, and enhanced care is an objective for all patients (Goodwin, 2003). Psychological interventions can reduce the risk of relapse when added to treatment as usual, and have a pragmatic emphasis on self-monitoring, self-management and education about the illness. Moreover, the indication for the content and the timing of treatment is being rationally defined and refined in controlled trials (Vita & Colom, 2004). The confusion around whether cognitive–behavioural therapy (CBT) is really useful for schizophrenia is telling (Turkington & McKenna, 2003; Durham et al., 2005). Moreover, although being adopted by NICE, as one might say, for the nation, the relevance of CBT for psychosis has been wildly amplified at grass roots level in a way that could never have occurred for bipolar disorder – common sense would prevail when the greatest therapeutic optimist met their first patient with florid mania.

DO PSYCHIATRISTS HAVE A FUTURE AS MEDICAL SPECIALISTS?

These differences illustrate what can occur in psychiatry when modern medical treatments
are allowed to develop unhampered by ideology, compared with what happens when they are not. The social model of schizophrenia was the minority, left-wing, ‘show biz’ cause of the 1960s, a formative time for our late-middle-aged policy makers, who have their own curious heartlands. Its current dominance is neither measured, nor moderate. Its essentially totalitarian spirit has even required a new language – an Orwellian ‘newspeak’ where no one must be said to have an illness, comply with treatment or be a patient. The only possible surprise is that the use of the word schizophrenia has not yet been banned from the National Health Service.

Other losses have been more subtle. Can it be sensible to invest a tenth of what we do on schizophrenia on research in bipolar disorder (Neurosciences and Mental Health Board Strategy and Portfolio Overview Group, 2005)? Why do so many junior doctors leave psychiatry because of the role it currently offers them (Lambert et al, 2006)? Most doctors may feel marginalised by managers with regard to resource allocation: psychiatry appears to us unusual in the extent to which managers literally think they know how we should do our jobs. Doctors have a training that brings scientific rigour to what they observe and how they treat. As doctors we also have a broader base in general medicine than most other disciplines involved in psychiatry. A good doctor must be able to make a difference to an individual patient. However, our assumptions and allegiances – our heartland – must be fruitful, not a barren wilderness of good intentions.

We could still develop a more interesting role for doctors in psychiatry because there are effective evidence-based treatments for a wide range of specific conditions, not just bipolar disorder. We happen to know bipolar disorder best and we have been appalled by the difficulties faced by people with bipolar disorder in the current model of secondary services. However, little seems likely to change if schizophrenia continues to occupy such a central and distorting position in our thinking. Why should one condition continue to be so dominant? In general medicine, it would seem ludicrous if the decision was made by the Department of Health to restructure all care around the model of diabetes. To continue to make schizophrenia the paradigm condition in psychiatry is against the interests of psychiatrists and, more importantly, of our patients.

REFERENCES


Correspondence: Dr Guy M. Goodwin, University Department of Psychiatry, Warneford Hospital, Oxford OX3 7JX, UK. Email: guy.goodwin@psych.ox.ac.uk

(First received 29 January 2007, final revision 23 March 2007, accepted 5 April 2007)
Affective neuroscience and psychiatry

NEIL A. HARRISON and HUGO D. CRITCHLEY

Summary

Affective neuroscience addresses the brain mechanisms underlying emotional behaviour. In psychiatry, affective neuroscience finds application not only in understanding the neurobiology of mood disorders, but also by providing a framework for understanding the neural control of interpersonal and social behaviour and processes that underlie psychopathology. By providing a coherent conceptual framework, affective neuroscience is increasingly able to provide a mechanistic explanatory understanding of current therapies and is driving the development of novel therapeutic approaches.

Declaration of interest

None.

Funding detailed in Acknowledgements.

Over the past 25 years there has been a revolution within neuroscience, characterised by recognition of the importance of emotion to behaviour and subjective experience (Dalglish, 2004). This revolution and the consequent snowballing of studies, now subsumed within affective neuroscience, was initiated and led by a key group of experimental, cognitive and social psychologists, anatomists, neuroscientists, clinical neurologists and psychiatrists. Affective neuroscience has challenged ‘standard’ cognitive models to account for empirical and clinical evidence of ‘emotion’ influencing, processes such as attention, perception, learning and memory at every level. Affective neuroscience focuses on brain function and how emotions are ultimately linked to genetic imperatives, primary motivations and reinforcement learning. More recently, affective neuroscience has extended its reach beyond the individual to address dynamic influences on social and economic behaviour.

Well-being and psychiatric dysfunction are necessarily measured against subjective emotional experience. Neurophysiological understanding of emotional disorders, including depression and bipolar disorder, provides a broad framework that may then usefully be applied to other psychiatric conditions, to address biological determinants of stress responses, disorders of personality and even prognosis across mental illness. This has contributed to a systematic integration of previously disparate genetic, neurochemical and psychodynamic models. This process continues to enrich the conceptual language empowering both clinician and patient and driving the development of novel diagnostic tools and therapeutic interventions. Neurobiological accounts of emotional behaviour, interpersonal and social interactions are increasingly plausible and no longer represent merely ‘retreats into organicity’ irreconcilable with psychosocial formulations. Affective neuroscience does not aim for a fully mechanistic account of emotional social mechanisms, but rather to provide insights into control and influence of emotion without the constraints of established disciplinary boundaries. This may be illustrated in our increasing understanding of the neural processes mediating the effect of social and psychological stress on mood disturbance, memory impairments and enhanced risk of mortality. These cross-disciplinary approaches open exciting new treatment options.

EMOTIONAL LEARNING

To illustrate how advances within affective neuroscience have the potential to influence psychiatry, we have predictably chosen to focus our discussion around the ‘most limbic of brain regions’, the amygdala. The contribution of the amygdala to affective behaviour was recognised within MacLean’s concepts of the visceral/limbic/mammalian brain (MacLean, 1990). Weiskrantz (1956), following observations by Klüver & Bucy (1937) on behavioural effects of temporal lobectomy, showed that focal bilateral amygdala lesions had a detrimental impact on the social and emotional behaviour of monkeys. More recent studies of patients with bilateral amygdala lesions reveal impairments in processing of social and emotional cues, notably the recognition of facial or auditory expressions of fear (Adolphs et al, 1994). As a consequence, amygdala dysfunction is linked both theoretically and empirically to psychiatric disorders in which social behaviour is compromised. These include autism, schizophrenia and psychopathy. Practitioners can now conceptualise ‘biological’ impairments in emotional understanding/behaviour as originating in focal dysfunction of regions such as the amygdala.

The amygdala supports the detection and learning of motivational significance. In animal experiments, this is typically illustrated in fear conditioning studies (i.e. the learning of threat). Fear conditioning represents a basic model for understanding the pathogenesis and maintenance of anxiety disorders, including post-traumatic stress disorder and panic. The role of amygdala in fear conditioning has been extensively described in rodents (LeDoux, 1996). In humans, the contribution of the amygdala to processing threat is evident in studies of patients with lesions and from neuroimaging, where amygdala activity is now almost treated as a biomarker of functional integrity within neural systems concerned with emotion. Findings such as these are now beginning to inform psychiatric treatment. Cognitive–behavioural therapy (CBT) remains a gold standard therapy for anxiety disorders by engendering ‘unlearning’ of fear responses through exposure and habituation. Animal studies indicate that both learning and unlearning (extinction) of threat are dependent on glutamate/N-methyl-D-aspartic acid (NMDA) receptors within the amygdala dependent, a pharmacological mechanism that neuroscientists have stimulated to enhance this behavioural extinction process. Early trials suggest that boosting glutamate/NMDA receptor function with cycloserine, a partial agonist at the NMDA receptor, enhances exposure-based CBT, with promising results in the treatment of acrophobia (fear of heights) and social anxiety (Hofmann et al, 2006). These findings result from a growing awareness of the important effects of emotional
processes on basic cognitive functions and could not have resulted from models of memory based on standard cognitive neuroscience.

**SOCIAL BIOLOGY**

Recent widespread clinical recognition that pervasive neurodevelopmental disorders, including Asperger syndrome, place demands on clinical services for adults of working age have forced a reappraisal of conventional psychiatric practice. In parallel, neuroscience has explored the neural mechanisms through which we understand other people (i.e. the cognitive component of interpersonal interaction, which was previously a preserve of psychodynamic psychotherapy). Theoretically, we can understand the experience or intentions of other people by ‘simulation’. Mirror neurons represent a convincing biological instantiation of simulation. Mirror neurons are located within premotor and motor cortex and in primates respond both when performing a specific action and when viewing another monkey performing that specific action (Rizzolatti et al, 1996). Many studies now show activity in the human brain, mimicking or mirroring the performance of perceived actions, which is consistent with a testable neural model for perspective taking, intentional stance and theory of mind. Emotional empathy has been related to co-activation of the same brain regions when observing actual physical pain and observing one’s partner in pain (Singer et al, 2004). Similarly, brain responses to the distress of others engage unconscious autonomic bodily responses that mimic, in sympathy, the observed emotional cues (Harrison et al, 2006). Together these studies expand perception-action principles beyond classical ‘mirror’ regions and suggest that a correspondence between observed and experienced sensations, actions and feelings may be a more general feature of the human brain. Behavioural and neural evidence for robust simulatory systems actually predicts individual differences in emotional empathy. Conversely, individuals with autistic-spectrum disorders and developmental psychopathy show attenuated activation of brain ‘mirror regions’ when observing emotion in others. Such findings can have implications for the future diagnosis and monitoring of disorders of empathy.

**INTEGRATION**

The capacity for affective neuroscience to cross disciplines is illustrated within psychosomatic medicine, a field as relevant to primary care as liaison psychiatry. Broadly speaking, advances are being made in unpacking the mechanisms through which emotional trauma and stress impair cognitive, emotional and physical well-being. Adaptive physiological responses to acute physical and psychological trauma may have pathological effects on the brain and body if the challenges are extreme or prolonged. Stress hormones such as cortisol represent one mechanism. Lifelong stress (even within a general healthy population) is associated with reductions in hippocampal volume that reflect diminished cognitive (especially memory) and behavioural resources. Such findings offer a perspective on clinical psychiatry, as they address core processes underlying vulnerability to psychopathology. Psychosomatic medicine, within affective neuroscience, examines the health consequences of mind–body interactions. This has particular relevance to psychiatry where high rates of physical morbidity require appraisal in the context of potentially cardiotoxic and metabolic effects of medication.

**BROADENING THE NET**

Advances in methods for human brain imaging have assisted the affective neuroscience revolution. The capacity to observe the ‘brain in action’ at fairly high spatiotemporal resolution increasingly informs our understanding of physiological mechanisms underpinning human experience. In the clinical context, identification of functional and structural biomarkers contributes to the characterisation of psychiatric conditions and may enhance the monitoring of clinical course, and even targeting of treatments. Functional signatures in brain activity associated with depression and obsessive-compulsive disorder (OCD) were recognised relatively early. Subsequent studies of both healthy people and patients, using positron emission tomography and functional magnetic resonance imaging (Drevets et al, 1997), support the proposal that abnormalities in subgenual cingulate function in depression may relate to low mood and can predict treatment responsiveness. Animal studies had already linked this region closely with vegetative homoeostatic control. These findings, coupled with the observation of abnormalities of subgenual structure in individuals with unipolar depression, led to neurosurgical targeting of this region in treatment-resistant depression using deep brain stimulation resulting in a marked and sustained symptomatic improvement (Mayberg et al, 2005). Likewise, findings from neuroimaging studies are contributing to targeting of selective fronto-striato-thalamic circuits for the control of severe OCD and Tourette syndrome.

**PROSPECTS**

Biological psychiatry, particularly in relation to emotional disorder, is recovering from a position of low status within the hierarchy of scientific priorities. Affective neuroscience embodies a resurgence of interest in emotional behaviour within biological and social sciences. Recognition that emotion influences a broad spectrum of human functioning suggests that much of biological psychiatry could be rebranded clinical affective neuroscience. Integration of knowledge across disciplines represents an exciting future for psychiatry research and the examples emerging from affective neuroscience are promising. Molecular science already contributes to this integrative process (e.g. in studies of gene–environment interactions which underpin emotional development, behaviour and psychiatric vulnerability). Consideration of biological, psychological and social aspects of psychiatric disorder is fundamental to clinical practice, and it is encouraging that these interactions are becoming central to both basic and clinical research.

Finally, affective neuroscience provides a robust investigative framework for exploring the fundamentals of adaptive emotional behavioural and psychiatric morbidity.

**ACKNOWLEDGEMENTS**

We thank Marcus Gray for helpful comments. N.A.H. and H.D.C. are funded by a Wellcome Trust programme grant to H.D.C.

**REFERENCES**

Management of insomnia: treatments and mechanisms

SUE WILSON and DAVID NUTT

Summary  Management of insomnia is an interesting subject at present. New drug treatments are now becoming available after a relatively static period since the development of the Z-drugs in the 1990s. Moreover, more evidence is coming to light about the length of drug treatment and the effectiveness of psychological therapies. This article briefly describes current treatments, both evidence-based and common practice, and goes on to describe some emerging approaches.

Declaration of interest  SW and D.N. have received funding from several pharmaceutical companies with an interest in the drug treatment of sleep disorder.

Insomnia is the subjective experience of poor or unrefreshing sleep, usually with some objective evidence of reduced time asleep or delayed sleep onset, although often the subjective experience of suffering may appear more than expected from the degree of sleep shortening. If this occurs in the absence of psychiatric disorder, it is called primary insomnia. Insomnia in other psychiatric disorders is called secondary insomnia and is very common in depression, with up to 90% of patients with sleep disturbance. Sleep abnormalities predict poor response to cognitive–behavioural therapy (CBT) (Thase et al, 1995) and often continue to disrupt life even when mood has improved, contributing to the risk of relapse. In mania, insomnia often precedes a relapse and may be a useful warning sign of imminent mood swings. The loss of sleep may accelerate the upward mood shift, just as sleep deprivation can elevate mood in people with depression. Insomnia might offer an important signal of impending illness and an opportunity for prophylactic interventions.

DRUG TREATMENT

When measures to improve sleep habits have failed (see Morin & Espie, 2003), insomnia is usually treated with a variety of drugs – either those that increase brain inhibition through the GABA-A benzodiazepine receptor system or those that decrease excitation through blocking 5-hydroxytryptamine (5-HT) or histamine H₁ receptors.

GABA–A receptor drugs

Benzodiazepines and the Z-drugs (zolpidem and zopiclone, and to a lesser extent zaleplon) are the most commonly used hypnotics which act on the benzodiazepine receptor, with the less selective agents clomethiazole and choral hydrate having poor safety profiles. The Z-drugs have a better pharmacokinetic profile than the older benzodiazepines (Nutt, 2005), but are equally efficacious (Dundar et al., 2004; http://www.nice.org.uk/TA077) and are the drugs of choice to avoid daytime carry-over effects. Doses are usually halved in older adults (see British National Formulary; http://www.bnf.org) but elderly patients do tend to get up in the night and even short-acting GABA-ergic drugs can compromise balance and cognition early in the night (Allain et al., 2005). The length of time that hypnotics show efficacy is less clear. In clinical practice many patients are treated with hypnotics for many months or longer. In a placebo-controlled trial of the active (S) enantomer of zopiclone (eszopiclone) efficacy was maintained over 6 months (Krystal et al., 2003) and there are now open-label continuation data for 12 months (Roth et al., 2005). These long-term controlled data showing continued efficacy of a hypnotic will be reassuring to patients and their treating doctors.

Nevertheless, withdrawal reactions with rebound insomnia are still seen in some patients, even with the newer agents, although the timing of these varies with the half-life of the drug. Thus rebound insomnia would be expected on the first drug-free night with drugs with short half-lives or several nights later with drugs with long half-lives.

Antidepressants

Tricyclic and some other classes of antidepressants as well as antipsychotics have long been used for the treatment of insomnia, whereas selective serotonin reuptake inhibitors (SSRIs) generally disrupt sleep early in a course of treatment. This effect of SSRIs on alertness can be offset by sedative antidepressants such as trazodone, probably because they block 5-HT₂ receptors which are being overstimulated by an increase in 5-HT (Kaynak et al., 2004). Other 5-HT₂ antagonist antidepressants such as nefazodone (Hicks et al., 2002) and mirtazapine (Winokur et al., 2003) have been shown to reduce insomnia in depression, especially early in treatment. There are no controlled studies of the hypnotic efficacy of low-dose amitriptyline but despite this it is fairly common in primary care practice to use 10 or 25 mg amitriptyline to promote sleep. At this dose amitriptyline is probably acting mostly as a histamine H₁ receptor antagonist, although a degree of 5-HT₂ and cholinergic muscarinic antagonism may also contribute.

Antihistamines

Antihistamines have sedative effects and are sold over the counter as sleeping medications. There is limited evidence that over-the-counter antihistamines work, although recently some benefits have been reported for diphenhydramine in mild insomnia. More profound effects on sleep have been reported for both promethazine and hydroxyzine, although neither is available over the counter and they have quite long half-lives so are likely to cause hangover.

Antipsychotics

For decades sedative antipsychotics, especially thioridazine and chlorpromazine, were used to treat more serious insomnia. More recently, as the cardiac safety of these drugs – especially thioridazine – has been questioned, atypical antipsychotics have been used in this role (most usually olanzapine and quetiapine). These act by blocking
5-HT₂ and H₁ receptors as well as muscarinic and α₂-adrenoceptors, all of which contribute to sedation, but there is no published controlled trial of atypical antipsychotics in insomnia.

**Melatonin receptor drugs**

Melatonin is a hormone that helps to regulate circadian rhythms. There is a strong public perception of melatonin as a sleep-promoting agent, which leads many people to self-treat with supplies obtained over the internet or from other countries. However, despite its obvious appeal, clinical efficacy data in primary insomnia are very slight (see Buscemi et al., 2005), although there is some evidence of efficacy in disorders such as jet lag and delayed sleep phase syndrome. Melatonin is also prescribed quite often by child psychiatrists because it seems to have good tolerability, but its efficacy has not been proven in large controlled studies (Buscemi et al., 2006). One such study is being conducted at present by the National Health Service Technology Assessment Programme. In 2005, a synthetic analogue of melatonin, ramelteon, was granted a licence for the treatment of insomnia in the USA. Ramelteon acts in the same way as the natural hormone to stimulate melatonin receptors and has been shown to promote the onset of sleep in placebo-controlled trials (Erman et al., 2006).

**Psychological therapy for insomnia**

Psychological therapies for insomnia have been shown to be effective (Morin et al., 2006; Espie et al., 2007). Once a person experiences insomnia, worry about sleep itself often perpetuates the problem. Currently used psychotherapy for insomnia is a combination of behavioural and cognitive strategies which are given to individuals (Jacobs et al., 2004) or to groups (Jansson & Linton, 2005). The behavioural approaches reinforce natural sleep-initiating and maintaining processes, such as sticking to routines, minimising time in bed spent awake (so limiting ruminations) and lowering physical and psychological arousal at bedtime. Cognitive therapy is based on knowledge of the presence of pre-sleep cognitions in patients who find it difficult to get to sleep. The CBT challenges these thoughts to bring about cognitive restructuring.

Cognitive-behavioural therapy for the treatment of insomnia has been evaluated in a number of studies with varying results, probably because of differences in patient groups and the nature, duration and setting of the treatment. Most reports reveal some improvements, which although modest in terms of sleep onset time and total sleeping time, are often greater in areas such as increased quality of life and decreased anxiety about sleep (Green et al., 2005).

A number of practical issues need to be considered when organising and evaluating these treatments, including selection of participants, access during working hours (many people with insomnia still work) and willingness to engage in the group. Therapists trained in CBT are rarely available as they are usually fully committed to patients with severe mental illness. However, CBT is effective and if available locally should be used for chronic insomnia. It is important to note that a person need not be off medication for the treatment to work.

**Conclusions**

Insomnia is common, especially in people with psychiatric disorder. It is often persistent and disabling and contributes to poorer treatment outcomes and lower quality of life. A summary of treatments for insomnia with their mechanisms of action is given in Table 1.

Current hypnotics that act as agonists of the GABA-A benzodiazepine receptor system, especially the Z-drugs, are safe and effective and new data with eszopiclone shows efficacy for 6 months. However, some people do have trouble in stopping these drugs, owing in part to rebound. Antidepressants with 5-HT₂ and H₁ receptor-blocking properties are useful for promoting sleep, but tricyclic antidepressants are considerably less safe in overdose so should be used with caution. Antipsychotics have a role in severe insomnia associated with other psychiatric disorders, especially depression and psychosis.

Melatonin has some utility in sleep phase disorders (e.g. jet lag). The new synthetic analogue of melatonin, ramelteon, is effective in sleep-onset insomnia. Psychotherapeutic approaches, especially CBT, can be effective for primary insomnia and can work well in group settings.

**Table I Efficacy and mechanisms in the treatment of insomnia**

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Effect on sleep</th>
<th>Brain action responsible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Z-drugs and benzodiazepines</td>
<td>↓ time to fall asleep ↓ total sleep time (depending on half-life)</td>
<td>↑ GABA function (allosteric modulation)</td>
</tr>
<tr>
<td>Trazodone, mirtazapine, olanzapine, quetiapine</td>
<td>↓ total sleep time ↓ awakening</td>
<td>Blocks 5-HT₂ receptors</td>
</tr>
<tr>
<td>Ramelteon</td>
<td>↓ time to fall asleep ↓ awakening</td>
<td>Melatonin agonist</td>
</tr>
<tr>
<td>Promethazine'</td>
<td>↓ total sleep time ↓ awakening</td>
<td>Blocks histamine H₁ receptors</td>
</tr>
<tr>
<td>CBT</td>
<td>↓ time to fall asleep ↓ total sleep time ↓ awakening</td>
<td>↓ worry → ↓ arousal</td>
</tr>
</tbody>
</table>

CBT, cognitive–behavioural therapy.

1. Only one, single-night, single-dose study available.

**References**


Antipsychotic treatment of primary delusional parasitosis

Systematic review

PETER LEPPING, IAN RUSSELL and ROLAND W. FREUDENMANN

Background Little is known about the treatment of delusional parasitosis with typical and atypical antipsychotics.

Aims To evaluate the effectiveness of typical and atypical antipsychotics in primary delusional parasitosis (delusional disorder, somatic type).

Method A systematic review was conducted.

Results No randomised trials were found and hence we collected the best evidence from 16 other trials and case reports, separating primary from other forms of delusional parasitosis. Studies using typical antipsychotics showed partial or full remission in between 60 and 100% of patients. Analysis of selected patients with primary delusional parasitosis showed that typical and atypical antipsychotics were effective in the majority, but that remission rates did not differ significantly between typical and atypical antipsychotics.

Conclusions In the absence of controlled trials there is limited evidence that antipsychotics are effective in primary delusional parasitosis. Rigorous studies are needed to evaluate their effectiveness and to compare typical and atypical antipsychotics directly.

Declaration of interest PL. has received fees from Lilly, Otsuka and AstraZeneca for educational talks.

Delusional parasitosis is characterised by the fixed belief that one is infested with parasites or small living creatures although there is no medical evidence for this (Arnold, 2000; Freudenmann, 2002). Patients usually complain about itching that they ascribe to the presence of animals in or under the skin. The belief is usually held with delusional intensity but the severity of the delusional intensity can vary. The annual prevalence of delusional parasitosis is estimated at 80 cases per million inhabitants, with a yearly incidence of 20 per million (Trabert, 1997).

History It is usually said that delusional parasitosis was first described by Thibierge in 1894 and Perrin in 1896. However, according to Trabert's comprehensive historical literature search (Trabert, 1993, 1997), the clinical picture was first mentioned by Robert Willan in 1799 and Johann Heinrich Jordens in 1801. According to our research neither author suspected a psychiatric aetiology. Moreover, Musalek (1991) discovered a patient with delusions of intestinal parasitosis (Enterozooswahn) in an article from 1843 (Charcellay, 1843).

Pathogenesis Delusional parasitosis is a non-specific syndrome rather than a single disorder. It can occur as a delusional disorder, meeting ICD–10 criteria for persistent delusional disorder (World Health Organization, 1993) and DSM–IV–TR criteria for delusional disorder, somatic type (American Psychiatric Association, 2000). This is what clinicians usually mean when they refer to ‘delusional parasitosis’, although it accounts only for about 40% of all patients with such symptoms (Trabert, 1995). We will call this type of delusional parasitosis ‘primary delusional parasitosis’ following Ganner & Lorenzi’s (1975) concept of ‘reiner Dermatozooswahn’ (German ‘rein’ = primary or pure). The diagnosis of primary delusional parasitosis can be made only after real infection or other underlying medical or psychiatric conditions have been excluded, because delusional parasitosis can be associated with several physical illnesses, psychiatric disorders or intoxications (Magnan & Saury, 1889; Ekberg, 1938; Huber, 1957; Berrios, 1985; Freyne & Wrigley, 1994; Freudenmann, 2002). Delusional parasitosis can also occur as a folie a deux or folie à trois (shared psychotic disorder; Trabert, 1995) as well as by proxy (Nel et al., 2001). Delusional parasitosis syndromes can thus be classified according to their pathogenesis (Appendix 1).

Clinical management

The clinical management of patients with delusional parasitosis is a challenge, as patients are often reluctant to engage in a meaningful therapeutic relationship because of their somatic concept of the illness. Thus they seek help from general practitioners, dermatologists or pest control companies but refuse psychiatric referral or therapy. Usually, it is difficult to obtain informed consent to treat patients with delusional parasitosis with antipsychotics. Therefore experienced clinicians tell their patients that the antipsychotics are effective ‘against the itch’ or the ‘problems with the pests’ in order not to have to lie. A few practical guidelines have been proposed (Musalek, 1991; Driscoll et al., 1993; Winsten, 1997; Freudenmann, 2002).

Another approach to achieve a better therapeutic relationship was developed in the late 1980s. Specialised out-patient clinics were located in dermatology clinics to acknowledge the patients’ non-psychiatric concept of their illness (Musalek & Kutzer, 1989; Musalek et al., 1989; Musalek, 1991; Trabert, 1993). However, even these ‘low threshold’ settings have often failed to allow the establishment of a sufficient therapeutic alliance. Trabert’s study in Homburg, Germany stated that 20 of 35 patients (57%) were seen for less than 3 months (Trabert, 1993). Despite all these efforts, many patients lose faith in professional medicine and resort to dangerous self-therapies such as excessive skin cleaning with chemicals or pesticides (Freudenmann, 2002).

Antipsychotic treatment

For adequate treatment of delusional parasitosis it is necessary to differentiate...
between the different forms (Berrios, 1985; Freudenmann & Schönfeldt-Lecuona, 2005). Although antipsychotics provide the main treatment for primary delusional parasitosis, they are used only symptomatically for delusional parasitosis secondary to somatic diseases, which mainly requires adequate therapy of the underlying disorder. Even in recent years, many sources recommended the use of the typical antipsychotic pimozide in delusional parasitosis (Driscoll et al., 1993; van Vloten, 2003), although pimozide is no longer a first-line antipsychotic because of concerns about drug safety (high risk of extrapyramidal symptoms, longer QTc interval and drug–drug interactions (Food and Drug Administration, 1996; National Institute for Clinical Excellence, 2002; Benkert & Hippius, 2005).

Several case reports have indicated the beneficial effects of atypical antipsychotics in primary delusional parasitosis, but evidence for these is still limited to risperidone (Gallucci & Beard, 1993; Freyne et al., 1999; Moretti & Varga, 2000), quetiapine (Kim et al., 2003), olanzapine (Le & Gonski, 2003) and amisulpride (Lepping et al., 2003).

Although it is often stated that there is a lack of randomised controlled trials of the use of antipsychotics (including pimozide) in delusional parasitosis (Driscoll et al., 1993; Trabert, 1995; Freudenmann & Schönfeldt-Lecuona, 2005), we know of no systematic review on this topic. Moreover, no antipsychotic is licensed for the treatment of delusional parasitosis. We therefore undertook the first systematic review of the effectiveness of typical and atypical antipsychotic treatment for primary delusional parasitosis (meeting ICD–10 F22.0 criteria for persistent delusional disorder or DSM–IV–TR criteria for delusional disorder of somatic type) in order to determine whether: typical antipsychotics are effective in treating primary delusional parasitosis and are more effective than placebo; atypical antipsychotics are effective in treating primary delusional parasitosis and more effective than placebo; atypical antipsychotics are more or less effective than typical antipsychotics.

**METHODS**

**Search strategy**

Our first priority was to discover randomised controlled trials that addressed the study questions. We tried to identify all available works on delusional parasitosis published in English, German, French, Spanish, Portuguese, Italian, Dutch or Hungarian before December 2005. A comprehensive search of EMBASE, Medline, Psycinfo, PsyCIt and Psyndex was performed using the search terms ‘delusion(s) of parasitosis’, ‘delusional parasitosis’, ‘delusion(s) of infestation’, ‘*parasitosis*’, ‘monosymptomatic hypochondriacal psychosis’, ‘parasitophobia’, ‘entomophobia’, ‘acarophobia’, ‘Dermatozoenwahn’, and ‘Ekbohm’s syndrome’ (discarding papers on the ‘burning feet syndrome’ which has also been labelled with this eponym). We checked the reference lists of identified articles. We searched the internet using Google, and textbooks of psychiatry, theses, unlisted journals and conference proceedings by hand. We wrote to pharmaceutical companies producing substances often used to treat delusional parasitosis and known authorities in the field of delusional parasitosis (e.g. Wolfgang Trabert, Emden, Germany, and Marc Bourgeois, Bordeaux, France) to identify unpublished data on the use of antipsychotics in delusional parasitosis. We searched for continuing trials via two websites – ClinicalTrials (http://www.clinicaltrials.gov) and Current Controlled Trials (http://www.controlled-trials.com).

**Assessment of literature**

P.L. and R.W.F. independently assessed whether all retrieved works dealt with delusional parasitosis in general or primary delusional parasitosis, and whether the intervention consisted of typical or atypical antipsychotics. We also assessed whether any of these were randomised controlled trials.

In the absence of randomised controlled trials we planned to gather sound evidence from other studies meeting defined inclusion criteria (see Appendix 2). In particular we sought well-designed quasi-experimental and observational studies relevant to our research questions. We included all open studies with either prospective design or more than 30 patients. We then summarised in structured form the main findings of the 16 studies meeting these minimal criteria (Table DS1, data supplement to online version of this paper). We assigned individual outcomes between three main categories: no effect (0); partial remission (i.e. some response) (1) and full remission (2). To strengthen our conclusions we also tried to separate primary delusional parasitosis from secondary delusional parasitosis.

P.L. and R.W.F. also selected all case reports containing information on diagnosis of primary delusional parasitosis, gender, age, antipsychotic medication used and dose, and clinical outcome on the same 3-point scale after 4 weeks or more. In this way we applied Trabert’s case-based meta-analysis which is designed for uncommon syndromes that cannot be studied in a traditional randomised controlled trial (Trabert, 1995). We also separated patients treated with typical and atypical antipsychotics. Although this approach is subject to publication bias because it considers only published cases, we aimed to increase comparability between studies. To judge the success of this strategy we tested whether clinical outcomes differed significantly between studies.

**RESULTS**

**Literature search**

We identified a total of 368 works on delusional parasitosis in general, including poster presentations, in December 2005 (Fig. 1). We found most on Medline (n=191), and in review articles and the comprehensive theses of Musalek (1991) and Trabert (1993). At least half were not in English – most of these were in German, but some were in Italian and French. The full bibliography can be obtained from the authors on request.

Before the psychopharmacological era, which began with the discovery of chlorpromazine in 1952, only 31 works on delusional parasitosis were retrieved. The majority (n=223) were published between 1952 and the launch of risperidone in about 1990. The remainder (n=114) were published after 1990, but many of these did not examine the use of atypical antipsychotics.

**Absence of randomised controlled trials**

Our systematic search found no randomised controlled trials on the effects of typical or atypical antipsychotics in either primary or other delusional parasitosis. When accessed in December 2005, the Clinical Trials (http://www.clinicaltrials.gov) and Current Controlled Trials (http://www.controlled-trials.com) websites gave no indication of unpublished or current
randomised controlled trials. Our evaluation of the literature therefore relied on results from other studies.

Effect of typical antipsychotics

Table DS1 (see data supplement to online version of this paper) summarises the 16 quasi-experimental or observational studies which primarily used typical antipsychotics and met our inclusion criteria (Frithz, 1979; Hamann & Avnstorp, 1982; Munro, 1982; Lyell, 1983; Ungvari, 1983, 1984; Ungvari & Vladar, 1984, 1986; Lindskov & Baadsgaard, 1985; Bourgeois, et al, 1986; Reilly & Batchelor, 1986; Musalek, et al, 1989; Paholpak, 1990; Trabert, 1993, 1995; Srivivasan et al, 1994; Zomer et al, 1998; Bhatia et al, 2000). The studies showed aggregate partial and full remission rates between 60 and 100% after treatment with typical antipsychotics. Unfortunately, the majority were not limited to primary delusional parasitosis. Nevertheless they suggest a generally good outcome for primary and other forms of delusional parasitosis whenever antipsychotic treatment can be established.

In primary delusional parasitosis, aggregate partial and full remission rates with pimozide ranged from 67% (n=66; Lyell, 1983) through 89% (n=9; Munro, 1982) to 100% (n=10; Ungvari & Vladar, 1984, 1986), n=18; Ungvari, 1983, 1984). In mixed samples with primary and other forms of delusional parasitosis, aggregate partial and full remission rates with pimozide were similar and varied from 61% (n=33; Zomer et al, 1998), through 87% (n=52; Bhatia et al, 2000) to 91% (n=11; Hamann & Avnstorp, 1982). A high rate of side-effects such as sedation, extrapyramidal symptoms and depression was noted in several studies using pimozide (38%; Ungvari, 1983, 1984 and 73%; Hamann & Avnstorp, 1982).

The only two placebo-controlled trials in delusional parasitosis both used pimozide (Hamann & Avnstorp, 1982; Ungvari & Vladar, 1984, 1986). However, they are limited by a lack of randomised allocation to the treatment groups and small samples (n=10 or 11 respectively). Only Ungvari & Vladar (1984, 1986) treated patients with primary delusional parasitosis.

Studies not specific for particular antipsychotics demonstrated aggregate partial and full remission rates between 82% (n=35, not only primary delusional parasitosis, Trabert, 1993) and 89% (n=19, only primary delusional parasitosis, but some patients were treated with electroconvulsive therapy; Srivivasan et al, 1994).

The only study that investigated traditional depot antipsychotics found an aggregate response and remission rate of 93%, even in patients that could not be treated with oral medication (Frithz, 1979).

The sample consisted only of patients with primary delusional parasitosis (n=15). Another study using haloperidol reported a 100% response rate (but no full remissions) (Paholpak, 1990); nine of ten patients in this study had primary delusional parasitosis.

Across the different typical antipsychotics used, an effect of antipsychotic medication in delusional parasitosis was noted after about 3–6 weeks (Hamann & Avnstorp, 1982; Trabert, 1995). Studies came to different conclusions as to whether or not it is necessary to continue antipsychotics after successful acute therapy.

Symptoms of delusional parasitosis that were associated with major depression could be treated successfully with antidepressants (Musalek et al, 1989; Trabert, 1993; Bhatia et al, 2000).

One small study indicated that electroconvulsive therapy might be effective in patients with primary delusional parasitosis (Srivivasan et al, 1994).

A survey of British dermatologists suggested that combining psychopharmacological and dermatological treatments (local and systemic) is superior to single therapeutic approaches. Of particular note is that therapy without psychotropic medication was ineffective (Reilly & Batchelor, 1986).

We abstracted data from case series and case reports on 92 patients with primary delusional parasitosis treated with typical antipsychotics who met our selection criteria – (see Table DS2 (data supplement to online version of this paper) (Riding & Munro, 1975; Gould & Gragg, 1976; Munro, 1978a,b, 1982; Frithz, 1979; Avnstorp et al, 1980; Ungvari, 1984; Berrios, 1998; Ungvari & Vladar, 1984; Frithz, 1993; Bhatia et al, 2000).

### Table 1 Summary of 92 case reports of the treatment of primary delusional parasitosis with typical antipsychotics

<table>
<thead>
<tr>
<th>Antipsychotic</th>
<th>Dosage</th>
<th>n</th>
<th>Full remission</th>
<th>Partial remission</th>
<th>No effect</th>
<th>Non-adherence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pimozide</td>
<td>1–12 mg/day</td>
<td>53</td>
<td>24</td>
<td>26</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Trifluoperazine</td>
<td>2–15 mg/day</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Haloperidol</td>
<td>1–10 mg/day</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chlorpromazine</td>
<td>150–300 mg/day</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perphenazine</td>
<td>4–12 mg/day</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sulpiride</td>
<td>50–150 mg/day</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Penthiazine</td>
<td>5 mg/day</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thoridazine</td>
<td>75 mg/day</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Depots</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluphenazine</td>
<td>7.5–25 mg</td>
<td>10</td>
<td>7</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flupenthixol</td>
<td>2–20 mg</td>
<td>5</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. See Table DS2 in data supplement to online version of this paper for further details.
Method used for identification and selection of studies of delusional parasitosis.

We were able to identify only 12 case reports on atypical antipsychotics that met our selection criteria (Table 2). Although atypical antipsychotics were effective for the majority of patients with primary delusional parasitosis only three patients achieved full remission. All six patients for whom risperidone was prescribed, achieved full (four) or partial (two) remission; the dosages used ranged from 1 to 8 mg per day.

**Effect of atypical antipsychotics**

We were able to identify only 12 case reports on atypical antipsychotics that met our selection criteria (Table 2). Although atypical antipsychotics were effective for the majority of patients with primary delusional parasitosis only three patients achieved full remission. All six patients for whom risperidone was prescribed, achieved full (four) or partial (two) remission; the dosages used ranged from 1 to 8 mg per day.

**Effect of typical vs. atypical antipsychotics**

The five main studies of patients with primary delusional parasitosis treated with typical antipsychotics report very heterogeneous outcomes (Table 3), suggesting that the studies themselves are very heterogeneous, for example in their inclusion and exclusion criteria (selection bias) or in their definitions of partial and full remission (measurement bias), or that there is publication bias. Furthermore, the eight studies of patients with secondary delusional parasitosis treated with typical antipsychotics reported heterogeneous outcomes. This reduces the value of comparing studies reporting the use of typical primary antipsychotics in secondary delusional parasitosis. However, there was a large difference between the failure rate of 5% in primary delusional parasitosis and that of at least 26% in secondary delusional parasitosis ($\chi^2=18.2$, d.f. = 2, $P<0.001$). It follows that studies of secondary delusional parasitosis have nothing to contribute to the issue of the most effective treatment for primary delusional parasitosis.

The heterogeneity in the reported outcome of patients with primary delusional parasitosis treated with typical antipsychotics reduces the value of comparison with reports of atypical antipsychotics. Although the difference is not statistically significant ($\chi^2=2.6$, d.f. = 2), we cannot conclude that typical and atypical antipsychotics are equally effective because of the innate biases already identified.

**DISCUSSION**

This is the first systematic review of the effectiveness of typical and atypical antipsychotics in the treatment of delusional parasitosis. Our review was based on 368 published works and covered almost twice as many papers as the 193 covered by the most comprehensive review published to date (Trabert, 1993). In contrast to previous review articles (e.g. Aw et al, 2004; Bourgeois & Nguyen-Lan, 1986; Driscoll et al, 1993; Freudenmann, 2002; Lynch, 1993; Slaughter et al, 1998; Wykoff, 1987; Zanol et al, 1998), our review focuses on primary delusional parasitosis. We separated this important form of delusional parasitosis (delusional disorder of somatic type) from symptomatic forms of delusional parasitosis (‘secondary delusional parasitosis’) which cover different nosological entities and require other forms of therapy that focus on the underlying illness. As we were unable to show homogeneity in the outcomes of primary and secondary delusional parasitosis, we did not review other studies examining the effect of antipsychotic medication in other types of ICD–10 F22.0 disorders, which also differ clinically from delusional parasitosis.

Our systematic review identified no randomised controlled trials of the efficacy of typical and atypical antipsychotics in primary delusional parasitosis, probably because the disorder is rare and it is difficult to recruit patients, obtain informed consent, and achieve sufficient adherence
Summary of 12 case reports of the treatment of primary delusional parasitosis with atypical antipsychotics

Our findings show that primary delusional parasitosis is the intramuscular application of traditional depot antipsychotics (Frithz, 1979), because the main problem in clinical management is to convince patients to take oral medication regularly. Injection may be more consistent with patients’ (false) somatic concept of their illness and require less cooperation than oral medication. If the patient agrees to a first depot injection, the delusion may well remit at least partially and further antipsychotic treatment will be accepted. As the only study of this approach (Frithz, 1979) has only 15 patients, larger samples are needed.

Use of typical antipsychotics

Our findings show that primary delusional parasitosis can be effectively treated with typical antipsychotics. Outcome is generally good, although this conclusion is limited by a possible publication bias. We confirm Trabert’s finding that the introduction of typical antipsychotics has substantially improved remission rates (Trabert, 1995). Although the better studies have so far used pimozide (Hamann & Avnstorp, 1982; Ungvari & Vladar, 1984, 1986), the evidence for its efficacy is weak by today’s standards. The level of evidence for its use in primary delusional parasitosis is IIa according to the criteria of the Agency for Health Care Policy and Research (1992), whereas other typical antipsychotics such as haloperidol have only level III evidence. Pimozide should not be used in patients with a high cardiac risk, together with other substances that prolong the QTc interval (Food and Drug Administration, 1996), or in elderly patients with delusional parasitosis.

Another important treatment in primary delusional parasitosis is the intramuscular application of traditional depot antipsychotics (Frithz, 1979), because the main problem in clinical management is to convince patients to take oral medication regularly. Injection may be more consistent with patients’ (false) somatic concept of their illness and require less cooperation than oral medication. If the patient agrees to a first depot injection, the delusion may well remit at least partially and further antipsychotic treatment will be accepted. As the only study of this approach (Frithz, 1979) has only 15 patients, larger samples are needed.

Use of typical antipsychotics

Our systematic review revealed only 12 usable case reports of the use of atypical antipsychotics in primary delusional parasitosis. These provide limited evidence that primary delusional parasitosis can be treated effectively with these drugs. To our knowledge, this is the first complete collection of patients with primary delusional parasitosis treated with atypical antipsychotics, whereas many patients with secondary forms of delusional parasitosis have been reported in recent years (e.g. De Leon et al, 1997; Safer et al, 1997; Kumbier & Kornhuber, 2002; Freudenmann, 2003; Le & Gonski, 2003; Scheinfeld, 2003; Wenning et al, 2003). Thus the evidence for the use of atypical antipsychotics in delusional parasitosis is even weaker than for typical antipsychotics.

Most case reports are available for risperidone, whereas we are not aware of reports on the use of clozapine, ziprasidone or aripiprazole in primary delusional parasitosis. Amsulpride might be a good alternative given that its selective D2/D3 antidopaminergic action resembles that of typical antipsychotics without the same high cardiac risk profile. Risperidone microspheres for intramuscular injection provide a potential new treatment for delusional parasitosis, as this is the only atypical antipsychotic in depot form, but this recommendation is entirely theoretical since there are no reports of the use of risperidone microspheres in delusional parasitosis at present.

Other treatment options

An alternative to these pharmacological strategies is electroconvulsive therapy. The use of electroconvulsive therapy in a patient with delusional parasitosis was first
Outcome of treatment of delusional parasitosis with antipsychotics

<table>
<thead>
<tr>
<th>Outcome study</th>
<th>No effect, refusal, etc., n (%)</th>
<th>Partial remission, n (%)</th>
<th>Full remission, n (%)</th>
<th>Total, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atypical antipsychotics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary delusional parasitosis (five studies)</td>
<td>2 (17)</td>
<td>7 (58)</td>
<td>3 (25)</td>
<td>12</td>
</tr>
<tr>
<td>Typical antipsychotics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary delusional parasitosis</td>
<td>7 (8)</td>
<td>40 (43)</td>
<td>45 (49)</td>
<td>92</td>
</tr>
<tr>
<td>Frithz (1979)</td>
<td>1</td>
<td>3</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>Munro (1982)</td>
<td>1</td>
<td>5</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Ungvari (1984)</td>
<td>0</td>
<td>7</td>
<td>12</td>
<td>19</td>
</tr>
<tr>
<td>Ungvari &amp; Vladar (1986)</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Srinivasan et al (1994)</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>I4 studies with ≤ 4 patients</td>
<td>3</td>
<td>11</td>
<td>14</td>
<td>28</td>
</tr>
<tr>
<td>Mixed delusional parasitosis outcome on 3-point scale</td>
<td>43 (26)</td>
<td>50 (30)</td>
<td>73 (44)</td>
<td>166</td>
</tr>
<tr>
<td>Lindskov &amp; Baadsgaard (1985)</td>
<td>0</td>
<td>4</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>Musalek et al (1989)</td>
<td>12</td>
<td>5</td>
<td>17</td>
<td>34</td>
</tr>
<tr>
<td>Trabert (1993)</td>
<td>6</td>
<td>18</td>
<td>9</td>
<td>33</td>
</tr>
<tr>
<td>Zomer et al (1998)</td>
<td>19</td>
<td>7</td>
<td>7</td>
<td>33</td>
</tr>
<tr>
<td>Bhata &amp; Batchelor (1986)</td>
<td>6</td>
<td>16</td>
<td>30</td>
<td>52</td>
</tr>
<tr>
<td>Mixed delusional parasitosis: binary outcome</td>
<td>27 (29)</td>
<td>67 (71)</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>Hamann &amp; Avnstorp (1982)</td>
<td>1</td>
<td>10</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Lyell (1983)</td>
<td>16</td>
<td>44</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Reilly &amp; Batchelor (1986)</td>
<td>10</td>
<td>13</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

1. \( x^2 \approx 22.1, d.f. = 8, P < 0.001. \\
2. Test for homogeneity of studies with outcome on 3-point scale \( x^2 \approx 43.5, d.f. = 8, P < 0.001. \)
3. Outcome dichotomised, i.e. partial and full remission combined. Test for homogeneity of all studies of mixed delusional parasitosis, \( x^2 \approx 34.6, d.f. = 7, P < 0.001. \)

APPENDICES

Appendix I: Aetiological classification of delusional parasitosis

(I) Primary delusional parasitosis: delusional disorder

Primary delusion according to Berrios (1985), first described by Huber (1957); diagnosis: persistent delusional disorder (ICD–10 F22.0); delusional disorder, somatic type (DSM–IV–TR 297.1)

Special form: as a shared psychotic disorder (ICD–10 F24.22, DSM–IV–TR 297.3)

(II) Secondary forms of delusional parasitosis: secondary to another condition.

(a) Concomitant psychotic symptom in another psychiatric disorder:

(i) Schizophrenia or other psychotic disorders; diagnosis according to underlying psychotic disorder (ICD–10 F22.x, DSM–IV–TR 295, etc.)

(ii) Major depressive disorder with psychotic symptoms or mania; diagnosis according to underlying affective disorder (ICD–10 F31.x, DSM–IV–TR 296, etc.)

(iii) Dementia; diagnosis ICD–10 F00–03, DSM–IV–TR 290, 294.

(b) Delusional parasitosis based on other brain pathologies (‘macroscopic’) or general medical condition:

(i) Brain disorders not mentioned in ICD–10 F0 (e.g. brain neoplasm/infection, stroke); organic damage with secondary delusions’ according to Berrios (1985), first described by Elkon (1938).

(ii) Somatic illness with pruritus or paraesthesia (e.g. diabetes mellitus with neuropathic pain, uraemia, jaundice, cancer); ‘paraesthesia or other somatic pathological sensations with secondary delusions’ according to Berrios (1985); diagnosis: organic hallucinosis or organic delusional disorder (ICD–10 F06.0 or F06.2); psychotic disorder due to ... [indicate the general medical condition] with delusions (DSM–IV–TR 293.81) or with hallucinations (293.82) or persistent delusional disorder (ICD–10 F22.0); delusional disorder, somatic type (DSM–IV–TR 297.1), when delusional parasitosis is not the direct physiological consequence of the somatic illness.

(c) Delusional parasitosis as a substance-induced ‘toxic’ psychosis:

Substance-induced ‘paraesthesia or other somatic pathological sensations with secondary delusions’ according to Berrios (1985), first described by Magnan & Sauvy in 1889 for inebriates (‘signe de Magnan’):

(i) Owing to psychotropic substance: e.g. cocaine, amphetamines; diagnosis: acute intoxication, psychotic disorder; predominantly delusional (ICD–10 F1X.51) or predominantly hallucinatory (ICD–10 F1X.52); substance-induced psychotic parasitosis in comparison with pimozide, the use of atypical antipsychotics might improve side-effects, and thus adherence and patient outcome.

It is important to strengthen this weak evidence in the future. We limited our selection of case series and case reports of delusional parasitosis to those including a minimum data-set for each recruited patient (age, gender, the nature and timing of diagnosis, the name and dose of medication, and the nature and timing of remission on a 3-point scale; Appendix 2).

Implications

Our systematic review generated weak evidence that antipsychotics are effective in treating primary delusional parasitosis. However, in view of the limited evidence, this recommendation is tentative and needs caution in implementation. Since the introduction of atypical antipsychotics, pimozide is no longer the treatment of choice for reasons of drug safety, even though it has the best evidence of effectiveness in treating primary delusional parasitosis. It is important to improve this evidence through rigorous, ideally randomised, studies which compare typical antipsychotics and atypical antipsychotics directly.

described by Harbauer in 1949 and has since occasionally been reported (Baumer, 1951; Bers & Conrad, 1954; Hopkinson, 1970). Srinivasan et al (1994) reported effectiveness in primary delusional parasitosis in a small sample. Electroconvulsive therapy might be a useful option in cooperative refractory patients when antipsychotics are contraindicated or problematic (e.g. in the elderly).

Synthesis

The very heterogeneous outcomes reported by the five main studies of treating primary delusional parasitosis with typical antipsychotics (see Table 3) suggest that the studies themselves suffer from some or all of selection bias, measurement bias and publication bias. Together with these flaws the paucity of evidence on treating primary delusional parasitosis with atypical antipsychotics undermines any comparison of typical and atypical antipsychotics. Despite weaker evidence for the effectiveness of atypical antipsychotics in treating delusional parasitosis in comparison with pimozide, the use of atypical antipsychotics might improve side-effects, and thus adherence and patient outcome.
Appendix 2: Proposed minimum information required for case series and case reports of primary delusional parasitosis

Case report criteria

(1) Diagnosis including confirmation date
(2) Gender of patient
(3) Age of patient
(4) Medication used with dosage
(5) Outcome on 3-point scale: (0) no remission, (1) partial remission, (2) full remission
(6) Length of follow-up (at least 4 weeks after date of diagnosis).

REFERENCES

Agency for Health Care Policy and Research (1992) Agency for Health Care Policy and Research Publication 92-0002. AHCPR.


Jørdens, J. H. (1801) Entomologie and Helminthologie des menschlichen Körpers oder Beschreibung der Würmer und Feinde desselben unter den Insekten und Würmern [Entomology and helminthology of the human body or the description of the inhabitants and enemies of the same among insects and parasites]. Grau.


PETER LEPPING, MD (FRG), MRCPsych, MSc, North Wales Section of Psychological Medicine, Wrexham
Academic Unit, Technology Park, Wrexham, Wales, UK; IAN RUSSELL, PhD, FRCPG, FRCPEd, FFPH, Institute of
Medical and Social Care Research, University of Wales Bangor; Gwynedd, Wales, UK; ROLAND W.
FREUDENMANN, MD (FRG), Department of Psychiatry, University of Ulm, Ulm, Germany

Correspondence: Dr Peter Lepping, North Wales Section of Psychological Medicine, Wrexham
Academic Unit, Technology Park, Wrexham, LL13 7YP, UK. Email: peter.lepping@new-tr.wales.nhs.uk
(First received 9 August 2006, final revision 22 February 2007, accepted 28 February 2007)
Guided self-help in primary care mental health
Meta-synthesis of qualitative studies of patient experience

NAGINA KHAN, PETER BOWER and ANNE ROGERS

Background  There is a gap between the supply of trained cognitive–behavioural therapists to treat depression and demand for care in the community. There is interest in the potential of self-help interventions, which require less input from a therapist. However, the design of effective self-help interventions is complex. Qualitative research can help to explore some of this complexity.

Aims  The study aimed to identify qualitative studies of patient experience of depression management in primary care, synthesise these studies to develop an explanatory framework, and then apply this framework to the development of a guided self-help intervention for depression.

Method  A meta-synthesis was conducted of published qualitative research.

Results  The synthesis revealed a number of themes, including the nature of personal experience in depression; help-seeking in primary care; control and helplessness in engagement with treatment; stigma associated with treatment; and patients’ understandings of self-help interventions.

Conclusions  This meta-synthesis of qualitative studies provided a useful explanatory framework for the development of effective and acceptable guided self-help interventions for depression.

Declaration of interest  None. Funding detailed in Acknowledgements.

METHOD

Although evidence exists for the effectiveness of guided self-help, this effectiveness varies significantly with different types of guided self-help (Anderson et al, 2005; Mead et al, 2005; Salkovskis et al, 2006), which means that there is significant ambiguity over the optimal design of such an intervention to maximise effectiveness and acceptability. For example, how much guidance should be provided? What is the optimal balance between activities that develop the therapeutic alliance and those focused on teaching cognitive–behavioural techniques? How can the materials be designed to engage patients most effectively? How much reliance should be placed on the agency and experience of the individual patient rather than that of the therapist?

Mental health interventions are increasingly complex, involving a number of different ‘active ingredients’ to achieve change (Medical Research Council, 2000). Understanding the contexts and ways in which such interventions achieve their effects is crucial for scientific understanding and effective clinical delivery (Campbell et al, 2000). The ‘phased’ development of complex interventions has been advocated (Medical Research Council, 2000), and there is interest in the role of qualitative methods alongside randomised controlled trials (Donovan et al, 2002). Qualitative research can help to explore some of this complexity and increase our understanding of the way in which interventions are used and experienced.

The qualitative methodology used in this study was meta-synthesis. The technique has some similarity to quantitative meta-analysis, involving the development of an overview of research, but is based on qualitative papers. Meta-synthesis assists knowledge synthesis through a process of re-conceptualisation of themes across a number of published qualitative studies (Noblitt & Hare, 1988). The method has been applied to particular examples in the area of healthcare (Britten et al, 2002; Campbell et al, 2003). The synthesis is derived through the transfer of ‘ideas, concepts and metaphors’ across different studies (Britten et al, 2002), where the interpretations and explanations in the original studies undergo a process of deconstruction, translation and reconstruction as ‘a means to grasp the particulars within the wholes’ (Thorne et al, 2004). A distinguishing characteristic of this method is that translations are not literal, but are concerned with the preservation of meaning across studies (Britten et al, 2002). Relationships between studies can be described in a number of ways (Noblitt & Hare, 1988). For the present purposes, a ‘line of argument’ approach was adopted, where statements about the phenomenon of interest were inferred from the selected studies (Noblitt & Hare, 1988). Our intention was to develop a line of argument about the likely response of patients to a guided self-help intervention.

The basic data were the main concepts reported in each of the individual studies. These concepts were synthesised across the studies to develop new ideas and interpretations. The results of this synthesis formed the basis of an explanatory framework concerning patients’ experience of
depression and its management in primary care. This framework was then applied to the specific pragmatic question driving the study: what factors might influence the effective implementation of guided self-help for depression in primary care?

**Stages of the meta-synthesis**

The meta-synthesis involved a number of stages:

(a) identifying the literature – topic selection, searching for the studies and appraisal of the studies;

(b) data analysis and interpretation – extraction of main findings from the published studies, synthesis of main findings into an explanatory framework, application of the explanatory framework to the guided self-help intervention.

**Identifying the literature**

The initial search for qualitative papers in this study focused on three broad themes: studies of patients’ and professionals’ perspectives on help-seeking and treatment for common mental disorders in primary care; the process of implementation of self-management interventions for chronic conditions; and the use of technologies related to self-management. To identify the primary studies, the Medline, EMBASE, CINAHL and Web of Knowledge databases were searched for the period 2000–2005 inclusive. Separate search strategies were created in line with the three initial themes, with the assistance of a specialist librarian (copies of the exact search strategies are available from the authors, and an example is given as a data supplement to the online version of this paper). Twenty-four potentially relevant papers were found. Initial analysis of the papers demonstrated that the original themes used to structure the search were not clearly reflected in the available literature. Instead, studies could be grouped according to three emergent categories: patients’ perspectives regarding the experience of coping with depression; patients’ perspectives regarding the management of depression within primary care; and patients’ attitudes towards and use of treatments commonly provided for managing depression in primary care (e.g. antidepressants and psychological therapy).

Papers were appraised using the British Sociological Association (BSA) criteria for the evaluation of qualitative research papers (BSA Medical Sociological Group, 1996). Exclusions were made if studies turned out to be insufficiently focused on the topic (e.g. not based on direct experience of depression) or if the paper was not essentially qualitative (as some studies had collected data using qualitative methods but did not analyse the data qualitatively).

**Data analysis and interpretation**

A grid was constructed in which each paper was entered into a separate row, and a description of the concepts derived from each paper added to the grid. The descriptions could involve the author’s own words, or a paraphrase (Britten et al., 2002), in order to reliably retain the meanings and concepts of each study. The entries in this grid were then synthesised by reading the concepts and interpretations off the grid, and establishing relationships between them across the studies, in order to arrive at a broader explanatory framework. This framework was then applied to the specific issue of the delivery of guided self-help in primary care.

**RESULTS**

**Synthesis of main findings into an explanatory framework**

Nine papers were included in the meta-synthesis. Elementary contextual information for the included studies is provided in a data supplement to the online version of this paper. The results of the synthesis are described below and in Table 1.

**Personal experience of depression**

External sources of stress or conflict were drawn upon most frequently to account for the presence of depression. These included conflict with work colleagues or family, chronic illness, events in childhood, material disadvantage and racism (Kadam et al., 2001; Rogers et al., 2001; Burr & Chapman, 2004; Grime & Pollock, 2004). Rather than emphasising symptoms or feelings of depression, respondents’ personal experience was characterised by expressions of being unable to cope, and in particular disturbances to everyday functioning and social roles (with negative consequences for other family members) (Kadam et al., 2001; Rogers et al., 2001; Knudsen et al., 2002; Burr & Chapman, 2004; Maxwell, 2005). Metaphors used by respondents to communicate the experience of depression included being ‘on edge’, ‘churned-up inside’, ‘boxed in’, ‘a volcano bursting’, ‘broken in half’, ‘shut in my own little shell’, ‘a wall of pain’ and ‘prisoner in my own home’ (Kadam et al., 2001). Attempts to overcome such feelings were expressed in terms such as ‘fight’ and ‘conquer’ (Kadam et al., 2001; Glasman et al., 2004).

**Ambivalent help-seeking and the covert presentation of psychological problems**

The experience of depression and failures to cope could lead people to seek help from formal agencies such as primary care. In relation to the decision to seek help, patients’ accounts exhibited a need to leave behind more passive periods when they felt overwhelmed by feelings of their inability to cope (Rogers et al., 2001) or where inaction was leading to negative consequences for other family members, which might in turn lead to guilt (Maxwell, 2005). However, engaging with primary care services was problematic. Patients used primary care because it represented the only place where help was seen to be on offer, rather than through a specific expectation that accessing these services would be helpful. Contact with primary healthcare was relatively insignificant for the individual in the context of pressing problems and adverse circumstances that respondents reported, and patients spent very little time in face-to-face contact with their doctors or with other health professionals (Rogers et al., 2001). Accessing help was viewed as a set of ‘moral actions’ (Maxwell, 2005), further complicated by feelings of shame and lack of legitimacy, which could lead to the covert presentation of psychological problems (Rogers et al., 2001; Burr & Chapman, 2004). Moreover, there was the possibility that accessing help could threaten an already weakened sense of self if it led to discussions about treatments that patients might find unacceptable (such as medication or referral to specialist mental health services). Some patients also exhibited an unquestioning attitude to the quality of care for their problems (Gask et al., 2003).

**Control and helplessness in engaging with treatment**

Patients reported the use of coping strategies, such as distraction or the use of particular locations associated with feelings of safety and control (Kadam et al., 2001).
### Table 1  Main results from the meta-synthesis

<table>
<thead>
<tr>
<th>Main findings from the studies in the review</th>
<th>Synthesis of main findings into an explanatory framework</th>
<th>Application of the explanatory framework to the guided self-help intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Depression related to feelings of powerlessness and lack of control, hopelessness and detachment from the world (Burr &amp; Chapman, 2004)</td>
<td>Personal experience of depression: depression is characterised by a sense of dissociation, loss of everyday function, and a feeling of loss of control</td>
<td>Incorporating personal experience in a structured intervention: materials can be sensitised to people’s own experience of depression through: (a) acknowledging the range of causal factors seen as constituting the phenomena of ‘depression’; (b) using similar metaphors to express feelings associated with depression; (c) highlighting intervention as a method of regaining control</td>
</tr>
<tr>
<td>Help sought because of a perceived inability to carry out everyday functions, loss of control, failure of previous coping strategies and impact of inaction on others (Rogers et al., 2001, 2004; Maxwell, 2005)</td>
<td>Help-seeking from GP related to feelings of legitimacy Rogers et al., 2001; Burr &amp; Chapman, 2004; Maxwell, 2005</td>
<td>Conceptualising the therapeutic environment for the purposes of engagement: the point at which people make contact, prior contact with sources of help and the point in their illness trajectory may be important to review in assessing the acceptability and appropriateness of guided self-help</td>
</tr>
<tr>
<td>Use of metaphors to convey a sense of struggle (Kadam et al., 2001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambiguity over the role of primary care (Kadam et al., 2001; Rogers et al., 2001)</td>
<td>Ambivalent help-seeking and the covert presentation of psychological problems: patients used primary care because it happened to be there or was the only place where help was seen to be on offer. Help-seeking is complicated by influences which result in the covert presentation of psychological problems</td>
<td></td>
</tr>
<tr>
<td>Help-seeking from GP related to feelings of legitimacy Rogers et al., 2001; Burr &amp; Chapman, 2004; Maxwell, 2005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expectations of health services based on prior experience, and affected by feelings of shame, pessimism about effectiveness and threat (such as disruption to sense of self) (Rogers et al., 2004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ambivalence about need for follow-up by doctor after initial diagnosis (Gask et al., 2003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondents concerned about psychological dependency, loss of autonomy and sense of failure to cope associated with use of antidepressants (Grime &amp; Pollock, 2004; Maxwell, 2005)</td>
<td>Control and helplessness in engagement with treatment: patients felt they had to give up control in order to engage with treatments based on biomedical principles, conflicting with their desire to recover a sense of self and usual function</td>
<td>Everyday self-management strategies and guided self-help: building on existing attempts to self-manage as part of an introduction to guided self-help may facilitate acceptance of the treatment</td>
</tr>
<tr>
<td>Self-help treatment enabled regained sense of self-control (Rogers et al., 2004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doubts of legitimacy of seeking help for depression (Burr &amp; Chapman, 2004; Maxwell, 2005)</td>
<td>Stigma associated with treatment: depression and receiving treatment for depression is stigmatised, which is related to a moral discourse about personal responsibility and the fear of loss of social function in everyday life</td>
<td>Managing identity and stigma: guided self-help interventions must be perceived as an acceptable way to deal with depression. This may be heightened by focusing on the degree to which the treatment can assist in regaining everyday function, and by highlighting the importance of the patient as the agent of change</td>
</tr>
<tr>
<td>Moral dilemma caused by need to access treatment to limit impact of depression on others, but also negative beliefs about use of medication (Maxwell, 2005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patients exhibit low self-esteem and low expectations of care (Gask et al., 2003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frustration and distress caused by failure to deal with the problem, led to feelings such as weakness (Knudsen et al., 2002)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>People generally not sympathetic (Kadam et al., 2001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Threat of depression seen as continuing presence — used enactive responses and modified CBT to use independently (Glasman et al., 2004)</td>
<td>Patients’ understandings of self-help interventions: acceptance of self-help depends on prior experience of services Awareness of the concept of self as the mechanism of change</td>
<td>Individual as change agent: professionals and their actions are ascribed greater authority and power than patients in bringing about therapeutic change. Dealing with these perceptions may be critical in any introduction to guided self-help</td>
</tr>
<tr>
<td>Medico-centred information, needs to be patient-tailored, based on individual need and experience, to help patients to self-manage (Grime &amp; Pollock, 2004)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CBT, cognitive–behavioural therapy; GP, general practitioner.
Seeking out treatment was associated primarily with the perceived failure of these strategies rather than with the negative feelings or symptoms more usually associated with a diagnosis of depression (Rogers et al., 2001, 2004). A key theme related to patients’ feelings that they had to give up personal control over coping in order to engage with treatments based on biomedical principles such as antidepressant pharmacotherapy, which led to distinctions between ‘feeling better’ (due to the benefits of antidepressants) and ‘being better’ (a state of improved emotional well-being in the absence of medication) (Grime & Pollock, 2004; Maxwell, 2005). Taking medication could lead to a tension between patients feeling a sense of relief because prescribed medication functioned as a prop to help them deal with difficulties in everyday life, and the need to reject such solutions as a means of taking back personal control and recovering a sense of self and social functioning (Knudsen et al., 2002; Rogers et al., 2004; Maxwell, 2005).

**Stigma associated with treatment**

Stigma in the context of mental health problems refers to an array of social processes focused on the personal and interpersonal aspects of creating a ‘spoiled’ identity (Rogers & Pilgrim, 2005). Although it has been suggested that appeals to stigma are inadequate in explaining a reluctance to disclose emotional problems to health professionals (Prior et al., 2006), a salient theme emerging from the synthesis was the felt stigma associated with engaging with primary care. Accessing treatment for depression was not straightforward. This was partly related to feelings of a loss of control and a lack of legitimacy in accessing care for a non-physical problem (Gask et al., 2003; Burr & Chapman, 2004), and partly because conventional treatment for depression (i.e. antidepressant medication) was associated with potential threats to the sense of self (Knudsen et al., 2002; Grime & Pollock, 2004). In general most participants were keen to portray themselves as the type of people who do not resort to medication use, or would rather not need to resort to medication use if they could really help it (Maxwell, 2005). Taking medication was related to a moral discourse about personal responsibility, the fear of a loss of function in everyday life and a need to accept help for the sake of others (Rogers et al., 2001; Knudsen et al., 2002; Grime & Pollock, 2004). It was only when the general practitioner or others (family or friends) offered advice to alleviate this moral dilemma were they willing to accept medication use, and even then this acceptance was contingent on the intervention being seen as short-term and temporary (Maxwell, 2005).

Respondents were unsure what to tell others about being prescribed an antidepressant and were wary of telling people that they were taking such drugs, because of the combined stigma associated with depression and the taking of antidepressants (Knudsen et al., 2002; Grime & Pollock, 2004). The importance of change to personal identity was also raised in the studies we reviewed. Medication users reported that they felt they had become a person who needed to take antidepressants in order to get through daily life and were therefore somehow deficient. Respondents spoke of guilt and of letting themselves or others down, and expressed concerns about long-term changes to their personality associated with treatment (Grime & Pollock, 2004).

**Patients’ understanding of self-help interventions**

Patients’ understandings of self-help interventions depend on prior experience and an awareness of the concept of self as the mechanism of change. Such awareness takes time to develop, and is difficult in the context of some of the symptoms of depression such as low self-esteem and motivation (Glasmann et al., 2004; Rogers et al., 2004). The presence of a therapist offering guidance in the use of self-help materials generated ambivalence in patients about the relative role of the therapist vs. their own use of self-help materials (Rogers et al., 2004). There is an expectation that discussions about problems are therapeutic in their own right (Kadam et al., 2001) and the development of an effective therapeutic alliance did show an impact on whether patients would subsequently use self-help (Glasmann et al., 2004). Nevertheless, there was evidence of tension between the positive impact of the therapist and the negative effect on patient understanding of the therapist’s role within self-help. Following contact with the therapist, patients did not always follow the principles and exercises as prescribed but reconstructed the principles of cognitive–behavioural therapy in eclectic ways that had meaning and applicability to living with psychological problems on an everyday basis. Self-help activity was described as ‘hard work’, and participants reported that there were times when they faced crises or lapses in their ability to use the techniques (Glasmann et al., 2004).

**Application of the explanatory framework to the guided self-help intervention**

**Incorporating personal experience in a structured intervention**

People acquire an expert body of knowledge about health which includes theories about ways of managing and predicting outcomes in physical and mental health (Davison et al., 1991; Rogers & Pilgrim, 2005), and which complements professional knowledge. This lay knowledge is concerned with people’s experience of dealing with a mental health problem and its effects on social functioning, and refers to life events, their present social and psychological circumstances and their past history. In contrast, guided self-help is designed to provide patients with standardised cognitive–behavioural techniques that are known to be effective in the management of depression, which are in turn based on a psychological model of the cause of depression. Patients’ descriptions of the cause of their problems differed from the psychological model, which underlies cognitive–behavioural therapy or the more biomedical notion underpinning the prescribing of antidepressants.

The metaphors used by patients in relation to depression convey a sense of struggle with thoughts and emotions, and issues of control are a dominant feature of patients’ subjective experience of depression with the need to restore social functioning being prioritised over symptoms. Self-help materials, and the guidance that supports them, could use similar language and metaphors to enhance communication between patients and professionals and maximise the resources patients already bring with them. However, working effectively with people’s own definitions not only involves the use of language within guided self-help materials, it also requires explicit acknowledgement of the wide range of causal factors and pathways which can account for the problems that are seen by patients as constituting the phenomena of ‘depression’, and active engagement with

209
the broad range of explanations that lay people provide.

**Conceptualising the therapeutic environment for the purposes of engaging patients**

Help-seeking from primary care is typified by people waiting for long periods in a state of distress, while trying out personal coping strategies and seeking other sources of help. Despite a sense of urgency about seeking assistance when personal coping strategies fail, there is ambiguity about the role of primary care and ambivalence about its benefits. The cultural norms operating in primary care about help-seeking for a mental health problem are salient here. The synthesis pointed to the way in which people conceptualised physical problems as the ‘correct’ problems to be presented and managed appropriately within primary care, and this may be one of the reasons for a high rate of referral of so-called ‘medically unexplained symptoms’ to primary care. Although on the face of things the findings concerning the management of depression within primary care may seem tangential to the main research question concerning the optimal way of implementing guided self-help, the context of primary care and how it is perceived by patients and professionals as an arena for the disclosure and subsequent management of mental health problems is not. Receptivity, and the norms and values operating in primary care about mental health, are likely to be central in considering issues concerning the delivery of guided self-help and the promotion of primary care settings as an appropriate place from which people can seek support. Issues such as the point in the illness trajectory at which people make contact with services, and their prior contact with other sources of help, may be important to review in assessing the acceptability and appropriateness of guided self-help.

**Everyday self-management strategies and guided self-help**

The articles suggested that patients developed individual strategies for controlling feelings – although these strategies varied significantly. As noted above, guided self-help is designed to provide specific strategies based largely on cognitive–behavioural therapy. However, it may not be optimal simply to replace everyday strategies with those conventionally considered to be ‘evidence-based’. Instead, acknowledging the importance of those everyday strategies and attempting to build on them as part of a negotiated introduction to guided self-help might encourage patient acceptance and involvement, if people were able to see the treatment as a progression from the activities that they themselves had initiated and found useful. For example, a major aspect of guided self-help is ‘homework’, where patients put aside time to use cognitive–behavioural techniques. Homework might be facilitated by linking to patients’ existing use of specific physical locations, which provide a sense of control and safety. In this way, individual therapeutic techniques could be delivered in a way that is both ‘evidence-based’ and ‘patient-centred’.

**Managing identity and stigma**

The way in which depressive problems and their management are perceived by others and affect identity emerged as prominent themes, and the extent to which the guided self-help philosophy avoids or acknowledges issues of stigma associated with medication, are likely to be key predictors of acceptability and appropriateness. If guided self-help can be discussed with individuals as a management strategy requiring a sense of acting on the world and enhancing self-worth, this might support efforts to engage people with depression by providing a means of management which is more acceptable to them and to others around them, and which allows them to avoid moral dilemmas concerning use of treatments. It may be that encouraging individuals to make a direct comparison of the benefits and drawbacks of guided self-help in contrast to those of antidepressant therapy is a means to achieve this.

**Individual as change agent**

Professional actions are ascribed greater authority and power than patients in bringing about therapeutic change. Generally an expected ‘cultural gap’ (Horowitz, 1983) exists between service users, who view their role as being in receipt of treatment rather than initiating therapy, and professionals, who are imbued with esoteric knowledge and charged with ministering to ‘patients’. This representation of the facilitator in a guided self-help model as the change agent was evident in one study (Rogers et al., 2004), but other papers reported that people spoke of having the strength to overcome negative feelings and felt that depression was something that they could control. This relates to the awareness of the self as the mechanism of change, which is the key to engagement with guided self-help. However, this perception was not widespread, and this may reflect a general view that treatments provided in a medical context do not require the patient to take a highly active role. There was evidence that developing the idea that the individual is the principal ‘active ingredient’ in guided self-help takes time, and there is a tension between the need to develop this idea and the relatively short-term nature of contact within guided self-help. It is possible that information provided before treatment begins could overcome some of the misconceptions patients might have about the nature of treatment, which could be reinforced further by contact with the therapist.

**DISCUSSION**

The synthesis is dependent on the particular studies included. The search was restricted to a relatively short period and might have excluded relevant studies, so the results cannot be considered definitive. The development of aspects of the explanatory framework often involved synthesising findings from only two or three studies rather than the majority; this partly reflects the fact that the research identified was fairly heterogeneous in nature. A more homogeneous set of studies might have allowed the development of a smaller number of themes, which could have been more fully described in the synthesis. The process of qualitative synthesis cannot be reduced to a set of mechanistic tasks, which raises issues of the transparency of the process. It is possible that other researchers would have synthesised the material differently. Some of the papers included in the meta-synthesis were published by the present authors, and the results of this synthesis may thus be weighted towards issues identified by us in previous work.

The meta-synthesis highlighted a number of key issues that may affect the success of the introduction of guided self-help in primary care. These include the importance of issues of control and social functioning among patients with depression; the need to ensure that the context of primary care is viewed as a suitable location for mental healthcare, and supports the active role of the patient that is required in guided
self-help; and the importance of engaging actively with patients’ own constructions of depression and their current coping strategies. An interesting issue concerns potential tensions between the results of the synthesis and current professional perspectives on mental health issues. For example, cognitive–behavioural approaches to symptoms (such as exposure and behavioural activation) are designed to combat avoidance behaviour, which may clash with patient behaviours that provide a sense of control and safety.

Although the synthesis made a distinction between antidepressant therapy and guided self-help in terms of issues of control, no paper discussed the experience of patients who were using both treatments; research into the views of such patients might usefully extend understanding of issues of control in engagement with treatments, and enable services to be provided that maximise the advantages of both types of treatment.

Some of the findings of the meta-synthesis may have implications beyond the design of the guided self-help intervention. For example, the traditional biomedical view of the importance of managing depressive symptoms is challenged, as the meta-synthesis highlighted the importance of issues of control and social functioning in the decision to seek help, in the evaluation of treatments, and enable services to be provided that maximise the advantages of both types of treatment.

Some of the findings of the meta-synthesis may have implications beyond the design of the guided self-help intervention. For example, the traditional biomedical view of the importance of managing depressive symptoms is challenged, as the meta-synthesis highlighted the importance of issues of control and social functioning in the decision to seek help, in the evaluation of treatments, and enable services to be provided that maximise the advantages of both types of treatment.

The study was conducted in the context of a trial funded by the Medical Research Council (principal investigator Professor Karina Lovell), in which P.B. and A.R. are grant holders. A.R. and P.B. are funded through the Department of Health, but the views expressed in this paper are those of the authors alone. N.K. is funded by the Medical Research Council. The authors would like to thank Rosalind McNally for assistance with the search strategies, and the referees for their useful comments on the initial draft.

REFERENCES


Medical Research Council (2000) A Framework for Development and Evaluation of RCTs for Complex Interventions to Improve Health, MRC.


White matter changes and late-life depressive symptoms

Longitudinal study


Background Evidence from cross-sectional studies suggests a link between cerebral age-related white matter changes and depressive symptoms in older people, although the temporal association remains unclear.

Aims To investigate age-related white matter changes on magnetic resonance imaging (MRI) as an independent predictor of depressive symptoms at 1 year after controlling for known confounders.

Method In a pan-European multicentre study of 639 older adults without significant disability, MRI white matter changes and demographic and clinical variables, including cognitive scores, quality of life, disability and depressive symptoms, were assessed at baseline. Clinical assessments were repeated at 1 year.

Results Using logistic regression analysis, severity of white matter changes was shown to independently and significantly predict depressive symptoms at 1 year after controlling for baseline depressive symptoms, quality of life and worsening disability (P < 0.01).

Conclusions White matter changes pre-date and are associated with the development of depressive symptoms. This has implications for treatment and prevention of depression in later life.

Declaration of interest None. Funding detailed in Acknowledgements.

Cerebral white matter changes are strongly associated with depressive symptoms in older people (de Groot et al, 2000). However, the direction of causation remains unclear because most studies are cross-sectional (O’Brien et al, 1996; Hickie et al, 1997; Steffens et al, 1999) and the possibility remains that both white matter changes and depressive symptoms are manifestations of a common pathological pathway, or that depressive symptoms pre-date the development of white matter changes.

The LADIS (Leukoaraiosis and Disability in the Elderly) Study is a large multi-centre pan-European longitudinal study of older adults without significant disability which was established to investigate the relationship between white matter changes on magnetic resonance imaging (MRI) and subsequent development of disability, cognitive impairment and depressive symptoms. We have previously reported cross-sectional associations between white matter changes and depressive symptoms in this sample (Firbank et al, 2005; O’Brien et al, 2006). In this report we examine white matter changes as an independent predictor of future depressive symptoms. We hypothesised that the severity of white matter changes at baseline would be associated with the development of depressive symptoms at 1 year follow-up, irrespective of known confounding variables.

METHOD

Recruitment

Participants were recruited from the 11 European centres participating in the LADIS Study (Amsterdam, Copenhagen, Florence, Graz, Gothenburg, Helsinki, Huddinge, Lisbon, Mannheim, Newcastle upon Tyne and Paris). Full details of the study design have been reported previously (Pantoni et al, 2005).

In brief, inclusion criteria were: (a) between 65 and 84 years; (b) living in the community; (c) no or mild disability (only one item compromised) as assessed by the Instrumental Activities of Daily Living scale (IADL; Lawton & Brody, 1969); (d) the presence of an informant; (e) any degree of age-related white matter changes on MRI scan according to a revised version of the scale of Fazekas et al (1987).

Exclusion criteria were: (a) the presence of severe illnesses (cardiac, hepatic or renal failure, neoplastic or other relevant systemic disease) which would increase the likelihood of drop-out; (b) severe unrelated neurological diseases; (c) leucoencephalopathies revealed by brain imaging that were of non-vascular origin (immunological, demyelinating, metabolic, toxic or infectious); (e) severe psychiatric disorders; (f) refusal or inability to give informed consent; (g) refusal or inability to undergo cranial MRI scanning.

There were 639 people who fulfilled the criteria and were recruited between July 2001 and January 2003. Most were recruited after presentation to centres with mild cognitive disturbances (n = 168), gait disturbances (n = 28), psychiatric complaints (n = 13), other neurological disturbances (n = 129), or minor stroke (n = 122). Other participants included those in whom white matter changes were incidentally found on computed tomography or MRI performed in other clinical settings (n = 107) and controls from other studies with brain white matter changes (n = 72). The total number of participants referred from each centre and the reasons for referral are given in Table DS1 of the data supplement to the online version of this paper. All participants gave informed consent.

Assessment

All participants had a comprehensive baseline demographic assessment by trained personnel. Information was collected on age, gender, education, occupational status, living conditions, previous medical conditions, including stroke (as defined according to the World Health Organization; Hatano, 1976) and hypertension (Chalmers et al, 1999), prescribed medication, lifestyle (alcohol and smoking) and vascular risk factors.

Further baseline clinical assessments included the following.

(a) Mini-Mental State Examination (MMSE) was used to assess cognition (Folstein et al, 1975).
(b) Functional status in terms of disability was measured by the IADL. This is a self-reported scale developed to monitor function and independent living among older adults and measures a broad set of daily activities, including shopping for personal items, preparing meals, performing housework and managing personal finances.

(c) Health-related quality of life (QoL) was measured using the self-reported Euro-QoL 5D (Euro-QoL Group, 1990). The participant indicates their current health state where 100 indicates best imaginable health and 0 the worst.

(d) Depressive symptoms were assessed by the self-completed 15-item Geriatric Depression Scale (GDS), which was developed as a screening instrument for depressive symptoms in elderly populations (Yesavage, 1988). The maximum possible score is 15.

(e) A history of depression was recorded, along with the date of any incident depression. History of depression was defined as the presence of a depressive episode requiring treatment or hospital admission. Incident depression was defined as any depressive episode requiring treatment or hospital admission over the study year. Both history and incident depression were obtained through interview of the participant and evaluation of the case notes.

(f) Standard neurological and cardiovascular examination was performed and included blood pressure measurements.

Participants were re-evaluated at 1 year and clinical assessments were repeated. To increase reliability, investigators were issued with a specifically designed handbook which contained guidelines for applying tools.

Magnetic resonance imaging
Participants had a baseline MRI scan at their respective centres. All centres used MRI scanners with a field strength of 1.5 T, except for one which had a 0.5 T system. A standard protocol was used (Pantoni et al, 2005). For the white matter rating, a FLAIR sequence was acquired with the following parameters: 250 mm field of vision; 256 x 256 or 256 x 192 matrix; 5 mm slice thickness; 0.5 mm slice gap; 19–28 slices; time to echo 100–140 ms; time to repetition 6000–10 000 ms; inversion time 2000–2500 ms, and echoes per shot 7–24.

Volumetric analysis of age-related white matter changes was performed on a Sparc 5 workstation (Sun, Palo Alto, California, USA; van Straaten et al, 2006). Lesions were marked using a ‘seed’ and local thresholding was performed using home-developed software (Show Images, version 3.6.1 using a Canay filter) on each slice. When necessary, borders could be adjusted by the operator by changing thresholds for upper and lower intensity values. If all lesions are delineated, the program calculates the total surface of the outlined area. By multiplying with the interslice distance, total volume of age-related white matter changes is established (Gouw et al, 2006).

Statistical analysis
Data were collected in each centre and entered into a central electronic database on a specifically developed website (http://www.unifi.it/LADIS). In a community-dwelling population it is normal for depression rating scales to be heavily skewed towards low values. Hence for analysis we divided our baseline and 1-year data into quintiles, using the same GDS range for each quintile as in our previous study (O’Brien et al, 2006). For the volume of age-related white matter changes we used a logarithmic transformation to produce normally distributed data.

Univariate analysis was used to examine correlations between different variables and depressive symptoms at 1 year. As depressive symptoms (as measured by GDS) are on an ordinal scale, we used Spearman’s rank order correlation coefficient (r_s) to determine the correlation between depressive symptoms at 1 year, the dependent variable and various independent demographic, clinical and MRI variables. Baseline independent variables included age, gender, baseline depressive symptoms (GDS), educational level, smoking status, MMSE score, stroke, hypertension, QoL and log total volume of white matter changes. Incident stroke (a new stroke over the study year) and worsening disability (an IADL score at 1 year less than at baseline) were included as further independent variables.

Multivariate analysis used ordinal logistic regression to determine predictors of the quintile of depression scale score at 1 year. The variables chosen were those which were significant in the univariate analysis and included GDS baseline quintile, worsening disability, QoL, MMSE score, years of education, incident stroke and total volume of white matter changes.

We constructed a further binary logistic regression to compare predictors of a depressive episode over the year using history of depression (instead of GDS quintile) together with the other significant independent variables as predictors.

Values of P < 0.05 were considered statistically significant.

RESULTS

Participants
Of the original cohort of 639 participants there were missing baseline data for MMSE score in 1, presence of stroke in 1, years of education in 1, hypertension in 2, total volume of white matter changes in 21 and QoL rating in 6. From the remaining 607, 6 people died during the year and 12 dropped out, leaving 589 participants at 1-year follow-up. There were missing follow-up data for a further 63 participants. Of these some had multiple variables missing; GDS score in 54; IADL score in 53; QoL score in 62; MMSE score in 57; and history of depression in 50. Thus there were 526 participants in the final analysis. To determine if there had been any effect of selective drop-out, we compared the baseline GDS, total volume of white matter changes, QoL and MMSE scores between those who did and did not have follow-up data. The baseline MMSE score was significantly lower in those without (mean 26.2, s.d.=2.5) compared with those with follow-up data (mean 27.6, s.d.=2.3). There were no significant differences in the other examined variables.

As expected for a community population, the group GDS score remained stable over time. The median GDS score was 2 at baseline and at 1 year (Wilcoxon signed-rank Z = -1.77, P=0.077). At baseline, 141 people had a history of depression for which medical help had been sought. Over the year, 85 people had an episode of depression. The number of people in each quintile of GDS at baseline and 1 year is given in Table DS2 of the data supplement to the online version of this paper.

Table 1 shows the demographic, clinical and MRI characteristics of the participants at baseline. Over the year, 14 participants had an incident stroke and 56 had worsening disability.
The principal finding of this longitudinal study was that severity of baseline white matter changes and GDS score at baseline were significantly and independently associated with GDS quintile at 1 year.

To control for the effect of MRI centre as a potential confounder we constructed a further logistic regression model including MRI centre as an additional independent variable, and found that baseline log total volume of white matter changes still significantly and independently predicted GDS quintile at 1 year ($P=0.046$).

Table 4 shows a further binary regression with occurrence of depression over the year as the dependent variable and baseline history of depression, QoL score, worsening IADL, MMSE score, years of education, incident stroke and log total volume of white matter changes as independent variables. Baseline history of depression, QoL and worsening IADL were significantly and independently associated with depression at 1 year. The association for log total volume of white matter changes did not quite reach statistical significance ($OR=1.63, P=0.1$).

**DISCUSSION**

**Main findings**

The principal finding of this longitudinal study was that severity of baseline white matter changes and GDS score at baseline were significantly and independently associated with GDS quintile at 1 year.

To control for the effect of MRI centre as a potential confounder we constructed a further logistic regression model including MRI centre as an additional independent variable, and found that baseline log total volume of white matter changes still significantly and independently predicted GDS quintile at 1 year ($P=0.046$).

Table 4 shows a further binary regression with occurrence of depression over the year as the dependent variable and baseline history of depression, QoL score, worsening IADL, MMSE score, years of education, incident stroke and log total volume of white matter changes as independent variables. Baseline history of depression, QoL and worsening IADL were significantly and independently associated with depression at 1 year. The association for log total volume of white matter changes did not quite reach statistical significance ($OR=1.63, P=0.1$).
matter changes predicted subsequent depressive symptoms at 1 year, even after controlling for baseline depressive scores, QoL, worsening disability, incident stroke, educational level and MMSE score.

Baseline white matter changes did not, however, significantly predict depressive episodes over the following year. The fact that we were unable to demonstrate that age-related change in white matter volume was a significant independent predictor of this clinically important measure might have been because of a lack of power in this study to detect such an effect. Incident depression may have been underreported because of the method used to screen for mood disorders. This is supported by the relatively low number of participants with incident depression (85) over the year and the fact that there was a trend towards significance (P=0.1).

Finally we established that QoL, worsening IADL and baseline GDS were significant and independent predictors of depressive symptoms, whereas incident stroke, MMSE score or years of education were not.

**Previous studies**

Our results support the findings of cross-sectional studies that white matter changes are of aetiological significance for depressive symptoms in older adults. Already a strong association between white matter changes and depressive symptoms has been established in both hospital patients (O’Brien et al., 1996) and those living in the community (Steffens et al., 1999; de Groot et al., 2000). White matter changes are also known to predict poorer outcome (O’Brien et al., 1998) and treatment response (Hickie et al., 1997) in elderly people with depression. Neuropathological study has shown deep white matter changes to have an ischaemic basis (Thomas et al., 2002) and it is thought that these changes represent a marker for vascular pathology in clinically important areas of the brain.

Following on from these findings the vascular depression hypothesis has been proposed (Baldwin & O’Brien, 2002; Alexopoulos, 2005), which states that disruption of fronto-striatal circuits (which reciprocally link prefrontal cortex to basal ganglia) by vascular changes predisposes, perpetuates or exacerbates depressive syndromes. Our study suggests that white matter changes are associated with the development of depressive symptoms in older adults and hence supports the vascular depression hypothesis.

Despite the convergence of evidence from cross-sectional studies, results from longitudinal studies remain unclear. In a large community population, the Cardiovascular Health Study (Steffens et al., 2002) found that small basal ganglia lesions and large cerebral cortical white matter changes at baseline predicted persistence of depressive symptoms; furthermore, subcortical white matter changes predicted worsening of depressive symptoms over time. Our data are in line with these results. However they are at variance with the findings of the study from the National Institute of Mental Health (NIMH; Taylor et al., 2003). In a cohort of 133 participants with severe depression, an increase in severity of white matter changes over 2 years, but not static baseline white matter change, predicted depressive outcome scores after controlling for baseline depressive symptoms (Taylor et al., 2003). We did not assess lesion progression; however, potential reasons for these discrepancies may be the different and highly selected population used in the smaller NIMH cohort.

Another longitudinal study, the PROSPER study (Versluis et al., 2006), rather surprisingly found that white matter changes were not related to baseline depressive features or development of depressive symptoms at follow-up. Furthermore, no association was found between progression of white matter changes and the progression of depressive symptoms. This is clearly in contrast to our results and those of other studies. Discrepancies with the current evidence may have been because of a relatively low total volume of white matter changes of participants in the PROSPER study (median 1.7 v. 13 ml in this study) and the low rate of participation of people with depressive symptoms (8% with GDS score > 3 v. 35% in this study).

**Nature of the relationship**

Our longitudinal results provide an important demonstration that white matter changes pre-date, and therefore may be causally related to, the development of depressive symptoms. We cannot, however, dismiss the possibility that depressive symptoms may also pre-date white matter changes. The relationship may be bidirectional, just as it is with vascular risk factors and depression in older adults (Thomas et al., 2004; Baldwin, 2005). It is known that baseline vascular burden independently predicts depression (Mast, 2004) and yet it is also clear from longitudinal studies that depression is an independent risk factor for later developing coronary artery disease (Pratt et al., 1996), stroke (Evenson et al., 1998) and cardiac mortality post-myocardial infarction (Frasure-Smith et al., 1999).

The relationship is likely to be complex, although we have helped to unravel one aspect. Interestingly, the temporal association demonstrated remained significant even after controlling for baseline cognitive scores and worsening disability. This suggests that this effect is not simply mediated through disability or secondary to a psychological reaction to declining cognitive performance, as has also been suggested (Cahn et al., 1996; Paterniti et al., 2002).

**Strengths and limitations of the study**

The strengths of the LADIS Study are the large number of community-based participants without significant disability and its multicentre design. Further strengths include the central measurement of all scans by a single operator and the use of a robust measure of the volume of age-related white matter changes which is less operator dependent and less susceptible to ceiling effects (van Straaten et al., 2006). Finally, a strong model for logistic analysis was constructed controlling for many potential confounders, including putative mediators of depressive features such as cognition, QoL, incident stroke and worsening disability.

Limitations include the recruitment of only those with white matter changes: consequently people with minimal or no health problems might have been excluded (although evidence suggests that up to 95% of older adults have some white matter changes; de Leeuw et al., 2001). Another potential limitation arises from the fact that some symptoms measured by the GDS, such as apathy and withdrawal, are not exclusive to a depressive syndrome and may also be manifestations of a mild cognitive dysexecutive syndrome. Despite this weakness we were able to demonstrate a significant association between white matter changes and depressive symptoms not only after controlling for cognitive confounders but also at higher GDS scores, when the symptoms are more likely to
reflect a depressive syndrome. Finally, our results might have been influenced by sampling bias as the study population represents a heterogeneous group with a variety of different complaints but without significant disability. However, such bias is unavoidable in large multicentre studies and this heterogeneity might actually increase the generalisability of our findings.

Clinical implications
The most important implications concern the prevention of late-life depressive symptoms. These extend more widely to the elderly population as a whole because most participants had less disability and less depression than seen routinely in clinical practice. Recent evidence shows that treatment of hypertension slows progression of white matter changes (Schiffrin, 2005). We would therefore advocate tighter control of vascular risk factors to prevent the development of late-life depressive symptoms.

The presence of white matter changes on MRI, if considered alone, is of little clinical value in predicting future depressive symptoms. However, when white matter changes are taken along with other independent predictors such as quality of life and previous depressive episodes, they raise the index of suspicion for further depressive symptoms. Thus in clinical practice it is wise to take account of the presence of white matter changes on MRI when attempting to predict future depressive symptoms, and this might influence decisions regarding the frequency of clinical monitoring and the need for prophylactic antidepressants.

Future research
Large studies in a community population are indicated to establish whether static lesion volume, rate of progression, or both are the important determinants in predicting future depressive symptoms. Further research should also be directed at looking more precisely at the time course of emergence of changes and of mood symptoms, and whether modification of white matter changes influences the course of depressive symptoms.

ACKNOWLEDGEMENTS
The LADIS Study is supported by the European Union within the V European Framework Programme ‘Quality of Life and Management of Living Resources’ (1998–2002), contract no. QLRT-2000-00446 as a concerted action.

APPENDIX
List of participating centres and personnel
Helsinki, Finland (Memory Research Unit, Department of Clinical Neurosciences, Helsinki University), Timo Erkintjunti, MD, PhD, Tarja Pohjasaava, MD, PhD, Pia Pihanen, MD, Raja Ylikoski, PhD, Hanna Jokinen, LPsych, Meija-Marjat Somerokoski, MLPsych, Riitta Mantyla, MD, PhD, Olli Salonen, MD, PhD; Graz, Austria (Department of Neurology and MRI Institute, Medical University Graz), Franz Fazekas, MD, Reinhold Schmidt, MD, Stefan Ropele, PhD, Brigitte Reus, MD, Kaja Petrovic, MagPsych, Ulfke Garmeh, Alexandra Seewann, MD; Lisbon, Portugal (Serviço de Neurologia, Centro de Estudos Egas Moniz, Hospital de Santa Maria), Jose M. Ferro, MD, PhD, Ana Verdelho, MD, Sofia Madureira, PsyD; Amsterdam, The Netherlands (Department of Radiology and Neurology, VU Medical Center), Philip Scheltens, MD, PhD, Ilse van Straaten, MD, Frederik Barkhof, MD, PhD, Alida Gouw, MD, Wiesje van der Flier, MD; Gothenburg, Sweden (Institute of Clinical Neuroscience, Gothenburg University), Anders Wallin, MD, PhD, Michael Jonsson, MD, Karin Lind, MD, Arto Nordlund, PsyD, Sindre Rolstad, PsyD, Ingela Isblad, RN; Huddinge, Sweden (Karolinska Institute, Neuros Department, Section of Clinical Geriatrics), Lars-Olof Wahlund, MD, PhD, Milta Crisby, MD, Anna Pettersson, Physiotherapist, Kaarina Amberla, PsyD; Paris, France (Department of Neurology, Hopital Larbousu), Huges Chabrait, MD, PhD, Karen Hernandez, Psychologist, Annie Kurzt, Psychologist, Dominique Herve, MD, Mannheim, Germany (Department of Neurology, University of Heidelberg, Klinikum Mannheim), Michael Hennerici, MD, Christian Blahak, MD, Hansjorg Baazner, MD, Martin Wiarda, PsyD, Suzanne Seip, RN, Copenhagen, Denmark (Memory Disorders Research Group, Department of Neurology, Rigshospitalet, Copenhagen University Hospital, Denmark), Gunhild Waldemar, MD, DMSc, Egil Rostrup, MD, MSc, Charlotte Ryberg, MSc, Tim Dyrbj, MSc, Olaf B. Paulson, MD, DPSc; Newcastle upon Tyne, UK (Institute for Ageing and Health, University of Newcastle), John O'Brien, DM, Sanjeev Pakrasi, MRCPsych, Mani Krishnan MRCPsych, Andrew Teodorczuk, MRCPsych, Michael Birbank, PhD, Philip English, DCRC, The Coordinating Centre is in Florence, Italy (Department of Neurological and Psychiatric Sciences, University of Florence), Domenico Inzitari, MD (Study Coordinator), Luciano Bartolini, PhD, Anna Maria Basile, MD, PhD, Elena Magnani, MD, Monica Martini, MD, Mario Mascali, MD, PhD, Marco Moretti, MD, Leonardo Pantoni, MD, PhD, Anna Poggesi, MD, Giovanni Pracucci, MD, Emilia Salvadori, PhD, Michela Simon, MD, The LADIS Steering Committee comprises Domenico Inzitari, MD (Study Coordinator), Timo Erkintjunti, MD, PhD, Philip Scheltens, MD, PhD, Marieke Visser, MD, PhD, Peter Langhorne, MD, BSc, PhD, FRCP who replaced Kjell Asplund, MD, PhD in 2005.

REFERENCES
Correspondence: Dr Teodorczuk, Institute for Ageing and Health, Newcastle General Hospital, Westgate Road, Newcastle upon Tyne NE4 6BE, UK. Email: andrew.teodorczuk@ncl.ac.uk

(First received 9 February 2007, final revision 26 April 2007, accepted 27 June 2007)

ANDREW TEODORCZUK, MRCPsych, JOHN T. O'BRIEN, DM, MICHAEL J. FIRBANK, PhD, Institute for Ageing and Health, University of Newcastle upon Tyne, UK; LEONARDO PANTONI, MD, ANNA POGGESI, MD, Department of Neurological and Psychiatric Sciences, University of Florence, Italy; TIMO ERKINJUNTI, MD, Memory Research Unit, Department of Neurology, University of Helsinki, Finland; ANDERS WALLIN, MD, Institute of Clinical Neuroscience, Gothenburg University, Sweden; LARS OLOF WAHLUND, MD, Karolinska University Hospital, Huddinge, Sweden; ALIDA GOUV; MD, Alzheimer Centre and Department of Neurology, VU University Hospital, Amsterdam, The Netherlands; GUN-HILD WALDEMAR, MD, Memory Disorders Research Group, Department of Neurology, Rigshospital, Copenhagen University Hospital, Denmark; REINHOLD SCHMIDT, MD, Department of Neurology and MRI Institute, Medical University Graz, Austria; JOSÉ M. FERRO, MD, Servicio de Neurología, Centro de Estudios Egas Moniz, Hospital de Santa Maria, Lisbon, Portugal; HUGUES CHABRIAT, MD, Department of Neurology, Hopital Lariboisiere, Paris, France; HANSJÖRG BÄZNER, MD, Department of Neurology, University of Heidelberg, Universitätshklinikum Mannheim, Germany; DOMENICO INITITARI, MD, Department of Neurological and Psychiatric Sciences, University of Florence, Italy


Genetic risk of depression and stress-induced negative affect in daily life

MARIEKE WICHERS, INEZ MYN-GERMEYS, NELE JACOBS, FRENK PEETERS, GUNTER KENIS, CATHERINE DEROM, ROBERT VLIETINCK, PHILIPPE DELESPAUL and JIM VAN OS

Background A bias to develop negative affect in response to daily life stressors may be an important depression endophenotype, but remains difficult to assess.

Aims To assess this mood bias endophenotype, uncontaminated by current mood, in the course of daily life.

Method The experience sampling method was used to collect multiple appraisals of daily life event-related stress and negative affect in 279 female twin pairs. Cross-twin, cross-trait associations between daily life mood bias and DSM-IV depression were conducted.

Results Probands whose co-twins were diagnosed with lifetime depression showed a stronger mood bias to stress than those with co-twins without such a diagnosis, independent of probands’ current depressive symptoms and to a greater extent in monozygotic twins than in dizygotic twins.

Conclusions Genetic liability to depression is in part expressed as the tendency to display negative affect in response to minor stressors in daily life. This trait may represent a true depression endophenotype.

Declaration of interest None.

The personality trait neuroticism is related to the tendency to develop negative affect in the face of stress (Bolger & Schilling, 1991; Gunthert et al., 1999; Van Os & Jones, 1999; Kendler et al., 2004) and has been referred to as a ‘mood bias towards negative affect’ (Hasler et al., 2004; Erickson et al., 2005). Although research suggests that this is a depression endophenotype (Fanous et al., 2002; Gottesman & Gould, 2003; Hettema et al., 2006), two major problems remain. First, neuroticism is invariably measured using a brief, cross-sectional self-report rating scale with limited ecological validity to the dynamic concept of negative mood bias in the face of daily life stress. Second, many of the items in rating scales of neuroticism are mood-related and therefore contaminated by current mood states, in particular depression (Horwood & Fergusson, 1986). To examine the endophenotype hypothesis, measures are needed that directly capture reactivity of mood states in response to daily life stressors. In this study experience sampling methodology (Csikszentmihalyi & Larson, 1987) was combined with the twin method in order to establish uncontaminated cross-twin, cross-trait associations between daily life mood bias on the one hand and depressive disorder on the other.

METHOD

The study sample was drawn from the general population of Flanders (Belgium) and consisted of 279 female twin pairs aged 18–46 years. Most (218) of these pairs were recruited from the East Flanders Prospective Twin Survey; this population-based survey has prospectively recorded all multiple births in the province of East Flanders since 1964 (Loos et al., 1998; Derom et al., 2002). The project was approved by the local ethics committee and all participants gave written informed consent. The sample was female only, given evidence for qualitative differences in the type of environmental stressors that are associated with depression in men and women (Kendler et al., 2001), and evidence for gender-specific genetic factors for both neuroticism and depression (Fanous et al., 2002).

Experience sampling method

The experience sampling method (ESM) is a structured diary technique to assess people in their daily living environment and has been validated for the purpose of studying the immediate effects of stressors on mood (Csikszentmihalyi & Larson, 1987; DeVries, 1992; Delepaun, 1995). Participants received a digital wristwatch and a set of ESM self-assessment forms collated in a booklet for each day. The wristwatch was programmed to emit a signal (‘beep’) at an unpredictable moment in each of ten 90 min time blocks between 07.30 and 22.30 on five consecutive days. After each beep participants were required to stop their activity and to fill out the ESM self-assessment forms previously handed to them, collecting reports of thoughts, current context (activity, persons present and location), appraisals of current situation and mood. All self-assessments were rated on seven-point Likert scales. Trained research assistants with ample experience in momentary assessment technique explained the ESM procedure to the participants during an initial briefing session and a practice form was completed to confirm that the latter were able to understand the Likert scale. Participants were given a telephone number to call in case they had questions or problems during the ESM sampling period. They were instructed to complete their reports immediately after the beep, thus minimising memory distortion, and to record the time at which they completed the form. In order to know whether the participants had completed the form within 15 min of the beep, the time at which they indicated they completed the report was compared with the actual time of the beep. All reports not filled in within 15 min after the beep were excluded from the analysis, since previous work (Delepaun, 1995) has shown that reports completed after this interval are less reliable and consequently less valid. In addition, participants with fewer than 17 valid reports (out of 50) were excluded from the analysis, as previous work has shown that measures of individuals with
Measurements

The Structured Clinical Interview for DSM-IV Axis I disorders (SCID; First et al, 1996) was administered in order to obtain current and lifetime diagnoses of major depressive disorder. Participants also filled in the Symptom Check List – 90 – Revised (SCL–90–R; Derogatis, 1983) in order to obtain a continuous measure of depressive symptoms. The SCL–90–R depression score was log-transformed in order to improve normality.

Measures of stress and negative affect were collected at each beep within the experience sampling method framework. In order to measure ESM event-related stress (hereafter referred to simply as ‘stress’), participants were asked to report the most important event that happened between the current and the previous beep. This event was subsequently rated on a seven-point bipolar scale (−3 very unpleasant, 0 neutral, 3 very pleasant). Responses were recoded to allow high scores to reflect stress (−3 very pleasant, 0 neutral, 3 very unpleasant). Negative affect was assessed at each beep with six mood adjectives (‘insecure’, ‘lonely’, ‘anxious’, ‘low’, ‘guilty’ and ‘suspicious’) rated on seven-point Likert scales as described above. The mean of the six items was taken as the measure of negative affect (Cronbach’s \( \alpha = 0.76 \) over the participants’ mean).

Cross-twin, cross-trait method

The cross-twin, cross-trait method was used in this study. This means that within a twin pair, variable \( x \) in the proband is associated with variable \( y \) in the co-twin (Fig. 1). The reason why the second variable is measured in the other twin is because measure \( x \) may be confounded by measure \( y \) when measured in the same person. This method thus ensures that the association between \( x \) and \( y \) – or in this case ‘negative mood bias’ and ‘depression’ – is uncontaminated.

Analyses

Experience sampling method data have a hierarchical structure. In this study, multiple observations (level 1) were clustered within individuals (level 2), who were part of twin pairs (level 3). Multilevel analysis takes the variability associated with each level of nesting into account (Snijders & Bosker, 1999).

In order to investigate the association between stress and negative affect in the course of daily life (i.e. stress sensitivity), multilevel linear regression analyses, using the XTMIXED command in Stata version 9.1 for Windows, were applied to the data. A cross-twin, cross-trait design was used in which negative affect in the proband twin was regressed on the interaction ‘stress’ of the proband and lifetime DSM–IV diagnosis of major depressive disorder in the co-twin (hereafter referred to as co-twin lifetime depression), to test the hypothesis that familial vulnerability for depression would have an impact on stress sensitivity (i.e. would moderate the association between stress and negative affect). The analysis was additionally corrected for the SCL–90–R depression score and past DSM–IV depressive disorder in the proband twin, and a sensitivity analysis was carried out excluding all proband twins with a current depressive state.

Next, the degree of dose–response relationship in the association between stress and negative affect as a function of co-twin lifetime depression was investigated, with the hypothesis that higher appraisals of stress would display greater interaction effects. Thus, it was examined whether higher stress appraisal was associated with a greater impact of co-twin lifetime depression on stress reactivity, the effect of stress on negative affect. For this purpose, using the score ‘very pleasant’ as the reference category, six dummies were created, since the appraisal scores for the stress of the event ranged from −3 to 3 (‘very pleasant’ to ‘very unpleasant’). From the model with the interactions between the stress dummy variables and co-twin lifetime depression, effect sizes were calculated for stress with and without co-twin lifetime depression, stratified for each separate level of stress by applying and testing the appropriate linear combinations using the Stata LINCOM command. Main effects and interactions were assessed by Wald test (Clayton & Hill, 1993).

Finally, in order to examine possible contribution of genes to observed cross-twin, cross-trait associations, the three-way interaction Stress \( \times \) co-twin lifetime depression \( \times \) zygosity was fitted and evaluated, followed by calculation of stratified effect sizes using the LINCOM command.

RESULTS

The total sample consisted of 621 White female participants. Thirty-one women were excluded because they had fewer than 17 valid ESM self-reports or missing ESM self-reports. Forty-five women were non-twin sisters and therefore were not included in the analyses. Another 34 participants were excluded because of missing data. This resulted in a data-set of 511 participants who were part of 259 different twin pairs, of which 158 were monozygotic, 100 were dizygotic and 1 was of unknown zygosity. The mean age of the twins was 27 years (s.d. = 7.4, range 18–46). Nearly two-thirds (63%) had a college or university degree, 35% had completed secondary education and 2% had primary education only. The majority were currently employed (63% employed, 32% student, 2% unemployed, 2% homemaker and 0.4% sick leave).

Depression vulnerability and negative affect reactivity to daily life stress

The total number of ESM observations was 14,323 and the mean number of observations within individuals was 28. Eighty-nine probands (17.4%) had a co-twin with a diagnosis of lifetime depression. The mean stress score was −1.00 (s.d. = 0.62) for participants with co-twin lifetime depression and −1.16 (s.d. = 0.75) for those without. A multilevel analysis showed no significant association between proband
stress and co-twin lifetime depression ($\chi^2=1.62, P=0.2$). The mean negative affect score was 0.82 (s.d.=0.19) and 0.82 (0.23) respectively for the two groups. The association between co-twin lifetime depression and proband mean negative affect was not significant ($\chi^2=1.66, P=0.2$).

Multilevel analyses showed a significant main effect of stress on negative affect ($\beta=0.034, P<0.001$). There was a significant interaction between proband stress and co-twin lifetime depression in the association with negative affect ($\chi^2=15.9, P=0.0001$), which remained significant after correction for proband SCL-90-R depression score and proband past depressive disorder ($\chi^2=16.2, P=0.0001$) and after additionally excluding all probands with current depression according to SCID interview ($n=482; \chi^2=9.95, P=0.002$).

**Dose–response relationship**

Compared with the baseline of 'very pleasant', events that were rated as 'pleasant' ($\chi^2=0.01, P=0.92$) or 'slightly pleasant' ($\chi^2=2.07, P=0.15$) did not interact with co-twin lifetime depression. The interaction effect size increased for events rated as 'neutral' ($\chi^2=3.49, P=0.06$) and 'slightly unpleasant' ($\chi^2=10.5, P=0.001$) and was smaller again for events rated as 'unpleasant' and 'very unpleasant' (respectively $\chi^2=1.25, P=0.26$, and $\chi^2=5.06, P=0.02$). Stratified effect sizes calculated with the LINCOM procedure are displayed in Fig. 2 (the numbers of observations and participants are detailed in Table 1). The two lines represent the effect sizes of stress on negative affect: one line includes and the other does not include the interaction term of stress $\times$ co-twin lifetime depression. Thus, the difference between the two lines represents the size of the interaction effect.

**Genetic contribution to the observed cross-twin, cross-trait relationship**

There was a significant three-way interaction between zygosity, co-twin lifetime depression and stress in the association with negative affect ($\chi^2=6.73, P=0.010$), indicating that the interaction effect between co-twin lifetime depression and stress was significantly stronger in monozygotic twin pairs ($\beta=0.026, P<0.001$) than in dizygotic pairs ($\beta=0.003, P=0.7$). Effect sizes stratified by zygosity calculated with the LINCOM procedure are displayed in Fig. 3.

---

**Table 1** Number of observations for each level of stress appraisal separate for subjects with and without co-twin lifetime depression, and the number of participants contributing to each number of observations (see Fig. 2)

<table>
<thead>
<tr>
<th>Event stress appraisal</th>
<th>$-3$</th>
<th>$-2$</th>
<th>$-1$</th>
<th>$0$</th>
<th>$1$</th>
<th>$2$</th>
<th>$3$</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-twin with depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations, n</td>
<td>510</td>
<td>579</td>
<td>399</td>
<td>720</td>
<td>130</td>
<td>79</td>
<td>96</td>
<td>2513</td>
</tr>
<tr>
<td>Participants, n</td>
<td>85</td>
<td>81</td>
<td>71</td>
<td>80</td>
<td>56</td>
<td>48</td>
<td>41</td>
<td>88</td>
</tr>
<tr>
<td>Co-twin without depression</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations, n</td>
<td>2956</td>
<td>2517</td>
<td>1910</td>
<td>2999</td>
<td>576</td>
<td>412</td>
<td>361</td>
<td>11731</td>
</tr>
<tr>
<td>Participants, n</td>
<td>359</td>
<td>376</td>
<td>341</td>
<td>372</td>
<td>261</td>
<td>200</td>
<td>190</td>
<td>418</td>
</tr>
</tbody>
</table>

---

**Fig. 2** Effect sizes of stressful events on negative affect: effect sizes of the affective response of stress levels 'pleasant' to 'very unpleasant' relative to the affective response of the reference category 'very pleasant', stratified by co-twin lifetime depression and corrected for continuous depression score.

**Fig. 3** Associations between stressful events and negative affect, stratified by zygosity and co-twin lifetime depression, corrected for proband continuous depression score and proband depressive disorder in the past (DZ, dizygotic; MZ, monozygotic).
DISCUSSION

Negative mood bias endophenotype

Probands with the greatest level of familial liability for depression displayed greater negative affect responses toward daily life stress. This finding cannot be accounted for by higher current depression score or diagnosis of depression in the proband since the interaction with co-twin depression persisted after controlling for proband continuous depression score and past depressive disorder, and additionally excluding probands with current depression. Thus, increased negative affect reactivity, measured as negative affect level in response to daily life stress, is also present in participants with high familial loading regardless of past or current depression. In addition, a partial dose–response relationship was found: appraisals of greater stressfulness associated with events were associated with increased interaction effects in the negative affect model. For the most stressful events the difference between the groups was smaller than for events that were appraised as less stressful; numbers, however, were smaller in the former. The conservative conclusion is that events that are moderately stressful cause very little negative affect in control participants; in

relatives of patients with psychosis showed intermediate levels of stress sensitivity (Myin-Germeys et al., 2001). These data suggest that the mood bias endophenotype may transcend the borders of traditional diagnostic classification.

Negative mood bias and genetic transmission

Our data additionally suggest that genetic factors are likely to have a role, and thus provide a link between genetic factors and a phenotype consisting of an interaction between individuals and their environment in the course of daily life. The finding is in accordance with a previous study in the same sample that found, using structural equation modelling, that the association between daily stressors and negative affect was influenced by genetic factors (Jacobs et al., 2006). The analyses suggest that the mood bias trait fulfils the criteria for endophenotype such as heritability and familial association. Whether the trait additionally shows specificity and co-segregation in families, another requisite of the ‘endophenotype’ definition (Gottesman & Gould, 2003), should become the subject of further investigation. The link between genetic factors and a person–environment momentary assessment phenotype may be helpful in tracing the link from genotype to clinical depression and may be productively further examined using molecular genetic approaches. The data show that genetic effects can be thought of as influencing person–environment interactions rather than rigidly defined psychopathological phenotypes, and underline the importance of including environmental measurements in genetic approaches towards psychiatric disorders (Kendler, 2005; Moffitt et al., 2005; Rutter, 2005).

Although the aim of the investigation was to examine the possible contribution of genetic factors to the mood bias phenotype – and evidence for such a role was found – a role of shared environmental factors is also likely. For example, had we found that negative affective response to stress was increased in people with a cotwin with depression and that this association was of the same magnitude in monozygotic and dizygotic pairs, then the conclusion would have been that only shared environmental factors contribute to the mood bias. The fact that the association was significantly greater within monozygotic pairs than in dizygotic pairs.

Table 2 Number of observations and corresponding number of participants for each level of stress appraisal for monozygotic and dizygotic participants with and without a co-twin with lifetime major depression (see Fig. 3)

<table>
<thead>
<tr>
<th>Event stress appraisal</th>
<th>–3</th>
<th>–2</th>
<th>–1</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monozygotic twins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-twin with depression</td>
<td>332</td>
<td>328</td>
<td>238</td>
<td>365</td>
<td>94</td>
<td>48</td>
<td>65</td>
<td>1470</td>
</tr>
<tr>
<td>Observation, n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants, n</td>
<td>45</td>
<td>45</td>
<td>40</td>
<td>43</td>
<td>35</td>
<td>29</td>
<td>26</td>
<td>48</td>
</tr>
<tr>
<td>Co-twin without depression</td>
<td>192</td>
<td>161</td>
<td>1168</td>
<td>1953</td>
<td>346</td>
<td>255</td>
<td>219</td>
<td>7475</td>
</tr>
<tr>
<td>Observation, n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants, n</td>
<td>232</td>
<td>233</td>
<td>216</td>
<td>234</td>
<td>166</td>
<td>126</td>
<td>123</td>
<td>262</td>
</tr>
<tr>
<td>Dizygotic twins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Co-twin with depression</td>
<td>178</td>
<td>251</td>
<td>161</td>
<td>355</td>
<td>36</td>
<td>31</td>
<td>31</td>
<td>1043</td>
</tr>
<tr>
<td>Observation, n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants, n</td>
<td>40</td>
<td>36</td>
<td>31</td>
<td>37</td>
<td>21</td>
<td>19</td>
<td>15</td>
<td>40</td>
</tr>
<tr>
<td>Co-twin without depression</td>
<td>1027</td>
<td>882</td>
<td>725</td>
<td>1030</td>
<td>225</td>
<td>153</td>
<td>141</td>
<td>4183</td>
</tr>
<tr>
<td>Observation, n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants, n</td>
<td>126</td>
<td>140</td>
<td>123</td>
<td>135</td>
<td>92</td>
<td>73</td>
<td>66</td>
<td>153</td>
</tr>
</tbody>
</table>

Fig. 3 (the numbers of observations and participants are detailed in Table 2).
provides positive evidence for the involvement of genetic factors, but does not rule out a contribution from the shared environment.

**Strengths of the study**

Strengths of the study include the fact that it was the first to use a prospective momentary assessment design in an attempt to capture a mood phenotype at the conceptual level at which it is defined: reactivity in the course of daily life. The cross-twin, cross-trait design allowed for assessment of unconfounded relationships and assessment of the role of genetic factors. Another important strength was the separate assessment of compliance and validation of the ESM procedure as published elsewhere (Jacobs et al, 2005).

**Limitations**

Some methodological limitations are apparent. First, it has been suggested that problems may arise in the ESM procedure as it depends on the compliance of participants (Kudielka et al, 2003; Broderick et al, 2004). In particular, fixed time sampling protocols may be problematic and can bias results. However, this report did not use a fixed time sampling frame, and our ESM procedure was validated in a previous report. Thus, the same sample as described in this analysis (Jacobs et al, 2005) was instructed to take, during the ESM procedure, saliva samples at each of the ten unpredictable moments during the five consecutive days. Participants recorded collection times, unaware that compliance with the sampling protocol was being investigated by means of electronic monitoring devices. Results showed that compliance was high (over 90%) and inclusion of the inaccurately timed samples did not distort the data (Jacobs et al, 2005). Therefore, results from the ESM procedure in this report can be considered valid.

Another issue is that higher levels of negative affect itself may represent the real vulnerability factor, rather than negative affect reactivity to stress. Healthy probands with co-twin siblings who have a lifetime diagnosis of depression may show increased levels of negative affect compared with those whose co-twin has no such diagnosis; higher levels of negative affect give rise to more variability, which in turn enhances the detection of stress sensitivity. However, no difference in negative affect was apparent between participants with and without family loading for depression and this explanation is therefore unlikely. Furthermore, this study was longitudinal but essentially assessed multiple cross-sectional relationships at each ESM moment, which made it impossible to establish causal relationships. Therefore, it is impossible to determine whether stress measures influence mood, or whether mood influences subjective appraisals of stress. However, either explanation bears clinical relevance, and the interpretation of stress at least in part contributing to measures of mood has face validity. Finally, since only women were included in this study, the results are not necessarily generalisable to the male population.

**ACKNOWLEDGEMENTS**

This research was supported by the Dutch Foundation for Scientific Research. The East Flanders Prospective Survey has been in part supported by grants from the Fund of Scientific Research, Flanders, and Twins, the Association for Scientific Research in Multiple Births, Belgium. We thank all the participating twins for their cooperation.

**REFERENCES**


Women with autistic-spectrum disorder: magnetic resonance imaging study of brain anatomy

MICHAEL C. CRAIG, SHAHID H. ZAMAN, EILEEN M. DALY, WILLIAM J. CUTTER, DENE M. W. ROBERTSON, BRIAN HALLAHAN, FIONA TOAL, SUZIE REED, ANITA AMBIKAPATHY, MICK BRAMMER, CLODAGH M. MURPHY and DECLAN G. M. MURPHY

Background Our understanding of anatomical differences in people with autistic-spectrum disorder, is based on mixed-gender or male samples.

Aims To study regional grey-matter and white-matter differences in the brains of women with autistic-spectrum disorder.

Method We compared the brain anatomy of 14 adult women with autistic-spectrum disorder with 19 controls using volumetric magnetic resonance imaging and voxel-based morphometry.

Results Women with autistic-spectrum disorder had a smaller density bilaterally of grey matter in the fronto-temporal cortices and limbic system, and of white matter in the temporal lobes (anterior) and pons. In contrast, they had a larger white-matter density bilaterally in regions of the association and projection fibres of the frontal, parietal, posterior temporal and occipital lobes, in the commissural fibres of the corpus callosum (splenium) and cerebellum (anterior lobe). Further, we found a negative relationship between reduced grey-matter density in right limbic regions and social communication ability.

Conclusions Women with autistic-spectrum disorder have significant differences in brain anatomy from controls, in brain regions previously reported as abnormal in adult men with the disorder. Some anatomical differences may be related to clinical symptoms.

Declaration of interest None.

The prevalence of autistic-spectrum disorder in the UK has recently been estimated to be approximately 1%, and is 3–4 times higher in males than in females (Baird et al, 2006). However, despite the relatively high prevalence and heritability of this disorder, its pathophysiology remains incompletely defined. Recent studies have helped define the neuroanatomical and functional abnormalities underlying the condition (reviewed by Toal et al, 2005); however, these data have been acquired in male-only (or predominantly male) samples, and at different ages. To date there has been no study of regional differences in grey and white matter in female-only samples, or in adult women (when brain development is complete). Hence the biological associates of autistic-spectrum disorder in adult women are largely unknown. It has been reported that the behavioural phenotype of females with the disorder is different from that of males (Lord et al, 1982; McLennan et al, 1993; Gillberg & Coleman, 2000). Also, there are gender differences in postnatal brain development and ageing (Murphy et al, 1996; Giedd et al, 1997; Gur et al, 2002). It is therefore possible that the neuropathology of autistic-spectrum disorder in females is different from that reported in males.

Brain anatomy in vivo can be measured using magnetic resonance imaging (MRI) and a variety of analytical approaches, including hand tracing methods and voxel-based morphometry (VBM). Hand tracing allows measurement of relatively large regional bulk volumes (i.e. with no differentiation of grey and white matter), whereas the latter technique allows analysis of subtle regional differences in grey and white matter. We therefore used MRI and VBM to investigate the brain anatomy of women with autistic-spectrum disorder.

METHOD

The sample consisted of participants in a clinical research programme enabled by the Medical Research Council UK Autism Imaging Multicentre Study (AIMS) network, and the study was jointly conducted by South London and Maudsley National Health Service (NHS) Foundation Trust and the Institute of Psychiatry, London. We included 19 women in a control group (mean age 35.0 years, s.d.=14.0) and 14 women with an autistic-spectrum disorder: 10 with Asperger syndrome and 4 with autism (mean age 37.9 years s.d.=11.4). Participants were diagnosed using the ICD-10 clinical research criteria (World Health Organization, 1992). This was achieved by consensus between two clinicians, experienced in diagnosis of autistic-spectrum disorders, and a nurse, all trained in the use of the autism diagnostic measures used in the study. The diagnosis was based on clinical interviews, collateral information from family members and review of other information available, such as school reports. In addition, we were able to use the Autism Diagnostic Interview – Revised (ADI–R; Lord et al, 1994) to assess 7 individuals whose parental informants were willing and available and the Autism Diagnostic Observation Schedule (ADOS; Lord et al, 1989) to assess a further 5 participants who were willing to undertake further interviewing. Thus, we confirmed clinical research criteria in all participants using ICD–10, and in 12 of the 14 individuals with the ADI–R or ADOS. All assessments were masked to MRI data.

All participants underwent a structured clinical examination and routine clinical blood tests to exclude biochemical, haematological or chromosomal abnormalities. Individuals were excluded if they had a history of major psychiatric disorder (e.g. psychosis), head injury, toxic exposure, diabetes, abnormalities in routine blood tests, drug or alcohol misuse, clinical abnormality on routine MRI, or a medical or genetic disorder associated with autistic symptoms (e.g. epilepsy, tuberous sclerosis or fragile X syndrome). All participants gave informed consent and/or assent (as approved by the Institute of Psychiatry and the South London and Maudsley NHS Trust research ethics committee). None was taking medication at the time of testing.

Neuropsychological testing

Overall intellectual ability (IQ) was determined using an abbreviated Wechsler Adult Intelligence Scale (WAIS–R; Canavan & Beckmann, 1993).
**Image acquisition**

All MRI data were obtained using a GE Signa 1.5 T neuro-optimised magnetic resonance system (General Electric, Milwaukee, USA). Whole-head coronal three-dimensional spoiled gradient recalled (3D-SPGR) images (repetition time = 13.8 ms, echo time = 2.8 ms, 256 × 192 acquisition matrix, 124 slices, thickness 1.5 mm) were obtained from all participants.

**VBM pre-processing**

Voxel-based morphometry pre-processing was performed on the 3D-SPGR data using Statistical Parametric Mapping software (SPM2; Wellcome Department of Imaging Neurosciences, University College London, UK). The image processing steps have been described in detail elsewhere (Abell et al., 1999; Good et al., 2001).

The segmentation algorithm implemented in SPM2 incorporates a priori knowledge of the likely spatial distribution of tissue types in the brain through use of prior probability tissue maps derived from a large number of individuals. To ensure the most accurate segmentation possible, we created study-specific customised prior probability maps based on all 33 participants. The pre-processing stages were as follows:

(a) scans were segmented into probabilistic maps of grey and white matter and cerebrospinal fluid using a modified mixture model clustering algorithm;

(b) the segmented grey matter map was mapped to a grey matter template and the derived warping parameters were applied to the original T1-weighted image in order to map it into standard space (this procedure prevents skull and other non-brain voxels from contributing to the registration, while avoiding the need for explicit skull-stripping);

(c) the registered image was then resegmented, which is necessary as the a priori knowledge incorporated into the SPM2 segmentation algorithm means that it works optimally on images in standard space. The segmented maps were then corrected for volume changes introduced during the registration and smoothed using a Gaussian filter of 5 mm full width at half-maximum. Total grey and white matter densities were calculated from the segmented maps in native space.

**VBM analysis**

For the VBM analyses, between-group differences in grey- and white-matter density were calculated by fitting an analysis of covariance (ANCOVA) model at each intracerebral voxel in standard space, covarying for total grey-matter (or white-matter) density. Structural brain changes are likely to extend over a number of contiguous voxels and therefore test statistics incorporating spatial information, such as three-dimensional cluster mass (the sum of suprathreshold voxel statistics), are generally more powerful than other possible test statistics which are informed only by data at a single voxel. Therefore, our approach was to provisionally set a relatively lenient P value (P ≤ 0.05) to detect voxels putatively demonstrating differences between groups. We then searched for spatial clusters of such voxels. At the cluster level, rather than set a single a priori P value below which we would regard findings as significant, we calculated for a range of P values the number of clusters that would be expected by chance alone. We then set the statistical threshold for cluster significance by data-driven permutation testing. This was done such that the expected number of false positive clusters is less than 1, and we quoted the P value at which this occurs (Bullmore et al., 1999; Sigmundsson et al., 2001).

**Post hoc analysis of behavioural scores**

Finally, we carried out a preliminary (post hoc) analysis to determine if differences in brain density were associated with behavourial abnormality within people with autistic-spectrum disorder. To do this, we related (using Pearson product-moment correlation coefficients) severity of clinical symptoms within people with the disorder as measured by the ADI-R to the density of brain regions, which differed significantly from controls.

**RESULTS**

The characteristics of the sample are given in Table 1. There was no significant difference between women with autistic-spectrum disorder and controls in age, IQ or total brain grey- and white-matter density (in native space generated by SPM2).

**Voxel-based morphometry**

The three-dimensional cluster maps of the between-group differences in grey- and white-matter volume were large and extended into several regions.

**Grey matter**

All grey-matter differences between the autistic-spectrum disorder group and the control group were significant at P ≤ 0.002, the value at which less than 1 false positive cluster was expected by chance alone (Table 2). Women with the disorder had a significantly smaller grey-matter density than controls bilaterally in the temporal lobes (including parahippocampal gyrus), orbito-frontal cortex (medial and lateral) and the basal ganglia (lentiform nucleus and caudate nucleus), in the right medial occipital (left cuneus) lobe, and in the left frontal (right anterior cingulate) lobe (Fig. DS1 in the data supplement to the online version of this paper).

**White matter**

All white-matter differences between the groups were significant at P ≤ 0.01, the value at which less than 1 false positive cluster was expected by chance alone (see Table 2). Women with the disorder had a significantly smaller white-matter density bilaterally in the anterior temporal lobes and brain-stem (pons). In contrast, they had a significantly increased white-matter density bilaterally in the association and projection fibres of the frontal, parietal, posterior temporal and occipital lobes, in the commissural fibres of the corpus callosum (splenium) and cerebellum (anterior lobe) (Fig. DS2).

**VBM analysis of correlations with ADI score**

There was a negative correlation (r = -0.767, n = 7, P = 0.04) between reduced grey matter in the right limbic regions (including anterior and posterior cingulate, parahippocampal gyrus and uncus) and qualitative abnormalities in reciprocal social interaction (Fig. 1).

**DISCUSSION.**

Our main study findings were that women with autistic-spectrum disorder have a significantly reduced density bilaterally of grey matter within the fronto-temporal cortices and limbic system, and of white matter in the anterior temporal lobes. In contrast, they have increased white matter bilaterally in the fronto-parietal, posterior temporal.
Grey matter and white matter volumes were compared between ASD (n = 14) and controls (n = 19). The table shows the mean (s.d.) values for age, FSIQ, PIQ, VIQ, grey matter, and white matter volumes. The table also includes the Clusters of significantly decreased and increased grey-matter (P = 0.004) and white-matter (P = 0.01) volume in women with autistic-spectrum disorder compared with controls.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Sample characteristics and volumes of grey and white matter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ASD (n = 14)</td>
</tr>
<tr>
<td></td>
<td>Mean (s.d.)</td>
</tr>
<tr>
<td>Age, years</td>
<td>37.9 (11.4)</td>
</tr>
<tr>
<td>FSIQ</td>
<td>103.4 (17.0)</td>
</tr>
<tr>
<td>PIQ</td>
<td>105.1 (16.6)</td>
</tr>
<tr>
<td>VIQ</td>
<td>100.0 (20.3)</td>
</tr>
<tr>
<td>Grey matter, ml</td>
<td>607.4 (51.2)</td>
</tr>
<tr>
<td>White matter, ml</td>
<td>368.1 (34.2)</td>
</tr>
</tbody>
</table>

ASD, autistic-spectrum disorder; FSIQ, full-scale IQ; PIQ, performance IQ; VIQ, verbal IQ.

Table 2 | Clusters of significantly decreased and increased grey-matter (P = 0.004) and white-matter (P = 0.01) volume in women with autistic-spectrum disorder compared with controls.

<table>
<thead>
<tr>
<th>Cluster centroid (and other regions included in cluster)</th>
<th>Brodmann area of centroid (and other BAs included in cluster)</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>Voxels</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grey matter Decreases</td>
<td>Cuneus (lingual gyrus, precuneus, extending to right occipital lobe)</td>
<td>31 (18,31)</td>
<td>1</td>
<td>-66</td>
<td>7</td>
<td>462</td>
</tr>
<tr>
<td>L. temporal lobe</td>
<td>Inferior temporal gyrus (middle temporal/fusiform gyri posterior cingulate, parahippocampal gyrus, uncus)</td>
<td>37 (18, 20, 28, 30, 31, amygdala, hippocampus)</td>
<td>59</td>
<td>-47</td>
<td>-9</td>
<td>880</td>
</tr>
<tr>
<td>L. temporal lobe</td>
<td>Superior temporal gyrus (transverse temporal gyrus extending to inferior frontal/subcallosal gyri, caudate, claustrum, lentiform nucleus)</td>
<td>38 (18, 22, 28, 30, 34, 42, 47)</td>
<td>34</td>
<td>0</td>
<td>-16</td>
<td>2314</td>
</tr>
<tr>
<td>R. temporal lobe</td>
<td>Middle temporal gyrus (transverse/inferior/superior temporal/fusiform gyri and subgyral extending to middle/superior frontal and subcallosal gyri, and lentiform nucleus)</td>
<td>21(10, 11, 20, 34, 38, 42)</td>
<td>44</td>
<td>-5</td>
<td>-16</td>
<td>3097</td>
</tr>
<tr>
<td>R. limbic lobe</td>
<td>Anterior cingulate (posterior cingulate, parahippocampal gyrus, uncus)</td>
<td>32 (30, 38, amygdyla)</td>
<td>8</td>
<td>43</td>
<td>4</td>
<td>341</td>
</tr>
</tbody>
</table>

White matter Decreases | Superior temporal gyrus | 46 | -17 | -10 | 232 |
| L. temporal lobe | Middle temporal gyrus (inferior temporal gyrus) | 43 | -13 | -21 | 340 |
| R. brain-stem | Medulla (extending into left pons) | 2  | -20 | -35 | 296 |

Increases | Precentral/subgyral frontal gyri, subgyral temporal, anterior / posterior cingulate, parahippocampus, inferior / post-central parietal gyri, cuneus, anterior lobe cerebellum | 24 | -23 | 24 | 5509 |
| L. temporal lobe | Middle temporal gyrus | 28 | -60 | 22 | 200 |
| L. frontal lobe | Medial frontal, middle/superior temporal/parahippocampal / cingulate gyri and precuneus | 28 | -16 | 30 | 6137 |
| R. temporal lobe | Middle temporal gyrus | 27 | -63 | 22 | 220 |
| R. cerebellum | Anterior lobe | 27 | -29 | 41 | 2180 |

BA, Brodmann area; L, left; R, right.
males. Furthermore, these brain regions are implicated in some of the higher cognitive functions reported as abnormal in people with this disorder (e.g. social cognition, language, motor control and ‘theory of mind’).

Our preliminary finding was that in women with autistic-spectrum disorder reduced grey-matter density in limbic regions is correlated with abnormal social behaviour. It is possible that this finding is attributable to a type I error as we carried out multiple comparisons; however, it is tentatively supported by reports of social and emotional deficits in (macaque) monkeys following lesions of the anterior cingulate (Bachevalier & Merjanian, 1994; Rudebeck et al., 2006), and social cognitive deficits in humans following damage to the limbic system (Stone et al., 2002); further and larger studies are required to examine this issue.

However, there are some differences between our findings and previous neuroanatomical imaging studies of adult males with autistic-spectrum disorder (Abell et al., 1999; McAlonan et al., 2002). For example, in this study we found that women with this disorder have no difference in density of cerebellar grey matter, but they have excess white matter. Prior studies of men have reported both excess (Abell et al., 1999) and reduced (McAlonan et al., 2002) grey matter, but no difference in white matter. Cerebellar pathology has been reported in many post-mortem case studies across a variety of ages and IQ scores (21 out of 29 studies have reported reduced Purkinje cells; Palmen et al., 2004) and cerebellar hypoplasia has been found by some structural imaging studies (Ciesielski et al., 1997; Levitt et al., 1999; Carper & Courchesne, 2000; Courchesne et al., 2001). Thus the cerebellum is most probably abnormal in both men and women with autistic-spectrum disorder – but it is unclear if the neuropathology is similar in both genders.

In the light of our finding that women with autistic-spectrum disorders have abnormalities in brain anatomy that are broadly similar to those previously reported in men, the reasons for the gender difference in the prevalence of this disorder remain unclear. It has been suggested that there is a relative failure to diagnose these disorders in females because of differences in clinical presentation. For example (as noted above), it has been reported that females with the disorder have a different behavioural phenotype to males, with a lower frequency of comorbid challenging behaviours (McLennan et al., 1993) and fewer abnormal special interests (Gillberg & Coleman, 2000); they are less likely to exhibit stereotypic behaviour during play (Lord et al., 1982) and have better superficial social skills and language (Gillberg & Coleman, 2000). Alternatively, it might be that the increased prevalence of autistic-spectrum disorder reported in males (and gender differences in clinical presentation) is due to significant differences in biological vulnerability. Thus the underlying genetic susceptibility for the condition may be similar in both genders, but there may be a lower ‘threshold’ to developing autism in males. If so, the putative increased vulnerability of the male brain is probably due to a number of complex (and interacting) factors, including genomic imprinting (Badcock & Crespi, 2006), hormonal milieu (Baron-Cohen et al., 2005) and gender differences in the normal maturational trajectory of the brain regions implicated in this disorder (Giedd et al., 1999; Gogtay et al., 2004). For example, the development of frontal and parietal grey matter peaks approximately 1 year earlier in adolescent girls than in boys, and the amygdala increases in density in healthy boys but not in girls.

A further biological explanation for gender differences in the prevalence of autistic-spectrum disorder builds on the concept that autism represents an ‘extreme male brain’ (Asperger, 1944) by applying empathising–systemising theory (Baron-Cohen, 2002); this theory suggests that the female brain is predominantly ‘hard-wired’ for empathy, and that the male brain is predominantly ‘hard-wired’ for understanding and building systems (systematising). It is therefore proposed that people with autism may have an ‘extreme male brain’ that is even stronger at systematising and weaker at empathising than the normal male brain, and that this is underpinned by a skew of the normal gender differences in neurodevelopment. This may be due to an ‘extreme’ variation in the typical gender differences observed in brain regions that modulate processes involved in empathy (e.g. the amygdala) and/or systematising. Further, it has been suggested that normal gender differences postulated by empathising–systemising theory might be primarily due to an increase in the ratio of local white-matter tracts (important for systematising) to longer-range, interhemispheric tracts (important for empathising) in males, and that this skewed balance in connectivity is further exaggerated in autism (Baron-Cohen et al., 2005).

The ‘extreme male brain’ theory implicitly suggests that the skew in normal gender differences (i.e. in the maturation of specific brain regions such as the amygdala and of the ratio of interconnected white-matter tracts) will need to be even more ‘extreme’ in females compared with males with this disorder. The design of the study reported here did not allow us to test this hypothesis directly. Nevertheless, our results do suggest that females with autistic-spectrum disorder have abnormalities in brain regions and systems associated with empathising – such the parietal cortex and limbic regions – which are consistent with the theory. However, further imaging studies are needed to examine directly the differences in the brain anatomy of men and women with this condition.

Our study was limited by a number of factors, including a relatively small sample size, a cross-sectional design and the application of multiple statistical comparisons (i.e. increased risk of type 1 error). However, we believe these limitations are unlikely to explain our results fully. In particular, type 1 errors are unlikely to account for our reported findings with the minimal-assumption, data-driven (permutation) methods we used. Studies of the assumptions of normal theory based methods have often raised issues about the validity of these assumptions (see, for example, Hayasaka & Nichols, 2003; Thirion et al., 2007). However, we used a two-stage
inferential procedure in which permutation inference: random field and permutation methods. are the best approach to inference in MRI analysis.

In summary, our study suggests that adult women with autistic-spectrum disorder have significant differences from controls in brain anatomy, and these abnormalities are broadly similar to those observed in predominantly male populations with this disorder of similar age and IQ. Larger studies are needed to relate anatomy to behaviour and directly compare females and males with autism across the life span.

ACKNOWLEDGEMENTS

The authors thank the people with autistic-spectrum disorder who took part in the study, the Medical Research Council UK APHP programme for infrastructure support, and Professor Nancy Minshew and Dr Marco Catani for their valuable comments during the preparation of this manuscript. This project was assisted by support from the South London and Maudsley NHS Trust.

REFERENCES


Sulcal thickness as a vulnerability indicator for schizophrenia

VINA M. GOGHARI, KELLY REHM, CAMERON S. CARTER
and ANGUS W. MACDONALD

Background People with schizophrenia may demonstrate cortical abnormalities, with gyri and sulci potentially being differentially affected.

Aims To measure frontal and temporal sulcal cortical thickness, surface area and volume in the non-psychotic relatives of patients with schizophrenia as a potential vulnerability indicator for the disorder.

Method An automated parcellation method was used to measure the superior frontal, inferior frontal, cingulate, superior temporal and inferior temporal sulci in the relatives of patients (n=19) and controls (n=22).

Results Compared with controls, relatives had reversed hemispheric asymmetry in their cingulate sulcal thickness and a bilateral reduction in their superior temporal sulcal thickness.

Conclusions Cingulate and superior temporal sulcal thickness abnormalities may reflect neural abnormalities associated with the genetic liability to schizophrenia. Cortical thinning in these regions suggests that liability genes affect the dendrites, synapses or myelination process during the neurodevelopment of the cortical mantle.

Declaration of interest None. Funding detailed in Acknowledgements.

Sulco-gyral patterns are thought to form through minimisation of axonal tension during cortico-cortical connection formation (Van Essen, 1997; Hilgetag & Barbas, 2005). This process results in morphological differences in the gyri and sulci, which include the number and shape of cells. Early disruption of axonal tracts can affect later development of the cortical mantle. Assessing differences in the cortical integrity of sulci compared with gyri may index neurodevelopmental abnormalities in cortical development in schizophrenia. Recent advances in brain imaging methods now allow efficient measurement of both sulci and gyri, including cortical thickness (e.g. Fischl et al, 2004). Investigations of the cortical mantle have demonstrated thickness abnormalities in temporal gyrus and in frontal, temporal and parietal sulci in children and adolescents with schizophrenia-spectrum disorders (White et al, 2003); increased right temporal cortical folding (Harris et al, 2004a) and reduced posterior cingulate folding in patients with schizophrenia (Wheeler & Harper, 2007); and increased right frontal cortical folding in high-risk individuals who later developed schizophrenia (Harris et al, 2004b). Studying adult non-psychotic relatives of patients provides a way to investigate the potential effects of genes or gene–environment interactions not confounded by the illness process or medication. Family studies have found increased frontal cortical folding in unaffected relatives (Falkai et al, 2007) and reduced left frontal cortical folding in high-risk individuals (Jou et al, 2005). We previously investigated gyri in the same sample of non-psychotic relatives of people with schizophrenia and found abnormalities in the cingulate, superior temporal, middle temporal and parahippocampal regions (Goghari et al, 2007). To characterise further the potential effect of susceptibility genes on cortical folding, we evaluated sulcal thickness, surface area and grey-matter volume in frontal and temporal regions. This complementary focus on sulci allows a more complete understanding of the potential structural abnormalities of the cortical mantle related to the schizophrenia diathesis.

METHOD

Participants Twenty-two first-degree relatives of patients with schizophrenia and 23 controls participated. First-degree relatives were recruited by first identifying probands receiving treatment through the Western Psychiatric Institute and Clinic in Pittsburgh, Pennsylvania. Diagnosis for probands was confirmed using the Struktured Clinical Interview for DSM–IV (First et al, 1996). Control group participants were recruited from the general population. Individuals were excluded if they had experienced head trauma, seizures or had a diagnosis of substance misuse or dependence within the preceding 6 months. Control group candidates were excluded additionally if they had a family history of psychosis. Relatives and controls were screened for psychiatric symptoms or disorders using the Structured Clinical Interview for DSM–III–R, Non-patient version (Spitzer et al, 1990). Relatives were additionally assessed with the Structured Interview for DSM–III Personality Disorders–Revised (SID–P) for cluster A personality disorders (Pfohl et al, 1982). One control met criteria for DSM–IV diagnosis in the past month (intermittent explosive disorder). Three relatives and three controls met criteria for lifetime diagnosis for major depressive disorder. In addition, one relative met criteria for schizophrenia personality disorder. This protocol was approved by the University of Pittsburgh institutional review board. After complete description of the study to the participants, written informed consent was obtained.

Image acquisition and processing Spoiled gradient recalled magnetic resonance imaging (MRI) scans were acquired in the axial plane using a GE Signa (General Electric, Milwaukee, Wisconsin, USA) 3 T magnetic resonance scanner (0.9375 mm x 0.9375 mm x 1.5 mm voxel size, 124 slices). Whole-brain volumes were extracted using MipStrip, an automated, consensus-based stripping algorithm (Rehm et al, 2004). Whole-brain volumes were intensity-corrected using non-parametric non-uniform intensity normalisation (N3) to improve the accuracy of tissue classification.
and cortical surface extraction (Sled et al., 1998). An automatic algorithm labelled the left and right hemispheres, cerebellum and brain-stem (Rehm et al., 2005). Defects were manually edited where necessary. These stripped, intensity-corrected brain images and grossly subdivided volumes were used in the subsequent steps.

Surface extraction and cortical parcellation were conducted using FreeSurfer version 1.3 (Fischl et al., 2004) (DSI, see data supplement to online version of this paper). Briefly, the stripped, intensity-corrected, subdivided volume was segmented to classify white matter and to approximate the grey–white matter boundary for each cortical hemisphere, from which a topologically correct grey-white matter boundary surface triangulation was generated (Dale et al., 1999; Fischl et al., 2001). Subsequently, a pial surface was generated using a deformable surface algorithm (Fischl & Dale, 2000). All surfaces were visually inspected and defects leading to major topological errors were manually corrected. The grey–white boundary surface was inflated and individual differences in curvature were normalised. For each brain examined the inflated surface was morphed into a sphere and registered to an average spherical surface (Fischl & Dale, 2000). All analyses were conducted both with and without this participant's data were removed.

Parcellation for each individual's left and right hemisphere surface was transmitted from the parcellation of the reference spherical surface to which they were aligned. Each parcellated region was mapped back onto each individual brain's inflated surface by inverting the algorithm that morphed the inflated surface to the average spherical surface representation (Kuperberg et al., 2003; Fischl et al., 2004). Eighty-five parcellation units were provided by FreeSurfer, based on the conventions of Duvernoy (1991).

Thicknesses were calculated using FreeSurfer software; they were computed for each vertex in the triangulated surfaces by finding the point on the white-matter surface that was closest to a given point of the pial surface (and vice versa) and the average was taken between these two values. Surface areas were calculated for both the white–grey boundary and pial surface using the formula of Heron (Gellert et al., 1975). The average of the two surface areas at each triangle was computed and was used in analyses to simulate an intermediate cortical surface. Grey-matter volumes were calculated by constructing a closed mesh joining a pair of linked triangles and computing the enclosed volume (Eberly et al., 1991). Individual triangle volumes were aggregated to compute cortical grey-matter volume for each region of interest. This methodology has been extensively validated (Kuperberg et al., 2003; Fischl et al., 2004). A trained rater manually traced regions of interest on ten white–grey boundary inflated surfaces. An index of similarity between the automated and hand-drawn regions was calculated. This index was defined as the ratio of twice the common area between the two methods relative to the sum of the individual areas; a value above 0.7 was considered excellent correspondence (Zijdenbos et al., 1994). All three sulcal regions had acceptable correspondence (superior frontal 0.71, cingulate 0.82, superior temporal, 0.87).

**Analyses**

All sulcal regions of interest were assessed using mixed-model analysis of covariance (ANCOVA), with hemisphere (left, right) entered as an intra-individual effect and group (control, relative) entered as an inter-individual effect. Age, gender and average cortical thickness were entered as covariates when analysing regional thicknesses; age, gender and total cortical surface area were entered as covariates when analysing regional surface areas; and age, gender and intracranial volume were entered as covariates when analysing regional volumes. All analyses were also conducted with education as an additional covariate; as this did not affect the results, the findings reported below do not include it as a covariate. Greenhouse–Geisser correction is reported for the mixed-model ANCOVAs. These planned comparisons were set to a significance threshold of $P=0.05$. Because of concerns related to multiple comparisons, effect sizes are reported as a supplemental indicator to significance values. Partial eta-squared effect sizes are presented: 0.01 is considered a small effect, 0.06 is considered a medium effect and 0.14 is considered a large effect (Stevens, 2002).

**RESULTS**

Data for the final sample are presented in Table 1. No significant difference between groups was found for any demographic variable. Although relatives tended to have slightly less educational attainment than the controls, parents were matched for education. The imaging data from three relatives and one control had to be excluded owing to poor quality. As one relative met criteria for schizotypal personality disorder, all analyses were conducted both with and without this participant: none of the findings changed substantially when this participant’s data were removed.

**MRI analyses**

**Frontal sulcal measures**

Planned comparisons of middle and inferior frontal sulcal thickness, surface area and volume revealed no significant main effect of group or hemisphere by group interactions (all $P>0.2$).

**Cingulate sulcal measures**

A significant hemisphere by group interaction was found in cingulate sulcal thickness ($F=3.50, \text{d.f.}=1,36, P=0.03$; partial $\eta^2=0.13$) with the relatives group having a reversal in hemispheric asymmetry pattern compared with the control group. The control group had thicker sulci in the left hemisphere, whereas those in the relatives group had thicker sulci in the right hemisphere.

**Table 1** Demographic characteristics of the sample

<table>
<thead>
<tr>
<th></th>
<th>Control group</th>
<th>Relatives group</th>
<th>Test statistic</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age, years (s.d.)</strong></td>
<td>34.1 (8.4)</td>
<td>34.2 (11)</td>
<td>$t=0.02$</td>
<td>0.98</td>
</tr>
<tr>
<td><strong>Male gender, n (%)</strong></td>
<td>10 (46)</td>
<td>8 (42)</td>
<td>$\chi^2=0.05$</td>
<td>0.83</td>
</tr>
<tr>
<td><strong>Minority ethnicity, n (%)</strong></td>
<td>6 (27)</td>
<td>9 (47)</td>
<td>$\chi^2=1.78$</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Education, years (s.d.)</strong></td>
<td>Participants</td>
<td>15.7 (1.9)</td>
<td>14.2 (3.1)</td>
<td>$t=-1.79$</td>
</tr>
<tr>
<td></td>
<td>Father’s education</td>
<td>13 (2.8)</td>
<td>13.6 (2.6)</td>
<td>$t=0.68$</td>
</tr>
<tr>
<td></td>
<td>Mother’s education</td>
<td>13.1 (2.2)</td>
<td>14.1 (2.7)</td>
<td>$t=1.24$</td>
</tr>
</tbody>
</table>

Note: Table 1. No significant difference between groups was found for any demographic variable. Although relatives tended to have slightly less educational attainment than the controls, parents were matched for education. The imaging data from three relatives and one control had to be excluded owing to poor quality. As one relative met criteria for schizotypal personality disorder, all analyses were conducted both with and without this participant: none of the findings changed substantially when this participant’s data were removed.
hemi-sphere (Fig. 1, Table 2). There were also significant effects of age (F = 7.26, d.f. = 1,36, P = 0.01) and average cortical thickness (F = 23.85, d.f. = 1,36, P < 0.001) covariates. No significant difference was found in surface area or volume (all P > 0.1).

**Temporal sulcal measures**

A significant effect of group was found in the superior temporal sulcal thickness (F = 8.17, d.f. = 1,36, P = 0.007; partial \( \eta^2 = 0.19 \)) with the relatives group having 2.4% bilateral decrease (Fig. 1, Table 2). There was also a significant effect of average cortical thickness (F = 119.87, d.f. = 1,36, P < 0.001) covariate. No significant difference was found in surface area or volume (all P > 0.1). Planned comparisons of inferior temporal sulci thickness, surface area and volume revealed no significant main effect of group or hemisphere by group interactions (P > 0.1).

**DISCUSSION**

This study evaluated whether frontal and temporal sulcal abnormalities were associated with a possible genetic vulnerability to schizophrenia. To examine possible grey-matter abnormalities comprehensively we measured cortical thickness, surface area and volume. The principal findings were that non-psychotic relatives of patients with schizophrenia demonstrated abnormalities in thickness in the cingulate and superior temporal sulci.

Consistent with previous findings of cingulate cortex abnormalities, we found that cingulate sulcal thickness was abnormal among the relatives, who were observed to have a reversed asymmetry pattern with the right hemisphere being thicker than the left. Similar to our findings, studies have demonstrated that controls had more leftward asymmetry in their paracingulate sulci compared with patients with schizophrenia (Yucel et al., 2002; Le Provost et al., 2003), whereas the patients with schizophrenia had more rightward asymmetry in their paracingulate sulci (Le Provost et al., 2003). Compared with a control group, individuals with a high risk of developing schizophrenia tended to have a significantly more interrupted left cingulate sulcus and less well-developed left paracingulate sulcus (Yucel et al., 2003). Recent MRI studies of cortical thickness have found the cingulate cortex to be thinner in people with first-episode schizophrenia (Narr et al., 2005) and specifically the anterior cingulate gyri to be thinner in patients with chronic schizophrenia (Kuperberg et al., 2003). A study assessing the gyrification index, defined as the ratio of the length of the contour of the gyri and sulci compared with the length connecting the gyral surface, demonstrated significant reductions in posterior cingulate cortical folding in schizophrenia (Wheeler & Harper, 2007). Consistent with neuroimaging results of cortical thinning, post-mortem work on the anterior cingulate cortex of people with schizophrenia has demonstrated glial cell loss and changes in cell size and density (Benes et al., 1991, 2001; Stark et al., 2004).

However, a greater number of axons in layer II and sub-lamina III A has also been reported (Benes et al., 1986). The cingulate gyri in the current sample have also been noted to be thinner and have reduced surface area and volume in the relatives group (Goghari et al., 2007). Taken together these results suggest neurodevelopmental abnormalities in this limbic region, which has an essential role in affective, cognitive and motor control systems (Vogt et al., 1992).

Consistent with temporal cortex abnormalities reported in the literature, we found a bilateral reduction in thickness in the superior temporal sulci in the non-psychotic relatives of patients with schizophrenia compared with controls. Consistent with this finding, the cell density in a variety of temporal lobe sulci was reduced in schizophrenia (Chance et al., 2004). A study of male patients with chronic schizophrenia demonstrated faster cerebrospinal fluid volume expansion in posterior superior temporal sulci compared with controls over a span of 4 years (Mathalon et al., 2001). The superior temporal sulci border both the superior and middle temporal gyri. Previously in the current sample, only surface area and not thickness was decreased in the superior temporal gyri (Goghari et al., 2007). In addition, the left middle temporal and bilateral parahippocampal gyri were greater in surface area and volume in the relatives group. Superior temporal lobe gyrus volume has also been shown to be reduced in the young offspring of patients (Rajarethinam et al., 2004). Together, these findings suggest that the temporal sulci and gyri are globally affected. Thus, the temporal cortex may be an important indicator of the schizophrenia diathesis.
We failed to find any significant structural differences in the middle and inferior frontal regions, although functional MRI data collected on the same sample of relatives and controls suggest abnormalities in prefrontal activity (Brodmann areas 9, 8 and 6) (MacDonald et al., 2006). In contrast to our structural findings, Falkai et al. (2007) conducted a gyrification study of patients with schizophrenia and unaffected relatives and found both groups to have a higher frontal gyrification index than a control group. Findings have been inconsistent, however, with one study finding that high-risk individuals who later developed schizophrenia had a higher gyrification index in the right frontal lobe (Harris et al., 2004a), and another study finding that high-risk individuals had no difference in the right frontal lobe, but rather a reduced gyrification index in the left frontal lobe (Jou et al., 2005). These discrepancies in findings may be sample-specific or due to methodological differences. Our sample size is modest and might not have been sufficient to detect smaller group differences. However, in our sample, the effect sizes indicated the magnitude of the difference between the groups was small: superior and inferior frontal sulcal thickness, partial $\eta^2=0.06$; surface area, partial $\eta^2=0.02$; volume, partial $\eta^2=0.01$.}

**Limitations**

We attempted to control for multiple comparisons by restricting our analyses to five planned comparisons of regions consistently associated with schizophrenia. There is a need for caution when interpreting results that derive from multiple comparisons. However, this caution has to be balanced by the need to be sensitive to group differences. Our approach to this balance was to report effect sizes, which reflect the magnitude of the difference between groups regardless of sample size. Our results indicated the abnormality in the cingulate sulcal thickness (effect size 0.13) and the bilateral reduction in superior temporal sulci (effect size 0.19) are both noteworthy findings. This study compared the non-psychotic relatives of patients with schizophrenia and controls. It would have been useful in addition to have collected patient pedigree information to quantify genetic liability on a continuum. We used an automated method to measure cortical thickness and to parcellate the cerebral cortex. The accuracy of the thickness values depends on the accuracy of the grey–white segmentation and therefore can be influenced by various artefacts. However, we followed validated data processing procedures provided by the FreeSurfer manual and all analyses were performed in consultation with a magnetic resonance physicist (K.R.).

**Implications**

Sulci and gyri are thought to form together, with gyri forming between densely connected regions and sulci forming between weakly connected regions (Hilgetag & Barbas, 2006). In addition, structural differences exist between the gyri and sulci, with gyri having significantly greater number of neurons in deep layers. The process of cortical folding can lead to differences in cell and dendrite morphology and the layout of cortical blood vessels in the gyri and sulci, thereby potentially resulting in further differences in functioning (Hilgetag & Barbas, 2005). Assessing gyri and sulci may reflect a further way to index neurodevelopmental abnormalities of cortical development in schizophrenia.

In this study we found the cingulate and the superior temporal sulci to be the most affected, and these may serve as useful structural endophenotypes relating the pathology to the aetiology. In addition, our results indicate that thickness, surface area and volume are differentially affected in the sulci and gyri. Combined with our previous gyrical findings, these sulcal results suggest that cortical thickness abnormalities are more prominent in the sulci and that surface area abnormalities are more prominent in the gyri. A number of factors could change the formation of sulci and gyri. Differential growth of cortical regions could displace adjacent gyrical regions. Alternatively, growth of brain regions could affect the trajectories of migrating axons, affecting sulci and gyri formation patterns. Lastly, time course disruption of interhemispheric and intercortical white-matter tracts, especially of the thalamocortical fibre system, might affect sulci and gyri formation patterns, even in distant areas (Welker, 1990). The schizophrenia diathesis most likely has an impact on all these processes, thereby having a complex effect on the neurodevelopment of cortical topography.

**Table 2** Thickness, surface area and grey-matter volume (raw values; mean, s.d.)

<table>
<thead>
<tr>
<th></th>
<th>Thickness (mm)</th>
<th>Surface area (mm²)</th>
<th>Volume (mm³)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Controls</td>
<td>Relatives</td>
<td>Controls</td>
</tr>
<tr>
<td><strong>Frontal sulci</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left superior frontal</td>
<td>2.54 (0.17)</td>
<td>2.56 (0.21)</td>
<td>2697 (409)</td>
</tr>
<tr>
<td>Right superior frontal</td>
<td>2.56 (0.17)</td>
<td>2.57 (0.19)</td>
<td>2165 (340)</td>
</tr>
<tr>
<td>Left inferior frontal</td>
<td>2.42 (0.18)</td>
<td>2.37 (0.22)</td>
<td>1828 (308)</td>
</tr>
<tr>
<td>Right inferior frontal</td>
<td>2.40 (0.15)</td>
<td>2.36 (0.19)</td>
<td>1544 (312)</td>
</tr>
<tr>
<td><strong>Cingulate sulci</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left cingulate</td>
<td>2.71 (0.14)</td>
<td>2.67 (0.22)</td>
<td>2497 (294)</td>
</tr>
<tr>
<td>Right cingulate</td>
<td>2.67 (0.14)</td>
<td>2.70 (0.19)</td>
<td>2672 (379)</td>
</tr>
<tr>
<td><strong>Temporal sulci</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left superior temporal</td>
<td>2.65 (0.15)</td>
<td>2.58 (0.17)</td>
<td>4312 (369)</td>
</tr>
<tr>
<td>Right superior temporal</td>
<td>2.47 (0.18)</td>
<td>2.42 (0.21)</td>
<td>4609 (546)</td>
</tr>
<tr>
<td>Left inferior temporal</td>
<td>2.64 (0.21)</td>
<td>2.60 (0.15)</td>
<td>1043 (273)</td>
</tr>
<tr>
<td>Right inferior temporal</td>
<td>2.56 (0.21)</td>
<td>2.59 (0.25)</td>
<td>900 (233)</td>
</tr>
</tbody>
</table>
REFERENCES


ACKNOWLEDGEMENTS

V.M.G. was supported by a Postgraduate Scholarship Master’s and Doctoral Award from the Natural Sciences and Engineering Research Council of Canada and by the Graduate Research Partnership Program, University of Minnesota. Additional support was provided by the National Institute of Mental Health (MH45156) and the University of Minnesota. The authors gratefully acknowledge the assistance of the International Neuroimaging Consortium: Dr. David Rottenberg, Kirt Schaper, Kristi Boesen and Tim Jarvis; the University of Pittsburgh Medical Center, Department of Radiology; the Cognitive Control Neuroscience Laboratory; the Western Psychiatric Institute and Clinic, University of Pittsburgh; and Theresa Becker, Jill Stanton, Ryan Walter, and Bruce Fischl. We especially thank James Porter for his help with reliability analyses.

Correspondence: Angus W. Macdonald, III, Department of Psychology, University of Minnesota, 2108 Elliott Hall, 75 East River Road, Minneapolis, MN 55455, USA. Tel: +1 612 624 3813; fax: +1 612 625 6668; email: angus@umn.edu

(FIRST RECEIVED 29 JANUARY 2007, FINAL REVISION 29 APRIL 2007, ACCEPTED 9 MAY 2007)
Insight in psychosis: influence of cognitive ability and self-esteem

MICHAEL A. COOKE, EMMANUELLE R. PETERS, KATHRYN E. GREENWOOD, PETER L. FISHER, VEENA KUMARI and ELIZABETH KUIPERS

Background Insight in psychosis has previously been associated with both depression and cognitive ability. Some studies have found a curvilinear relationship between insight and cognitive ability, but the roles of self-esteem and depression have not been taken into account.

Aims To investigate the relationships between insight and IQ, depression, and self-esteem.

Method Correlations between self-reported and observer-rated insight, and measures of IQ, depression and self-esteem were examined in 67 people with psychosis.

Results Better self-reported insight was associated with higher IQ and poorer self-esteem, but not depression. There was some evidence for a curvilinear relationship between IQ and self-reported insight, specifically the ‘awareness of illness’ dimension, which survived correction for symptom variables.

Conclusions The relationship between insight and IQ might reflect both the basis of insight in intellectual ability and the influence of a psychological mechanism that preserves self-esteem.

Declaration of interest None. Funding detailed in Acknowledgements.

Lack of insight is a clinically important phenomenon in psychosis and has predictive value for treatment outcome (Kemp & David, 1996). Understanding the basis of poor insight might improve interventions. Studies of the relationship between insight and cognition have produced inconsistent results. Some have found poorer insight to be associated with lower scores on measures of executive function and IQ whereas others have not (Morgan & David, 2004; Cooke et al, 2005; Aleman et al, 2006). A curvilinear relationship between insight and cognition has also been noted (Startup, 1996): high cognitive ability was associated with both high and low extremes of insight. A meta-analysis also suggests that there is a modest relationship between better insight and greater depressive symptomatology (Mintz et al, 2003). However, few studies have examined the relationships between dimensions of insight and self-esteem, depression, and cognitive ability. The aim of this study was to investigate the relationships between insight and IQ, self-esteem and depression in people with psychosis.

METHOD

Sample

Sixty-seven participants (41 men, 26 women) were recruited and assessed at baseline prior to a randomised controlled trial of cognitive-behavioural therapy for psychosis. All participants were out-patients, had been stable on medication for at least 3 months and had at least one distressing positive symptom scoring three or higher on the Positive and Negative Syndrome Scale (PANSS; Kay, et al, 1987). Good insight was not required for recruitment. The mean number of previous hospital admissions was 2.6 (range 0–20), and the mean duration of illness was 8.1 years (range 0–32). The mean age was 38.1 years (range 23–62). Of the 64 participants for whom data on medication were available, 54 (84.4%) were taking an atypical antipsychotic, 7 (10.9%) a typical antipsychotic, and 3 (4.7%) were not taking any antipsychotic. The overall mean dose was 53.9% of the maximum dose.

Assessments

Insight was assessed using the self-report Insight Scale (Birchwood et al, 1994), which measures three dimensions of insight: awareness of illness (2 items), awareness of symptoms (2 items) and awareness of the need for treatment (4 items). Each sub-scale is given equal weight when calculating the total score. Item 4 (‘My stay in hospital is necessary’) was excluded because all participants were out-patients. The remaining three items from the ‘awareness of the need for treatment’ dimension were used to calculate a score for this sub-scale with equal weight to the other two sub-scales, allowing a total score to be calculated which has the same range (0–12) as the full Insight Scale. Higher scores indicate better insight.

The PANSS was used to assess positive, negative and general symptoms as part of the randomised controlled trial, and the PANSS G12 item was also used as an observer-rated measure of insight. Higher scores on the G12 item indicate poorer insight. Current IQ was measured using the Quick Test (Ammons & Ammons, 1962), a picture vocabulary test consisting of 50 items. Self-esteem was assessed using the Rosenberg Self-Esteem Scale (Rosenberg, 1965), a 10-item self-report questionnaire (higher scores indicate poorer self-esteem). Level of current depressive symptomatology was assessed using the Beck Depression Inventory (BDI; Beck et al, 1961), a 21-item self-report questionnaire. These measures were selected for their previous use in published studies of people with psychosis and their relatively brief administration time as part of an extensive clinical assessment.

Data analysis

Linear relationships between insight scores and other variables were examined using Pearson’s correlations. Following Startup (1996), the possibility of quadratic relationships between insight and IQ was investigated using hierarchical regression analyses in which IQ was treated as the dependent variable and insight score (either Insight Scale or PANSS G12 item) as a predictor variable. The insight score was...
first entered into the regression to test for a linear relationship, followed by the square of the insight score to test for a quadratic relationship.

RESULTS

Descriptive statistics for all measures used in this study are displayed in Table 1. The mean total Insight Scale score of 8.5 was comparable to that of the inpatient sample (at discharge) on which the scale was originally standardised (mean 8.1; Birchwood et al., 1994). Splitting our sample according to the cut-off score of 9 for good insight which was recommended in the original study (Birchwood et al., 1994), 30 individuals (44.8%) would be classified as having poor insight, whereas 37 (55.2%) would be classified as having good insight. This suggests that in our sample insight was slightly lower than in a recent neuropsychological study which utilised the Insight Scale (30% with poor insight; Donohoe et al., 2005). As expected, the total Insight Scale and PANSS G12 measures, which are scored in opposite directions, were highly negatively correlated (r = -0.646, P < 0.001).

Self-reported insight

Total Insight Scale score was correlated with higher IQ (r = 0.264, P < 0.05) and poorer self-esteem (r = 0.382, P = 0.001), but not depression (r = 0.189, P > 0.05). The relationship between insight and self-esteem remained significant after controlling for depression (r = 0.342, P < 0.005), as did that between insight and IQ (r = 0.31, P = 0.01). IQ and self-esteem were not related (r = 0.12, P = 0.3).

In the regression analysis, the linear Insight Scale component was significant (R² change 0.070, F change (1,65) 4.86, P < 0.05), reflecting an association between higher Insight Scale score and higher IQ. The quadratic Insight Scale component was also highly significant (R² change 0.127, F change (1,64) 10.11, P < 0.005). The full model accounted for 19.6% of the variance, with an adjusted R² of 0.17 (F (2,65) = 7.82, P < 0.001). The quadratic Insight Scale component of the model remained significant when positive, negative and general PANSS sub-scale scores, as well as BDI scores, were entered into the model at the first step (R² change 0.163, F change (1,60) 14.63, P < 0.001). There was one IQ outlier more than 1.5 times the interquartile range above the upper quartile, and no Insight Scale outliers. When the analysis was run again excluding this outlier, the quadratic component of the regression remained significant.

Separate regressions for the three dimensions of the Insight Scale indicated that it was the ‘awareness of illness’ dimension which was driving the curvilinear relationship between total Insight Scale score and IQ (Fig. 1), as this was the only dimension for which the quadratic component was significant (R² change 0.181, F change (1,64) 7.763, P < 0.005). Better insight in all three dimensions was significantly associated with poorer self-esteem (P < 0.015 for all), but there were no significant correlations with depression.

Observer-rated insight

The PANSS G12 score was not significantly correlated with IQ, self-esteem or depression. Neither the linear nor the quadratic component of the PANSS G12 regression was significant.

DISCUSSION

Main findings

This study found significant linear correlations between better self-reported insight and both higher IQ and poorer self-esteem, but not depression, in a sample of individuals with psychosis. A curvilinear (quadratic) relationship between self-reported insight and IQ was also found, and survived correction for symptom variables. This relationship was driven by the ‘insight into illness’ dimension. Despite a high concordance between the scales, no correlations with observer-rated insight were found.

It has previously been suggested that the association between insight and IQ reflects the inability of people with low IQs to communicate the presence of insight to the satisfaction of an interviewer (Lewis, 1934; Rossell et al., 2003). This hypothesis is not supported by the present study, which did not find an association between observer-rated insight and IQ.

Table I  Scores on measures of insight, cognitive ability, depression and self-esteem administered to 67 people with psychosis

<table>
<thead>
<tr>
<th>Measure</th>
<th>Mean score</th>
<th>s.d.</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Insight Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>8.5</td>
<td>3.3</td>
<td>0–12</td>
</tr>
<tr>
<td>Awareness of symptoms¹</td>
<td>2.7</td>
<td>1.4</td>
<td>0–4</td>
</tr>
<tr>
<td>Awareness of illness¹</td>
<td>2.7</td>
<td>1.4</td>
<td>0–4</td>
</tr>
<tr>
<td>Awareness of the need for treatment¹</td>
<td>3.1</td>
<td>1.8</td>
<td>0–4</td>
</tr>
<tr>
<td>PANSS G12 insight item</td>
<td>2.7</td>
<td>1.6</td>
<td>1–6</td>
</tr>
<tr>
<td>Quick Test IQ</td>
<td>94.4</td>
<td>13.1</td>
<td>65–135</td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>20.6</td>
<td>11.8</td>
<td>1–54</td>
</tr>
<tr>
<td>Rosenberg Self-Esteem Scale</td>
<td>25.9</td>
<td>6.0</td>
<td>11–38</td>
</tr>
<tr>
<td>PANSS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive symptoms</td>
<td>17.6</td>
<td>4.9</td>
<td>11–32</td>
</tr>
<tr>
<td>Negative symptoms</td>
<td>12.2</td>
<td>5.0</td>
<td>7–29</td>
</tr>
<tr>
<td>General symptoms</td>
<td>31.8</td>
<td>7.0</td>
<td>19–50</td>
</tr>
<tr>
<td>Total</td>
<td>61.3</td>
<td>12.2</td>
<td>39–87</td>
</tr>
</tbody>
</table>

PANSS, Positive and Negative Syndrome Scale.
¹ Weighted sub-scale mean.

Fig. 1  Scatterplot of IQ against Insight Scale scores with fitted quadratic regression curve.
However, the linear association between total Insight Scale score and IQ is consistent with the findings of a number of previous studies (for a review see Cooke et al, 2005), as well as a recent study which utilised the Insight Scale (Donohoe et al, 2005). The significant curvilinear relationship between total Insight Scale score and IQ also supports the hypothesis of Green et al (2000) that there may be two influences on insight. First, high cognitive ability is conducive to, but in itself not sufficient for, having good insight. Second, some individuals may cope with psychosis in a way that promotes their own positive self-evaluation and thus manifest poor insight. The association found in this study between higher Insight Scale scores and poorer self-esteem is consistent with this, and suggests a further psychological mechanism whereby self-esteem is increased at the expense of insight, which may apply to all individuals across the IQ range. The combination of these two factors might explain why some people with high cognitive ability display good insight and others display poor insight, particularly in the ‘awareness of illness’ dimension.

In the only previous study to find a curvilinear relationship between insight and cognitive ability, Startup (1996) found that a composite ‘cognitive deficits’ score derived from factor analysis of neuropsychological test scores explained 56% of the variance in insight. No linear relationship with cognitive ability was found. Although we employed similar statistical methods to Startup (1996), our use of a single cognitive measure and a different insight scale might explain why the amount of variance accounted for by the regression model in this study was modest (20%), and why a linear as well as a curvilinear relationship, was found. Furthermore, although the Quick Test has been used in a number of studies of people with psychosis (e.g. Kondel et al, 2003), some evidence (Mortimer & Bowen, 1999) suggests that it overestimates IQ in this population. This may have influenced the results, particularly the high IQ outliers. Nevertheless, these data support the curvilinear relationship between cognitive ability and insight in a sample of outpatients with psychosis and complement the results from a sample of in-patients investigated by Startup (1996). The findings support the view that the relationship between insight and cognition may be complex, and may reflect an interaction between cognitive abilities and other factors, such as the way a person copes with psychosis. Poor insight, especially in people with good intellectual function, may be adaptive and serve to protect against low self-esteem.

Implications

Self-report measures of insight may be more sensitive to associations with variables relevant to the aetiology of insight than single items from general symptom rating scales. Different dimensions of insight may have different psychological and cognitive correlates.

The finding that better insight is associated with poorer self-esteem, but not greater depression, suggests that insight might be more strongly related to stable, core beliefs about self-worth in people with psychosis than their current level of depressive symptomatology, which may be a result of a number of other factors, such as current social situation. If a person believes that they are mentally well despite disagreements with clinicians (‘poor insight’), this might help to maintain positive core beliefs about the self and promote good self-esteem. Therefore poor insight could be viewed as an adaptive response to a diagnosis of a serious mental illness with respect to psychological well-being, although it can have maladaptive effects on other aspects of functioning, such as engagement with services and taking medication.

Limitations

There are limitations to the generalisability of this study’s sample, since it was composed of out-patients who were recruited to a trial of cognitive–behavioural therapy for psychosis. Although the mean level of insight was comparable with samples reported in other studies using the Insight Scale (Birchwood et al, 1994; Donohoe et al, 2005), our participants might not have been representative in other ways, such as being more engaged with services, more motivated and less cognitively impaired. Detailed information on the psychiatric diagnosis was not available. Moreover, the curvilinear relationship between self-reported insight and cognitive ability was modest and requires replication.

ACKNOWLEDGEMENTS

MC. was supported by a Wellcome Trust Prize Studentship (071298/z/03/z). VK. is supported by a Wellcome Trust Senior Research Fellowship (067427/z/02/z).

REFERENCES


Evaluating the cost-effectiveness of reduced tardive dyskinesia with second-generation antipsychotics

R. A. ROSENHECK

Background  Second-generation antipsychotics may have few advantages over older, cheaper drugs, except for possibly reduced risk of tardive dyskinesia.

Aims  To evaluate the cost-effectiveness of second-generation antipsychotics with regard to reducing tardive dyskinesia.

Method  Literature was reviewed on risk of tardive dyskinesia with second-generation antipsychotics; on severity, duration and impairment of tardive dyskinesia; and on the relationship of this disorder to quality of life and quality-adjusted life-years (QALYs). Diverse cost and benefit assumptions and of 1-year and 5-year planning horizons were examined in a deterministic sensitivity analysis.

Results  Estimating 0.143 QALYs lost per case of severe tardive dyskinesia, 1-year cost-effectiveness estimates for second-generation antipsychotics ranged from £185 000 ($370 000) to £850 000 ($1.7 million) per QALY, and 5-year cumulative estimates ranged from £74 000 ($149 000) to £342 000 ($683 000) per QALY, all above the conventional policy threshold of £25 000 ($50 000).

Conclusions  Reduction of tardive dyskinesia with second-generation antipsychotics appears unlikely to meet standards for cost-effectiveness.

Declaration of interest  R.A.R. has received research support from and/or been a consultant to AstraZeneca, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Janssen Pharmaceutica and Wyeth.

Second-generation antipsychotic medications have emerged as one of the most costly drug classes, with annual spending in the US of £5 billion annually ($10 billion) (Rosack, 2006), or £45 ($90) per US household.

Initial randomised trials, primarily sponsored by industry, suggested that these drugs were superior to first-generation antipsychotics in the treatment of schizophrenia – reducing both symptoms and neurological side-effects such as extrapyramidal symptoms and tardive dyskinesia (Tamminga & Woerner, 2002; Davis et al, 2003; Tarsy & Baldessarini, 2006) as well as reducing healthcare costs (Hamilton et al, 1999). However, more recent studies, conducted by independent investigators, cast doubt on these conclusions, finding virtually no substantial difference in health outcomes, quality of life, neurological side-effects or non-drug health costs between these two classes of drugs (Leucht et al, 2003; Rosenheck et al, 2003, 2006; Lieberman et al, 2005; Jones et al, 2006; Swartz et al, 2007). The earlier studies commonly relied on haloperidol as the comparator – a high-potency antipsychotic often used at higher than recommended dosages (Hugenholtz et al, 2006) and without prophylactic anticholinergic medication.

Using such treatment as a comparator posed an especially high risk of extrapyramidal symptoms, akinetic depression and possibly tardive dyskinesia (Rosenheck, 2005). One literature review found that in studies that used dosages of 12 mg or less of haloperidol (or the equivalent), no benefit in efficacy or overall tolerability was observed for second-generation antipsychotics, although a reduction in extrapyramidal symptoms was noted (Geddes et al, 2000). Two reviews have also suggested that industry-sponsored studies of second-generation antipsychotics are especially likely to favour the manufacturers’ product (Montgomery et al, 2004; Heres et al, 2006), a phenomenon familiar in other areas of medicine as well (Bekelman et al, 2003; Lexchin et al, 2003).

Cost studies have now demonstrated that the higher costs of second-generation antipsychotics increase annual total health costs by £1200–3000 ($2400–6000) per patient (Rosenheck et al, 2003, 2006; Duggan, 2005; Jones et al, 2006). A consensus has also emerged that these newer drugs are more likely than first-generation antipsychotics to cause weight gain, diabetes and metabolic syndrome (American Diabetes Association et al, 2004), although short-term cost consequences may be small (Leslie & Rosenheck, 2005).

The hope remains that second-generation antipsychotics are superior to the earlier drugs on at least one important neurological side-effect: lowering the risk of tardive dyskinesia (Kane, 2006; Casey, 2006; Tandon & Constantine, 2006). A recent meta-analysis of eleven 1-year follow-up studies, including four clinical trials, found a lower annual incidence of tardive dyskinesia with second-generation drugs, but its authors acknowledged that their findings might have been biased by the use of high doses of haloperidol in comparison treatments. Like other recent summaries (Kane, 2006), the outcome of interest was simple incidence of tardive dyskinesia without consideration of severity of disorder, functional impairment, recovery, relationship to quality of life, or longer-term planning horizons. Clinical decision-making and formulary policy must make use of data on these broader health outcomes.

METHOD

We reviewed published research comparing the risks of tardive dyskinesia in treatments with the two classes of antipsychotic among adults with schizophrenia, noting methodological limitations and biases. Data were then derived from diverse published studies along with some previously unpublished data that addressed the expectable ranges of severity and duration of tardive dyskinesia and its relationship to functional capacity and quality of life. Although no single trial has addressed all of these relevant issues, we combined data from several sources to develop an estimate of the potential cost-effectiveness of second-generation antipsychotics in reducing risk of tardive dyskinesia in adults with...
schizophrenia. A range of alternative assumptions and planning horizons were considered in a deterministic sensitivity analysis that examined best- and worst-case scenarios. Unfortunately, available data were insufficient to support a full probabilistic sensitivity analysis of the type recommended for regulatory review (Briggs, 2005; Claxton et al, 2005).

RESULTS

Comparative incidence of tardive dyskinesia with first- and second-generation antipsychotics

The review by Correll et al (2004) summarised data on 1707 patients treated in four randomised trials for an average median duration of 8.8 months, plus 1571 patients treated in seven observational studies. Consistent with older studies (Kane et al, 1984; Chouinard et al, 1986; Glazer et al, 1993); the annual incidence of tardive dyskinesia with first-generation antipsychotics was estimated to be 5.4%. The annual risk with second-generation drugs, in contrast, was estimated at 0.8%, yielding a 4.6% greater attributable risk of this complication with the older drugs. However, further examination of these studies reveals methodological biases that might have led to overestimation of the benefits of the newer antipsychotics.

Of the four randomised trials – the studies that provide the highest quality of evidence – the one that showed the greatest advantage for second-generation antipsychotics compared ziprasidone, which had an annualised incidence of tardive dyskinesia of 6.8%, with placebo, which was reported to have an annualised incidence of this side-effect of 35.7% (Arato et al, 2002). Although tardive dyskinesia occasionally emerges in schizophrenia in the absence of antipsychotic medication (Chakos et al, 1996; Fenton, 2000), the 35.7% annualised incidence rate reported with placebo most probably reflects withdrawal dyskinesia due to residual effects from previous antipsychotic medications, coupled with discontinuation of anticholinergics (Woods, 1999), rather than a true association of placebo and tardive dyskinesia.

The other three randomised trials used haloperidol as the comparator at relatively high doses (13–15 mg), without prophylactic anticholinergics. Using data from the trial with the largest number of participants (Beasley et al, 1999), Correll et al (2004) included only those patients with 6 weeks or more of treatment which left only 36% of the original 1714 participants, a notable attrition from the original randomised sample. In the original publication that included the entire sample (Beasley et al, 1999) 35 of 48 (73%) cases of tardive dyskinesia occurred during the first 6 weeks, further suggesting that withdrawal dyskinesia or extrapyramidal side-effects were not well differentiated from tardive dyskinesia. In a third study, incidence of tardive dyskinesia was based on simple adverse events reporting, an imprecise and unstandardised measure of this complex syndrome (Czernansky et al, 2002).

Although the experimental data supporting reduced risk of this side-effect with second-generation antipsychotics are thus flawed, the conviction that these drugs lower the risk of tardive dyskinesia is not entirely based on such direct evidence. Many have inferred that less frequent occurrence of extrapyramidal symptoms with these drugs results in a lower risk of tardive dyskinesia (Tamminga & Woerner, 2002; Tarsy & Baldessarini, 2006; Tenback et al, 2006). However, a meta-analysis of 31 randomised controlled trials that included 2320 patients found no difference in risk of extrapyramidal symptoms between low-potency first-generation drugs and second-generation drugs other than clozapine (Leucht et al, 2003). Nor has the hypothesis been sustained that anticholinergics, prescribed to prevent extrapyramidal symptoms, themselves increase risk of tardive dyskinesia (Gardos & Cole, 1983). It is thus likely that the reduction in risk of extrapyramidal symptoms as well as of tardive dyskinesia with second-generation antipsychotics has been overstated.

Perhaps because larger numbers of patients have now been exposed to the newer antipsychotics, recent observational studies (Halliday et al, 2002; Lee et al, 2005; Rochon et al, 2005; Woods, personal communication, 2007) as well as randomised trials (Lieberman et al, 2005; Jones et al, 2006) have found rates of extrapyramidal side-effects and tardive dyskinesia that were no greater with first-generation than second-generation drugs. Since evidence for the superiority of the newer antipsychotic (other than the infrequently used drug clozapine) is weak, and contrary evidence is increasing, the 4.6% attributable risk advantage that we use in our analysis is likely to be an optimistic, upper-bound estimate for adults with schizophrenia.

Severity

Severity of illness and functional impairment are central to health outcome assessment but have been largely neglected in research on tardive dyskinesia. The most widely used measure of this disorder, the 10-item Abnormal Involuntary Movements Scale (AIMS; Guy, 1976), rates involuntary movements in seven topographic body zones (e.g. mouth, arms, trunk) and along three global dimensions (objective severity, subjective experience and functional impairment). The definition of tardive dyskinesia used in most studies is a dichotomous measure representing onset of either moderately severe movements in one of the seven body zones or mild movements in two (Schooler & Kane, 1982). Only a few published studies have reported data from the global severity ratings that are included in the AIMS, although severity should clearly be considered in the evaluation of any health outcome. One study of out-patients with first-episode schizophrenia reported that of the 16% of patients who developed tardive dyskinesia, 89% had mild symptoms, 11% moderate symptoms and none severe symptoms of this side-effect (Chakos et al, 1996). A recent European study of 8739 chronically ill out-patients found that of 9.4% with evidence of tardive dyskinesia, 80% had no significant interference in functioning or quality of life from it (Tenback et al, 2006). Thus 80–90% of out-patients with schizophrenia diagnosed with tardive dyskinesia appear to have mild symptoms.

For contrast, we present new data from participants in two large US Department of Veterans Affairs (VA) trials of psychiatric in-patients – patients whose illness is presumably more severe – for whom both baseline and 1-year follow-up data were available. The first trial involved participants with refractory schizophrenia who were receiving in-patient treatment with clozapine or haloperidol (Rosenheck et al, 1997); and the second, in-patients who were assigned to olanzapine or haloperidol (Rosenheck et al, 2003). Since there were robust differences in tardive dyskinesia between treatment groups in the first study, we present data separately for each treatment group. Data from the second study are pooled (Table 1). Over the 1-year follow-up, the prevalence of any tardive dyskinesia (mild to severe) was more than halved among patients taking clozapine, declined only slightly among those taking
Table 1  Distribution of severity of tardive dyskinesia at baseline and at 1-year follow-up in three samples

<table>
<thead>
<tr>
<th></th>
<th>In-patients with refractory schizophrenia</th>
<th>In-patients with non-refractory schizophrenia</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clozapine (n=121)</td>
<td>Haloperidol (n=110)</td>
</tr>
<tr>
<td>Any TD, %</td>
<td>36</td>
<td>28</td>
</tr>
<tr>
<td>Percentage of patients with any TD, %</td>
<td>17</td>
<td>19</td>
</tr>
<tr>
<td>Personal injury</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Mild TD</td>
<td>73</td>
<td>53</td>
</tr>
<tr>
<td>Any patient awareness</td>
<td>78</td>
<td>33</td>
</tr>
<tr>
<td>Moderate–severe TD</td>
<td>27</td>
<td>45</td>
</tr>
<tr>
<td>Moderate–severe subjective distress</td>
<td>30</td>
<td>24</td>
</tr>
<tr>
<td>Any functional impairment</td>
<td>39</td>
<td>16</td>
</tr>
<tr>
<td>Moderate–severe impairment</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

TD, tardive dyskinesia.  

haloperidol in the same trial and was unchanged in the study of patients treated with olanzapine and haloperidol (Table 1, row 1).

Of patients with any tardive dyskinesia, a half to a quarter had moderate–severe ratings initially (Table 1, row 4), and samples that had no change in the overall rate of this condition showed sharp declines in rates of moderate–severe symptoms – by well over a third in the clozapine–haloperidol study and by almost half in the olanzapine–haloperidol study – so that at 12 months 71–87% of those with tardive dyskinesia had mild illness (Table 1). Even smaller proportions of these patients reported moderate to severe subjective distress about their tardive dyskinesia – 17% to 30% across the three samples at baseline – and only 3–16% reported moderate to severe distress 1 year later (Table 1, row 6). Ratings of moderate to severe functional impairment due to tardive dyskinesia affected about 10% of those with this condition at baseline and declined to substantially less than 5% at 1 year (Table 1; row 7). Thus published data, although limited, suggest that only 10–20% of identified tardive dyskinesia among out-patients is rated as more than mild (Chakos et al, 1996; Tenback et al, 2006), and data from hospitalised patients show only a third to a half of those with the disorder to have moderate or more severe ratings, dropping to about a quarter 1 year later. Prevalence of moderate subjective distress or dysfunction was even less frequent, affecting no more than 10%. The fact that tardive dyskinesia can be severe and debilitating in infrequent cases must not be minimised, but a health outcomes perspective must recognise that most tardive dyskinesia is mild and causes limited distress or impairment.

Changing incidence and recovery

In a long-term study of 971 patients with schizophrenia, initiated during the 1970s, the incidence of tardive dyskinesia with first-generation antipsychotic therapy declined with time, from 6.1% in the first year of illness to 2.1% by year 20 (Tammenga & Woerner, 2002). Unfortunately, similar longitudinal data are not available concerning treatment with second-generation antipsychotics.

A recent 1-year clinical trial that involved first-episode schizophrenia found that among those treated with haloperidol, 9% had persistent tardive dyskinesia compared with 5% of those treated with risperidone (P=0.28) (Scholler et al, 2005). However, these persistent cases represented only 19% of all diagnosed tardive dyskinesia, 81% of which lasted less than 3 months, with similar recovery rates for the two drugs (80% v. 82%). In an earlier era, Chakos et al (1996) reported that in first-episode schizophrenia 24% of people experiencing episodes of tardive dyskinesia showed recovery within 3 months, rising to 35% of those with onset in year 2, and then falling to 11% of those with onset in year 4. In the previously cited long-term study of first-generation antipsychotics (Tammenga & Woerner, 2002), 40% of those tardive dyskinesia with onset during the first year of illness recovered within 3 months; this figure dropped to 26% of incident cases in year 5 and to 19% in year 15.

Although estimates of onset and recovery from tardive dyskinesia thus vary widely across studies, it appears that its annual incidence in patients taking conventional antipsychotics drops from 6% to 2% over 20 years, whereas recovery can be as high as 80% in the first year but also declines, to as little as 15%, in later years. Although comparable data are not available for second-generation drugs, evaluation of the health consequences of tardive dyskinesia should consider recovery rates as well as incidence rates.

Treatment of tardive dyskinesia with second-generation drugs

There have been reports that tardive dyskinesia can be effectively treated with second-generation antipsychotics. Although it has long been claimed that, paradoxically, either lowering or increasing the dosage of first-generation drugs can sometimes reduce symptoms of this disorder (Tammenga & Woerner, 2002), recent studies involving clozapine (Tammenga et al, 1994) and olanzapine (Kinon et al, 2004) suggest more complete recovery with the newer drugs. After one sample of patients with tardive dyskinesia were treated for several months with olanzapine, they showed little evidence of recurrence when dosages were subsequently reduced. The 4.6% reduction in risk estimated for second-generation antipsychotics in the
comprehensive review (Correll et al, 2004) does not address the potential benefit of switching to these drugs that may be available to patients who develop tardive dyskinesia while taking a conventional antipsychotic.

It is notable in this respect that the cost estimates we use below for first-generation antipsychotics from the Clinical Anti-psychotic Trials of Intervention Effectiveness (CATIE) (Rosenheck et al, 2006), incorporate the cost of the transition of about half of the patients receiving these drugs to treatment with second-generation antipsychotics, as occurred during the trial. The cost data used in our descriptive cost-effectiveness analysis thus favour the newer drugs because they incorporate the naturalistic cost impact of patients’ switch in treatment. Thus, even though tardive dyskinesia was not the reason for most of the treatment switches in CATIE, our cost estimates include the costs that would be incurred through switching patients from conventional to atypical antipsychotics – switching that would further lower the risk of tardive dyskinesia in real-world practice and that might foster recovery from cases of this side-effect that had occurred. The 4.5% benefit in tardive dyskinesia reduction with second-generation antipsychotics would thus be reduced in real-world practice.

**Tardive dyskinesia and quality of life**

Although assessment of quality of life is an essential feature of health outcomes research, only one small study (n=60) has examined the association of tardive dyskinesia and reduced quality of life. This study found modestly lower quality of life in patients with this condition, but the sample was too small to allow statistical adjustment for the greater severity of schizophrenic symptoms that was also found in these patients and could have been the cause of their lower quality of life (Browne et al, 1996).

To examine this issue further we present more data from the two VA clinical trials referenced above (Rosenheck et al, 1997, 2003). In these analyses average scores on the Heinrichs–Carpenter Quality of Life Interview (QoLI; Heinrichs et al, 1984) were compared across different levels of tardive dyskinesia, as defined by the overall global severity scale of the AIMS. The QoLI is a 20-item, schizophrenia-specific scale that assesses social and instrumental functioning, community activities and intrapsychic (subjective) well-being. Items are scored from 0 to 6 (total score range 0–120), with higher scores indicating better quality of life. Because (as noted above) symptoms of schizophrenia are likely to have a potentially confounding impact on analysis of the relationship between tardive dyskinesia and quality of life, we also analysed these data with multivariate adjustment for the Positive and Negative Syndrome Scale (PANSS; Kay et al, 1987).

Data from the clozapine–haloperidol study (Rosenheck et al, 1997) showed a small but progressive decline in quality of life with more severe tardive dyskinesia (Table 2). Patients with moderate or severe tardive dyskinesia had 12.3% lower QoLI scores than those without this side-effect, and severity of dyskinesia explained less than 1% of QoLI variance. After covarying for schizophrenia symptoms there was no significant difference in QoLI scores across tardive dyskinesia severity levels (P=0.43) (Table 2), but the model explained 36% of the variance. Differences in symptom severity in this sample thus entirely accounted for the modest differences in quality of life associated with tardive dyskinesia.

Data from the VA study of patients with non-refractory disorder (Rosenheck et al, 2003) also showed significantly lower quality-of-life scores with tardive dyskinesia (by 5–9%; Table 2), again explaining only 1% of the variance. After adjusting for symptom severity, differences in QoLI scores with tardive dyskinesia remained statistically significant (P=0.002), but the variance explained by the model increased to 24%. Severity of tardive dyskinesia thus explained only 4.5% of total explained variance in quality of life when symptoms of schizophrenia were also considered.

Although these analyses demonstrate a significant relationship between tardive dyskinesia and a disease-specific measure of quality of life the effect of this disorder was small, with a 5–12% reduction in magnitude of QoLI scores and an \( r^2 \) value of about 1%, suggesting that it explained only 1% of the variance in quality of life, a far smaller proportion than symptoms of schizophrenia.

**Tardive dyskinesia and quality of life in cost-effectiveness analysis**

Cost-effectiveness analysis is based on the use of a common metric for health outcomes that is not specific to any particular illness. The standard measure of quality of life for such analyses, as recommended by the 1996 US Public Health Service Task Force (Gold et al, 1996), is the quality-adjusted life-year (QALY), a year of life rated on a scale from 0 (worst possible health) to 1 (perfect health). These units are assessed with methods in which disease-specific health states are evaluated with universal health equivalents such as the risk of death or years of life lost, and with procedures such as the ‘standard gamble’ or the ‘time trade-off’.

A recent series of studies has estimated QALYs in schizophrenia using a multistep procedure (Lenert et al, 2003, 2004; Mohr et al, 2004). First, on the basis of cluster analysis of a large PANSS dataset, eight empirically derived schizophrenia health states were defined. Next, a group of experts working with professional actors developed video scripts to convey these health states to raters, along with five common drug side-effects, including severe tardive dyskinesia. Using these video presentations, the health states were rated by 620 members of the general public using the standard gamble, the favoured method for measuring QALYs (Gold et al, 1996). In the standard gamble people are queried about the risk of death that they would accept to be cured of each health state, e.g. ‘If you were in this health state, how great a risk of dying would you accept to be fully cured?’ In particular, one video clip, determined by the panel of experts to represent a case of severe tardive dyskinesia, showed a man with relentless jaw and arm movements, who had difficulty speaking and with whom his doctor expresses sympathy but no hope of cure. People imagining that they were in this health state were willing to accept a 14.3% chance of death on average if they could be cured of this problem, a modest effect consistent with the analysis of QoLI data reported above. The resultant QALY rating, determined by subtracting the accepted risk of death from 1.0 (perfect health) would be 0.857. The QALYs derived through this series of videotaped vignettes using this method ranged from 0.44 to 0.88 for the eight schizophrenia health states, whereas QALYs for side-effects ranged from a low of 0.857 for severe tardive dyskinesia to 0.959 for weight gain.

**Cost-effectiveness**

In summarising the diverse array of information presented above, we again follow
the method of calculation recommended by the US Public Health Service Task Force (Gold et al., 1996) for comparing the cost-effectiveness of treatments: the incremental cost-effectiveness ratio (ICER), the ratio of differences in the cost of treatments to differences in their benefits. This ratio reflects how much one would have to pay, on average, for a certain benefit. We apply a range of estimated ICERs in a sensitivity analysis of both-case and worst-case scenarios, to evaluate the consistency of the results across different measures and assumptions based on the diverse data reviewed above.

**Cost differences**

For the incremental cost of second-generation antipsychotics we use three annualised estimates derived from the CATIE study (Rosenheck et al., 2006) (Table 3), a randomised trial of over 1400 patients assigned to receive one first-generation antipsychotic (perphenazine) or one of four second-generation drugs (olanzapine, risperidone, quetiapine or ziprasidone) and followed for 18 months. The three cost estimates include:

(a) a lower-bound estimate of £1200 ($2400) per year based on stable drug cost differences of £100 ($200) per month between perphenazine and second-generation antipsychotics during the last 9 months of the trial, when many patients assigned to perphenazine were taking the newer drugs (Rosenheck et al., 2006);

(b) an intermediate cost estimate of £1700 ($3500) per year (the annualised difference in monthly total health costs between perphenazine and olanzapine, the next least costly treatment, over the entire trial);

(c) an upper-bound cost difference of £3100 ($6200) per year, representing the annualised total health cost difference between perphenazine and quetiapine, the most expensive treatment in the trial.

These annualised cost estimates are similar to those suggested in two sets of published cost estimates presented for application to the CATIE data (Basil et al., 2006; Davis, 2006) as well as to the results of two other clinical trials that compared the costs of first- and second-generation antipsychotics (Rosenheck et al., 2003; Jones et al., 2006). None of the specific trials that evaluated risk of tardive dyskinesia included cost estimates, but we believe these estimated cost differences represent an appropriate range and are consistent with many studies.

### Cost-effectiveness

In our first estimate of cost-effectiveness we used the best-case scenario for second-generation antipsychotics, i.e. the smallest estimate of increased costs and the greatest estimate of health gain. The lowest cost difference from CATIE (£1200 or $2400) and greatest benefit for second-generation drugs (4.6% fewer cases) as suggested by (Correll et al., 2004) yielded an ICER of £26 000 ($52 000) per case of tardive dyskinesia avoided, increasing to £68 000 ($135 000) per case avoided with the higher cost estimates (Table 3, row 2). In terms of cost per QALY, this analysis is the equivalent of assuming that each case of tardive dyskinesia prevented represents a 1.0 QALY gain, i.e. the equivalent of avoiding death and maintaining perfect health for 4.6% of those treated. In our second approximation we used the more realistic published estimate of a 0.143 reduction in QALYs per case of severe tardive dyskinesia (Lenert et al., 2004) rather than the 1.0 QALY estimate used in the first analysis. In this second analysis we assumed that all cases of tardive dyskinesia were severe (in spite of the evidence presented above that this is not the case). In this set of analyses ICERs ranged from £186 000 ($373 000) per QALY to £483 000 ($965 000) per QALY (Table 3, row 3).
Since most tardive dyskinesia is, as we have shown, mild in severity, the third set of analyses estimated—again conservatively—that two-thirds of cases were mild and a third to moderate to severe. We further assumed that those with mild disorder lose half as many QALYs (0.7 QALYs) as those with moderate to severe symptoms (the 0.143 QALY estimate presented above). With two-thirds of cases losing 0.7 QALYs and one-third losing 0.143 QALY, the weighted average loss across the entire population of patients with tardive dyskinesia would be 0.093 QALYs gained per case avoided. Using this severity-adjusted estimate of QALY benefits with second-generation antipsychotics, ICERs were found to increase further, ranging from £280 000 ($561 000) per QALY to £727 000 ($1 453 000) per QALY (Table 3, row 4). Finally, we incorporated the assumption that at least 15% of cases recover and thus included only 85% of second-generation antipsychotic benefits. Estimated ICERs under these assumptions increased further to a range of £330 000 ($660 000) per QALY to £855 000 ($1 700 000) per QALY (Table 3, row 5).

### Extending the planning horizon

The analysis we have presented represents a 1-year time horizon occurring roughly at the mid-point of potentially lifelong schizophrenia. Because longitudinal data are not available on the long-term risk of tardive dyskinesia with second-generation antipsychotic medication, it is not possible to estimate precisely the reduction in risk with these drugs over a lifetime. It is appropriate, however, to focus on a 5-year projection at present, because between 2007 and 2012 patents on both risperidone and olanzapine are likely to expire, bringing substantially lower prices for these drugs.

In a best-case clinical scenario for second-generation antipsychotics the annual attributable risk of 4.6% would accumulate linearly over 5 years. The incremental risk for first-generation drugs is thus 5 x 0.046 = 23% at the end of 5 years, with an average annualised attributable risk of half this magnitude (11.5%). Since costs and risks accrue simultaneously, discounting is not necessary. Assuming that cost differences remain the same over the 5 years, we estimate cost-effectiveness ratios of £75 000 ($149 000) per QALY to £193 000 ($386 000) per QALY, assuming all cases of tardive dyskinesia are severe (Table 4); £112 000 ($224 000) per QALY to £291 000 ($581 000) per QALY, assuming two-thirds of cases are mild; and £132 000 ($264 000) per QALY to £342 000 ($684 000) per QALY assuming 15% of cases recover. This projection is conservative because rates of tardive dyskinesia will be lower than projected among the group taking first-generation drugs because the new-case incidence of this side-effect declines with age, and because we have built into the cost estimates the assumption that half the sample taking first-generation drugs switches to the newer antipsychotics, which will both prevent some cases of tardive dyskinesia and facilitate recovery in some incident cases.

### Table 3 Cost-effectiveness ratio – second-generation vs first-generation antipsychotics: sensitivity analysis

<table>
<thead>
<tr>
<th>Difference in annualised health cost v. perphenazine</th>
<th>CATIE lost cost estimate</th>
<th>CATIE total olanzapine</th>
<th>CATIE total risperidone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UK £</td>
<td>US $</td>
<td>UK £</td>
</tr>
<tr>
<td>Difference in risk of TD cases (Correll et al., 2004)³</td>
<td>0.046</td>
<td>26 087</td>
<td>52 174</td>
</tr>
<tr>
<td>Assume all cases are severe (0.14 QALY loss; Lenert et al., 2004)</td>
<td>0.14</td>
<td>186 335</td>
<td>372 671</td>
</tr>
<tr>
<td>Assume 2/3 mild (QALY loss 0.07), 1/3 moderate/severe (QALY loss 0.14)</td>
<td>0.093</td>
<td>280 505</td>
<td>561 010</td>
</tr>
<tr>
<td>Assume 15% of cases last &lt; 3 months</td>
<td>0.85</td>
<td>330 006</td>
<td>660 012</td>
</tr>
</tbody>
</table>

CATIE, Clinical Antipsychotic Trials of Intervention Effectiveness; QALY, quality-adjusted life-year; TD, tardive dyskinesia.

³ These estimates represent the cost per QALY assuming each case of TD is poor health equivalent to death (QALY loss due to TD = 10).

### Table 4 Cost-effectiveness ratio – second-generation vs first-generation antipsychotics: 5-year projection

<table>
<thead>
<tr>
<th>Difference in annualised health cost v. perphenazine</th>
<th>CATIE lost cost estimate</th>
<th>CATIE total olanzapine</th>
<th>CATIE total risperidone</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UK £</td>
<td>US $</td>
<td>UK £</td>
</tr>
<tr>
<td>Difference in risk of TD cases¹</td>
<td>0.115</td>
<td>10 435</td>
<td>20 870</td>
</tr>
<tr>
<td>Assume all cases are severe (0.143 QALYs lost)</td>
<td>0.14</td>
<td>74 534</td>
<td>149 068</td>
</tr>
<tr>
<td>Assume 2/3 mild (QALY loss 0.07)</td>
<td>0.093</td>
<td>112 202</td>
<td>224 404</td>
</tr>
<tr>
<td>Assume 15% of cases lasts &lt; 3 months</td>
<td>0.85</td>
<td>132 002</td>
<td>264 005</td>
</tr>
</tbody>
</table>

CATIE, Clinical Antipsychotic Trials of Intervention Effectiveness; QALY, quality-adjusted life-year; TD, tardive dyskinesia.

¹ These estimates represent the cost per QALY assuming each case of TD is poor health equivalent to death (QALY loss due to TD = 10).
DISCUSSION

With recent evidence that second-generation antipsychotics may be no more effective and pose no greater risk of Parkinsonian side-effects than first-generation drugs, reduced risk of tardive dyskinesia may be their principal remaining advantage. Although several recent independent trials as well as several observational studies have not found substantial advantages for the newer antipsychotics in reducing the risk of tardive dyskinesia, we based our evaluation on a meta-analysis of older studies that found an annual incidence of this side-effect that was lower by 4.6% with second-generation drugs. We considered important components of health outcome neglected in previous studies of tardive dyskinesia, including severity, duration and quality of life attributable to this condition, and approximated a 5-year planning horizon.

Because available data are limited we used estimates from a variety of sources, in each instance presenting both best-case and worst-case scenarios. Even under the best-case clinical 5-year scenario, the cost per QALY for second-generation antipsychotics based on reducing risk of tardive dyskinesia ranged from £75 000 ($150 000) per QALY to £190 000 ($380 000) per QALY.

The approximate cut-off for reimbursement of drugs in the UK, Australia and Canada has been reported to range from £20 000 ($40 000) per QALY to £25 000 ($50 000) per QALY (Neumann, 2005: pp. 99, 102–3, 105). Thus, reduction of risk of tardive dyskinesia with second-generation antipsychotics as estimated here, even for the best-case scenario, does not seem likely to meet conventional standards for cost-effective treatments. These findings could be an argument for lowering payments for these products to a level at which tardive dyskinesia benefits would be worth the price.

Our study is limited by the small number of randomised trials that have compared the risk of tardive dyskinesia between the two classes of antipsychotics. Data on severity, duration and QALYs of states of this side-effect are also quite limited. The applicability of US cost data to the UK situation is unclear, but cost differences in the US CATIE trial (Rosenheck et al., 2006) were very similar to those of the UK (Cost Utility of the Latest Antipsychotics in Schizophrenia Study (CUtLASS; Jones et al., 2006). We also did not account for diabetes or other metabolic risks, which would result in poorer health and increased costs, further shifting the cost-effectiveness balance against the second-generation drugs.

Our estimates also do not apply to elderly people, in whom studies have shown greater risks of tardive dyskinesia with first-generation antipsychotics (Jeste et al., 1995; Correll et al., 2004), although most older patients do not receive antipsychotic treatment for schizophrenia, the clinical focus of this review. One recent study of older American patients found no net benefit of second-generation drugs even compared with placebo in Alzheimer’s disease (Schneider et al., 2006).

It must be acknowledged that owing to the limited data available and the necessity of estimating outcomes and costs from different studies, this presentation is based on a bounding argument in which we found that even in the best-case scenario for second-generation antipsychotics benefits were unlikely to justify the increased costs by conventional standards. Since we lack information on the probability distribution of the various outcomes, our data are insufficient to support a full probabilistic sensitivity analysis (Briggs, 2005), the standard of cost-effectiveness analysis for policy making (Claxton et al., 2005). However, such analysis may not be as critical in this case as in other situations, because even when we examined the best-case deterministic scenario for the newer antipsychotics, costs per QALY exceeded the generally acceptable threshold. Probabilistic sensitivity analysis, which tests whether our results could have resulted from chance alone, would not be likely to change this conclusion to one favouring these drugs.

There has been considerably more controversy in the USA than in the UK about whether cost-effectiveness analysis should influence clinical practice or healthcare policy (Ubel, 2000; Neumann, 2005). The analyses presented here suggest that, in view of recent studies showing little or no advantage of second-generation antipsychotics for symptoms or extrapyramidal side-effects, reduced risk of tardive dyskinesia does not appear likely to provide sufficient health benefit by itself to justify the predominant use of these agents in treatment of schizophrenia. However, in view of the many limitations noted above, the implications of these findings for either policy or practice must be applied with considerable caution and must ultimately be determined by public debate among relevant stakeholders and governmental agencies (Daniels & Sabin, 2002).

ACKNOWLEDGEMENTS

David Paitel, PhD, provided helpful suggestions on cost-effectiveness methodology.

REFERENCES


DYSKINESIA AND SECOND-GENERATION ANTIPSYCHOTICS

ROBERT A. ROSENHECK, MD, Yale Department of Psychiatry, New Haven, Connecticut, USA

Correspondence: Dr R. Rosenheck, Northeast Program Evaluation Center (182), VAConnnecticut Health Care System, 950 Campbell Avenue, West Haven, CT 06516, USA. Tel: +1 203 937 3850; fax: +1 203 937 3433; e-mail: Robert.Rosenheck@yale.edu

(First received 26 December 2006, final revision 30 March 2007, accepted 19 April 2007)
Treatment of social phobia through pure self-help and therapist-augmented self-help

RONALD M. RAPEE, MAREE J. ABBOTT, ANDREW J. BAILLIE and JONATHAN E. GASTON

Background Self-help for social phobia has not received controlled empirical evaluation.

Aims To evaluate the efficacy of pure self-help through written materials for severe social phobia and self-help augmented by five group sessions with a therapist. These conditions were compared with a waiting-list control and standard, therapist-led group therapy.

Method Participants with severe generalised social phobia (n=224) were randomised to one of four conditions. Assessment included diagnoses, symptoms and life interference at pre-treatment, 12 weeks and at 24 weeks.

Results A larger percentage of patients no longer had a diagnosis of social phobia at post-intervention in the pure self-help group than in the waiting-list group, although this percentage decreased slightly over the next 3 months. Symptoms of social anxiety and life interference did not differ significantly between these groups. Augmented self-help was better than waiting list on all measures and did not differ significantly from group treatment.

Conclusions Self-help augmented by therapist assistance shows promise as a less resource-intensive method for the management of social phobia. Pure self-help shows limited efficacy for this disorder.

Although treatments for social phobia (social anxiety disorder) have shown good efficacy (Gould, et al, 1997; Fedoroff & Taylor, 2001), traditional models of treatment delivery are associated with several limitations. These include their cost, the limited availability of mental health workers and most especially the small percentage of people with this disorder who seek traditional treatment (Meltzer et al, 2000; Issakidis & Andrews, 2002). Self-help and minimal therapist treatments provide a possible alternative to traditional therapy models. Bibliotherapeutic interventions have been applied with success to a wide range of difficulties (Scogin et al, 1990; Marrs, 1995), including anxiety disorders (Newman et al, 2003; Barlow et al, 2005). In contrast to research into other anxiety disorders, there has been little investigation into bibliotherapy for social phobia (e.g. Newman et al, 2003). The few studies that have been conducted have all included some degree of researcher involvement, hence limiting the conclusions that can be drawn about ‘pure’ self-help. Our study was designed to determine the value of two forms of self-help through the use of bibliotherapeutic materials in the reduction of social phobia: pure bibliotherapy that involved almost no contact with the researchers, and therapist-augmented bibliotherapy in which printed material was supplemented with five group sessions conducted by a therapist. Benchmarks for these conditions were provided by comparison with a no-treatment waiting list and standard ten-session group therapy conducted by a therapist.

METHOD

Participants Participants for the study were 224 individuals meeting DSM-IV criteria (American Psychiatric Association, 1994) for social phobia, randomly allocated to one of four treatment conditions: standard group treatment, ‘pure’ self-help, self-help augmented with minimal therapist assistance, and waiting list. Participants were included if they were aged 20–65 years, met criteria for social phobia as their main (or most interfering) disorder, and had sufficient English and education to read a tabloid newspaper in English. In order to maximise external validity, exclusions were kept to a minimum. The only planned exclusions were problems requiring immediate attention such as clear suicidal intent, severe substance misuse or dependence, or florid psychosis, assessed during the structured interviews. Concurrent pharmacotherapy or psychotherapy was allowed as long as dosages had been consistent for 3 months and there was no plan to change. No participant was in concurrent psychotherapy. However, 6.8% were taking benzodiazepines or other anxiolytics, 21.2% were taking selective serotonin reuptake inhibitors or other antidepressants and 9.9% were taking other prescription medications.

Diagnoses of Axis I disorders were made by graduate students in clinical psychology using a structured clinical interview, the Anxiety Disorders Interview Schedule for DSM–IV (ADIS–IV; Di Nardo et al, 1994). Data from our laboratory using this interview and including a proportion of the current sample have indicated a moderate to strong interrater reliability for diagnoses of anxiety and mood disorders, including a very high reliability for a diagnosis of social phobia (κ=0.89). In addition, the avoidant personality disorder questions from the ICD–10 International Personality Disorder Examination (Loranger et al, 1997) were also asked of all participants. Interrater reliability for a diagnosis of avoidant personality disorder was moderate (κ=0.65).

Among the current sample, 95.7% met criteria for the generalised subtype of social phobia and 55.8% met criteria for a diagnosis of avoidant personality disorder. As would be expected in such a severely affected sample, Axis I comorbidity was also high: 42.9% met criteria for an additional anxiety disorder, 33.9% met criteria for an additional mood disorder and 4.0% met criteria for an additional substance use or alcohol disorder. The mean age of the sample was 35.5 years (s.d.=11.0) and 50.4% were female.

Measures Participants were assessed with the following measures at a pre-treatment interview.
and 12 weeks later. Participants in active treatment were also followed up 12 weeks after that (24 weeks after the initial assessment).

Social Interaction Anxiety Scale
and Social Phobia Scale
The Social Interaction Anxiety Scale (SIAS) and the Social Phobia Scale (SPS) (Mattick & Clarke, 1998) are companion scales that assess the main fears and avoidance of social phobia, focusing respectively on interaction fears and more specific performance-based fears. They have excellent psychometric properties (Peters, 2000).

Brief Fear of Negative Evaluation scale
The Brief Fear of Negative Evaluation scale (BFNE; Leary, 1983) assesses the cognitive aspects relevant to social phobia, especially those related to negative evaluation. Psychometric properties are sound and it has shown stronger validity that the previous Fear of Negative Evaluation scale (Rodebaugh et al., 2004).

Albany Panic and Phobia Questionnaire
The Albany Panic and Phobia Questionnaire social phobia sub-scale (APPQ–S; Rapee et al., 1994) is a brief set of items designed to tap social fears that are relatively distinct from overlap with agoraphobic fears. Later examination has shown consistent factor structure, solid reliability and clear concurrent validity (Brown et al., 2005).

Self Consciousness Scale
The Self Consciousness Scale social anxiety sub-scale (SCS–A; Fenigstein et al., 1975) is a six-item scale containing items tapping a variety of broader aspects of shyness and social reticence. It has shown solid psychometric properties in a number of translations and has been widely used in the social anxiety literature.

Life Interference Scale
To provide a measure of the life impact of individuals’ social fears, six Likert scales (scored 0–8) asked respondents to indicate the impact of their fears on various components of their life including work, family life and leisure activities. The scales were summed to provide a total interference rating from 0 (no interference) to 48 (maximum interference). Previous analysis in our centre has shown that the six items show excellent internal consistency ($\alpha=0.90$) and the total correlates significantly with the 12-item Short-Form Health Survey (SF–12; Ware et al., 1996) mental component sub-scale ($r=0.56$).

Treatment conditions
Standard group treatment
Standard treatment was included to represent the ‘gold standard’ treatment effect. Treatment was conducted in groups of approximately six participants with two graduate psychology student therapists who received minimal supervision from an experienced clinical psychologist. Therapy extended for ten 2 h sessions across 12 weeks. Treatment was manualised, based on principles and components described in a book by Rapee & Sanderson (1998). Components included those typically found in empirically validated treatments for social phobia including cognitive restructuring of negative evaluation beliefs, exposure to feared social situations, realistic feedback of social performance, and attention training. Participants engaged in home exercises and received various handouts as relevant.

‘Pure’ self-help
Participants were given a copy of the book Overcoming Shyness and Social Phobia: A Step by Step Guide (Rapee, 1998) and told to read it and work their way through the exercises described in the book. The strategies outlined paralleled those in the standard group treatment, and practice sheets and exercises formed part of the book. In order to encourage a reasonable rate of progress, participants were given a cover letter with the book welcoming them to the programme and providing a suggested rate of progress in order to complete it in 12 weeks. They were told that post-treatment assessment would occur at 12 weeks and thereafter they would have no additional contact with the researchers. This condition was designed to simulate conditions under which a person might obtain written materials without professional assistance (for example, buying the book in a shop or being given a manual while waiting for treatment).

Self-help augmented by therapist assistance
Participants in the augmented self-help condition were given a copy of the same book as those in the pure self-help group and told to read and practise the exercises described. They also met in groups of five to seven participants with a therapist (a graduate psychology student) on five occasions across the 12 weeks. Each group session ran for 2 h; thus, the total therapist time was exactly half of that in the standard group therapy programme. The same therapists participated in this condition and in the standard treatment condition. The aims of the group sessions were to problem-solve application of the principles described in the book to the personal context of each participant and to provide motivation and encouragement to apply these principles.

Waiting list
Participants on the waiting list were told that they had been randomly allocated to receive no treatment for 12 weeks. At the end of the 12-week period they were offered our best available treatment.

Procedure
Potential participants contacted the Macquarie University Anxiety Research Unit through the usual referral sources, including general practitioners, mental health professionals, occasional media coverage and word of mouth. These volunteers were screened by telephone and those who appeared to have social anxiety-related difficulties were invited to attend for a structured interview. Those who met inclusion criteria were randomly allocated to one of the four conditions. Randomisation was done using a pre-assigned random number generator in blocks of eight to allow for group delivery. Participants in the pure self-help group were given a copy of the book and the cover letter and were then simply contacted again after 12 weeks for a second assessment. Participants in the augmented self-help group were given a copy of the book and a schedule of group meetings. Participants in the standard treatment group were simply given a schedule of meeting times. The procedures were approved by the Macquarie University human research ethics committee.

Statistical analysis
Primary outcomes for this trial were a reduction in clinical diagnoses of social phobia as assessed by the ADIS–IV, reduction in a composite of social phobia symptom
measures, and reductions in self-rated life interference.

Following earlier research (Clark et al., 1994), several related symptom measures were grouped together and combined into a standardised composite to reduce the number of statistical tests performed and hence the type 1 error rate. A composite score was produced to represent total social phobia symptom severity. This comprised scores on the clinician-rated severity of social phobia derived from the ADIS-IV, the SPS and SIAS, the social anxiety subscale of the APPQ, the social anxiety subscale of the SCS and the BFNE. Scores on each scale were standardised across all groups on all measurement occasions before being summed and re-standardised to give a mean for all groups and all measurement occasions of 0 and a standard deviation of 1.

The proportions of participants whose phobia was in remission at the post-treatment assessments and 3-month follow-up in each treatment condition were compared using Fisher's exact test. Differences between treatments in change in the standardised social phobia symptom composite and rating of life interference were examined using mixed models containing random intercept and random slope terms as well as fixed effects for treatment received (Gibbons et al., 1993). All analyses were conducted using the Statistical Package for the Social Sciences version 13.0.1 for Windows. Confidence intervals for the number needed to treat were calculated following Altman (1998).

Missing data

The number of participants who provided no data at post-treatment and at 3-month follow-up is shown in Fig. 1. The last value carried forward strategy was used to substitute missing data if data were not available at the 3-month follow-up or at both post-treatment and 3-month follow-up. Interpolation was used if post-treatment data only were not available. As a precaution against biasing effects of these methods of handling missing data, analyses were conducted with and without missing data substituted. Analyses with missing data substituted are equivalent to intent-to-treat analyses. However, analyses without missing data substituted are not equivalent to so-called 'compler' analyses. In most clinical trials completer analyses include only those participants who receive a sufficient 'dose' of treatment (e.g. attend enough treatment sessions); however, this cannot be determined in participants undergoing self-help. Therefore analyses without missing data substitution may include some participants who did not implement any of the self-help, even though they returned data. There was no significant difference in the pre-treatment social phobia symptom composite score between those who provided post-treatment and follow-up data and those who did not (t(222) = −1.135, P > 0.05).

RESULTS

Demographic data on the sample broken down across the four allocated groups are presented in Table 1.

**Diagnosis-free status**

The number and percentage of participants from each treatment condition who no longer met criteria for a diagnosis of social phobia at the post-treatment and follow-up assessments are shown in Table 2. At post-treatment assessment, participants who received active treatments showed significantly greater diagnosis-free rates (group treatment 22%, n = 13; augmented self-help 19%, n = 11; pure self-help 20%, n = 11) than the waiting-list group (6%, n = 3; Fisher's exact test P < 0.008). There was no significant difference in diagnosis-free rates at post-treatment assessment between those who received some form of group therapy (augmented self-help 19%, n = 11; group treatment 22%, n = 13), and those who received pure self-help (20%, n = 11; Fisher's exact test P = 0.522).

At 3-month follow-up there were significantly more participants who no longer met ADIS-IV diagnostic criteria for social phobia in the group treatment and augmented self-help conditions (22%, n = 13; 26%, n = 15, respectively) compared with the self-help condition (11%, n = 6; Fisher's exact test P < 0.05).

The number needed to treat comparing pure self-help with augmented self-help is 7 (1/(0.107 − 0.263) = 6.4) with a 95% confidence interval of 3.4 to 62.8. This indicates that seven patients with social phobia need to be treated with bibliotherapy augmented by face-to-face group interventions before one additional patient achieves a reduction in social phobia over and above that achieved from bibliotherapy alone.

**Change in composite outcome measures**

Changes in the mean standardised composite of social phobia symptom measures and standardised life interference ratings are shown in Table 3, as are the changes from pre-treatment to 24-week follow-up expressed as standardised mean difference effect sizes. In order to facilitate comparison of our data with other research, an additional table has been included that lists means and standard deviations for several of the main outcome measures (Table 4). To maintain a reasonable type 1 error rate, these individual scores were not subjected to independent statistical analyses – they are for descriptive purposes only.

Hierarchical linear or mixed models containing random intercept and slope parameters were fitted to the standardised composite social phobia symptom measures, and to standardised ratings of life

| Table 1 | Characteristics of the four groups of participants |
|---|---|---|---|---|
| Waitinglist | Pure self-help | Self-help augmented with therapist assistance | Group treatment |
| Age, years: mean (s.d.) | 36.2 (11.6) | 36.5 (10.1) | 34.8 (10.1) | 34.8 (12.1) |
| Female, n (%) | 23 (44) | 33 (59) | 27 (47) | 30 (51) |
| Married, n (%) | 15 (29) | 26 (46) | 21 (37) | 12 (20) |
| Post high-school education, n (%) | 32 (63) | 42 (75) | 36 (64) | 43 (74) |
| Avoidant personality disorder, n (%) | 31 (60) | 23 (41) | 35 (61) | 36 (61) |
| Any secondary Axis I diagnosis, n (%) | 33 (66) | 37 (66) | 37 (65) | 38 (64) |
| Prescription medication, n (%) | 20 (39) | 13 (23) | 17 (30) | 20 (34) |
interference due to social anxiety. The random intercept parameter allows for individuals differing within groups on their level of severity, whereas the random slope parameter allows for within-group variance in the rate of change over time. Models with an autoregressive covariance structure for the random slope effect were attempted but did not converge, so results from models with diagonal covariance structure for the random slope are described. Because of a trend toward differences between groups in the pre-treatment diagnosis of avoidant personality disorder ($\chi^2=6.196$, d.f.=3, $P=0.102$), pre-treatment clinician-rated severity of avoidant personality disorder was included as a covariate. The best-fitting model containing random intercept, slope, pre-treatment avoidant personality disorder and treatment effects gave a $-2 \log$ likelihood of 1104.298. (Because around 30% of the sample were taking medication, we tested models that contained medication use as a covariate; there was no improvement in model fit, and hence medication use was not included in the analyses.)

There was a significant group by time interaction on the social phobia composite ($F_{(14,216.9)}=16.131$, $P<0.001$), so planned follow-up tests were conducted to examine specific differences between groups. There was a trend toward differences between the pure self-help and waiting-list interventions post-treatment ($t_{(246.443)}=1.69$, $P=0.093$), and significant differences between waiting list and augmented self-help ($t_{(247.133)}=4.457$, $P<0.001$) and group treatment ($t_{(247.060)}=4.131$, $P<0.001$) also both at post treatment.

At the 24-week follow-up assessment, augmented self-help and group treatment resulted in significantly lower levels of the standardised social phobia composite than the pure self-help condition (augmented $v$. pure self-help $t_{(254.693)}=-3.582$, $P<0.001$; group treatment $v$. pure self-help $t_{(254.120)}=-3.447$, $P<0.001$). There was no significant difference between the augmented self-help and group treatment for this measure at the 24-week assessment ($t_{(254.900)}=0.137$, NS). The same pattern of results was observed when missing data were excluded.

Similar results were observed for ratings of the extent to which social anxiety interfered with a range of activities (total score on the Life Interference Scale). A mixed model containing random intercept and slope terms and including baseline clinician-rated severity of avoidant personality disorder as a covariate was the best fit to the data and gave a $-2 \log$ likelihood of 1368.75. There was a significant group by time interaction ($F_{(14,258.892)}=7.4398$, $P=0.001$) indicating that participants in the four conditions changed at significantly different rates. At the post-treatment assessment, augmented self-help and group treatment led to significantly lower ratings of life interference than the waiting-list control (augmented self-help $v$. waiting list $t_{(234.272)}=-2.577$, $P<0.01$; group treatment $v$. waiting list, $t_{(234.243)}=-2.41$, $P<0.02$) while there was no significant difference between the pure self-help and waiting list groups ($t_{(233.989)}=-0.716$, NS). At 24-week follow-up both augmented self-help and group treatment led to significantly less life interference compared with the pure self-help condition (augmented $v$. pure self-help, $t_{(249.894)}=-2.514$, $P<0.02$; group treatment $v$. pure self-help $t_{(249.671)}=-2.236$, $P<0.05$) with no significant difference between the two interventions involving group therapy ($t_{(249.972)}=0.294$, NS). Again results were consistent when missing data were not substituted.

**Mediation of change in bibliotherapy**

Participants in the two conditions that involved use of the self-help book differed significantly in the number of chapters they reported reading: pure self-help 4.11,
Table 3  Continuous outcome measure scores over time

<table>
<thead>
<tr>
<th></th>
<th>Pre-treatment</th>
<th>Post-treatment</th>
<th>3-month follow-up</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (s.d.)^1</td>
<td>Mean (s.d.)</td>
<td>Mean (s.d.)</td>
<td>Pre to post^2 Pre to follow-up^3</td>
</tr>
<tr>
<td>Waiting list (n=52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardised life interference</td>
<td>0.305 (0.148)</td>
<td>0.226 (0.126)</td>
<td>0.069 (0.130)</td>
<td>0.074</td>
</tr>
<tr>
<td>Standardised social phobia symptom composite</td>
<td>0.445 (0.111)</td>
<td>0.285 (0.100)</td>
<td>0.045 (0.101)</td>
<td>0.200</td>
</tr>
<tr>
<td>Pure self help (n=56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardised life interference</td>
<td>0.397 (0.143)</td>
<td>0.099 (0.123)</td>
<td>0.069 (0.130)</td>
<td>0.278</td>
</tr>
<tr>
<td>Standardised social phobia symptom composite</td>
<td>0.401 (0.107)</td>
<td>0.050 (0.097)</td>
<td>0.045 (0.101)</td>
<td>0.438</td>
</tr>
<tr>
<td>Augmented self-help (n=57)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardised life interference</td>
<td>0.464 (0.141)</td>
<td>0.225 (0.121)</td>
<td>0.309 (0.128)</td>
<td>0.667</td>
</tr>
<tr>
<td>Standardised social phobia symptom composite</td>
<td>0.397 (0.106)</td>
<td>0.330 (0.095)</td>
<td>0.459 (0.099)</td>
<td>0.908</td>
</tr>
<tr>
<td>Group treatment (n=59)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Standardised life interference</td>
<td>0.531 (0.139)</td>
<td>0.191 (0.119)</td>
<td>0.338 (0.126)</td>
<td>0.676</td>
</tr>
<tr>
<td>Standardised social phobia symptom composite</td>
<td>0.443 (0.104)</td>
<td>0.281 (0.094)</td>
<td>0.441 (0.098)</td>
<td>0.906</td>
</tr>
</tbody>
</table>

1. Means are estimated marginal means with the level of the clinician-rated severity of avoidant personality disorder set at the overall pre-treatment mean (2.783) and missing data substituted by the last observed value or the interpolation of adjacent values (described in more detail in the method section).

Table 4  Main social anxiety symptom measure scores

<table>
<thead>
<tr>
<th></th>
<th>Pre-treatment</th>
<th>Post-treatment</th>
<th>3-month follow-up</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (s.d.)</td>
<td>Mean (s.d.)</td>
<td>Mean (s.d.)</td>
<td>Pre to post^4 Pre to follow-up^5</td>
</tr>
<tr>
<td>Waiting list (n=52)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIAS</td>
<td>54.686 (13.679)</td>
<td>54.417 (13.508)</td>
<td></td>
<td>0.020</td>
</tr>
<tr>
<td>SPS</td>
<td>37.529 (15.142)</td>
<td>35.337 (15.830)</td>
<td></td>
<td>0.145</td>
</tr>
<tr>
<td>BFNE</td>
<td>49.823 (7.543)</td>
<td>48.843 (7.528)</td>
<td></td>
<td>0.130</td>
</tr>
<tr>
<td>Pure self-help (n=56)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIAS</td>
<td>51.125 (13.339)</td>
<td>46.057 (17.421)</td>
<td>47.143 (15.953)</td>
<td>0.380</td>
</tr>
<tr>
<td>SPS</td>
<td>36.429 (17.173)</td>
<td>30.962 (17.679)</td>
<td>31.464 (17.676)</td>
<td>0.318</td>
</tr>
<tr>
<td>BFNE</td>
<td>50.125 (8.594)</td>
<td>45.755 (9.589)</td>
<td>47.679 (9.280)</td>
<td>0.509</td>
</tr>
<tr>
<td>Augmented self-help (n=57)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIAS</td>
<td>55.789 (13.227)</td>
<td>43.236 (16.650)</td>
<td>41.596 (16.387)</td>
<td>0.949</td>
</tr>
<tr>
<td>SPS</td>
<td>34.456 (16.250)</td>
<td>24.429 (16.867)</td>
<td>23.491 (17.068)</td>
<td>0.617</td>
</tr>
<tr>
<td>BFNE</td>
<td>51.125 (6.735)</td>
<td>44.625 (9.316)</td>
<td>43.053 (9.420)</td>
<td>0.965</td>
</tr>
<tr>
<td>Group treatment (n=59)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIAS</td>
<td>54.237 (12.343)</td>
<td>44.333 (15.047)</td>
<td>42.203 (15.883)</td>
<td>0.802</td>
</tr>
<tr>
<td>SPS</td>
<td>38.475 (14.536)</td>
<td>28.630 (15.010)</td>
<td>26.034 (15.375)</td>
<td>0.677</td>
</tr>
<tr>
<td>BFNE</td>
<td>51.254 (6.997)</td>
<td>45.000 (8.795)</td>
<td>43.254 (9.325)</td>
<td>0.894</td>
</tr>
</tbody>
</table>

BFNE, Brief Fear of Negative Evaluation; SIAS, Social Interaction Anxiety Scale; SPS, Social Phobia Scale.
1. Means are calculated with missing data substituted by the last observed value or the interpolation of adjacent values (described in more detail in the method section).
2. Difference between the pre-treatment and post-treatment means divided by the pre-treatment standard deviation (the pre-treatment standard deviation was used because missing data substitution might have reduced the variance of later observations and thus unduly inflated effect size estimates).
3. Difference between the pre-treatment and follow-up means divided by the pre-treatment standard deviation.
4. Effect size calculated with the level of the clinician-rated severity of avoidant personality disorder set at the overall pre-treatment mean (2.783) and missing data substituted by the last observed value or the interpolation of adjacent values.
5. Effect size calculated with the level of the clinician-rated severity of avoidant personality disorder set at the overall pre-treatment mean (2.783) and missing data substituted by the last observed value or the interpolation of adjacent values.
the differences in the rate of change between the book only and book plus group interventions.

**DISCUSSION**

**Value of pure self-help**

The major question addressed in this study was whether marked reductions in social phobia could be achieved through self-help delivered in the form of printed material. The results provided mixed support for the value of bibliotherapy in reducing both social fears and the degree of life interference caused by social anxiety. Specifically, the extent of the reductions was markedly influenced by the method of delivering bibliotherapy.

When bibliotherapy was delivered in a 'pure' form – that is, with no significant involvement from a therapist – results were relatively modest. A reasonable proportion of patients no longer met diagnostic criteria for social phobia using pure self-help, although this proportion appeared to be declining by the follow-up point. Changes in symptoms showed a trend to be greater than those of the waiting-list control alone and were maintained reasonably over time, but reductions in life interference were not significantly greater than in the waiting-list group. Hence as a clinical intervention, pure bibliotherapy appears to show limited value for social phobia. However, the most indicative of efficacy (e.g. the moderate effect size change in life interference) suggest that pure bibliotherapy could have a role in population-level interventions or in provision of help to groups who might not have access to extensive mental health services. However, such a suggestion would require more thorough investigation including sample sizes sufficient to detect the small effect sizes that might still have benefits across an entire population.

From a theoretical perspective the modest efficacy of pure self-help for social phobia stands in interesting contrast to the stronger effects shown with many other disorders (Scogin et al., 1990; Marrs, 1995; Newman et al., 2003; Barlow et al., 2005). Social phobia is one of the most chronic of the anxiety disorders (Bruce et al., 2005) and has marked personality-like characteristics (Rapee & Spence, 2004). Hence self-help may be far more difficult to conceptualise and implement for this ego-syntonic condition than for disorders that involve more overt shifts from normal functioning. Our sample was also especially severely phobic and contained a large proportion of people with avoidant personality disorder. It is possible that individuals with more circumscribed forms of social phobia might be more amenable to self-help, although interestingly our data indicated that it was the more severely affected individuals who read more chapters of the book. Finally, the underlying fears in social phobia (e.g. ‘if I make a mistake people will think badly of me’) are typically far less veridical and hence more open to biases in interpretation than many of the concerns in other disorders (e.g. ‘riding on a bus will lead to a heart attack’). This feature may make social phobia less amenable than other disorders to pure self-help. Although the current study provided one of the most valid tests of pure self-help, it is not possible to test a true model of self-help as it would be used in the real world. Specifically, self-help in our study differed from real-world use through the inclusion of pre-treatment assessment and contact, a contact letter, the ‘structure’ of a research trial, and post-treatment assessments. These inclusions might have led to overestimation of the efficacy of pure self-help.

**Augmented self-help**

In contrast to pure self-help, augmentation of self-help with five therapist-led group sessions resulted in marked improvements in symptoms of social phobia and life interference that were as great as those produced by standard group treatment. The lack of a five-session, therapist-only condition does not allow complete conclusions to be drawn about the role of written materials. Although unlikely, it is possible that five group sessions with a therapist might have resulted in equivalent benefit to the augmented bibliotherapy. Nevertheless, this method may provide a template for a highly resource-effective method of treatment delivery. The effect size change in social phobia symptoms produced by augmented self-help (1.08) was larger than the typical effects of cognitive–behavioural therapy shown in meta-analyses (around 0.8) (Fedoroff & Taylor, 2001). Interestingly, a recent treatment for social phobia using internet-delivered self-help combined with some therapist input and in vivo exposure demonstrated an effect size of 0.87 (Andersson et al., 2006). Although treatment based on more recent models of social phobia has shown larger effects, this is accompanied by a markedly increased cost (e.g. Clark et al., 2003). Hence we can begin to flesh out the range of options available to mental health services. At one extreme, expert therapists treating individual patients under detailed supervision can produce extremely efficacious results at a higher cost and limited accessibility. At the other extreme, simple provision of printed materials can produce small changes at extremely low cost and broad accessibility. Augmentation of printed materials with a few therapist-led sessions provides one mid-point alternative. Future research needs to explore further alternatives that might provide the best balance between efficacy and resource use. As an example, John Walker and colleagues (personal communication) have shown good effects from augmenting bibliotherapy with group sessions led by lay facilitators.

**Mechanisms of change**

Further improvements in the efficacy of bibliotherapy could come from research into mediators of change. The results of our study showed that the amount of reading was positively related to outcome. Although this is not surprising, it does imply that identifying methods to increase reading of materials might increase the efficacy of bibliotherapy. Surprisingly, although the use of therapist augmentation was associated with a considerably greater amount of reading, this difference did not explain significant variance in the differences between groups. It appears that therapist augmentation of bibliotherapy provides benefits over and above simple motivation to read the materials. Candidate variables could include better interpretation of procedures, training in additional strategies or more positive outcome expectancies. Several other methods of augmentation have shown promise, including return of weekly homework tasks, ‘check-in’ and reminders through post, telephone, palmtop computers or email. Electronic delivery of self-help is enjoying popularity and may result in some benefits. In many cases internet systems simply consist of written materials in electronic form and will provide no greater benefit than printed materials. However, the use of sophisticated computer programs does allow several interesting features such as individually tailored applications, regular feedback and tracking of progress, and built-in reminders (Griffiths & Christensen, 2006).
Research into the efficacy of bibliotherapy would benefit from systematic examination of predictors. Significant predictors should be used both to screen participants who are most likely to benefit (Baillie & Rapee, 2004) and to inform the development of future modes of delivery.

**Implications**

Mental health services around the world are limited in their reach and scope. In addition, a large proportion of people with anxiety disorders including social phobia do not seek help from traditional mental health services (Meltzer et al., 2000; Issakidis & Andrews, 2002). Many of these people report preferring to deal with difficulties themselves (Issakidis & Andrews, 2002). For these people in particular, self-help might provide an acceptable alternative to traditional therapy. Advantages of self-help include freeing up mental health professionals to allow them to deal with individuals who do require more intensive intervention (Baillie & Rapee, 2004) and providing a more easily accessible and less stigmatising alternative for individuals who are unwilling or unable to access traditional services. Thus, continued investigation into the efficacy of self-help methods can have major implications for public health. Several studies have demonstrated the value of self-help for a variety of anxiety disorders. The current data suggest that pure self-help appears to be less efficacious for social phobia than for other anxiety disorders. Nevertheless, the indications shown here for small effects suggest that larger studies with clearer implications for population health would be of value.

**REFERENCES**


RONALD M. RAPEE, PhD, Department of Psychology, Macquarie University, Sydney; MAREE J. ABBOTT, PhD, Department of Psychology, University of Sydney; ANDREW J. BAILLIE, MPsych, PhD, Department of Psychology, Macquarie University, Sydney; JONATHAN E. GASTON, MClinPsych, Macquarie University Anxiety Research Unit, Macquarie University, Sydney, New South Wales, Australia

Correspondence: Dr Ronald M. Rapee, Department of Psychology, Macquarie University, Sydney, NSW 2109, Australia. Email: Ron.Rapee@mq.edu.au

(First received 29 June 2006, final revision 10 April 2007, accepted 14 May 2007)
Firearm legislation reform in the European Union: impact on firearm availability, firearm suicide and homicide rates in Austria

NESTOR D. KAPUSTA, ELMAR ETZERSDORFER, CHRISTOPH KRALL and GERNOT SONNECK

Background The availability of firearms in homes and at aggregate levels is a risk factor for suicide and homicide. One method of reducing access to suicidal means is the restriction of firearm availability through more stringent legislation.

Aims To evaluate the impact of firearm legislation reform on firearm suicides and homicides as well as on the availability of firearms in Austria.

Method Official statistics on suicides, firearm homicides and firearm licences issued from 1985 to 2005 were examined. To assess the effect of the new firearm law, enacted in 1997, linear regression and Poisson regressions were performed using data from before and after the law reform.

Results The rate of firearm suicides among some age groups, percentage of firearm suicides, as well as the rate of firearm homicides and the rate of firearm licences, significantly decreased after a more stringent firearm law had been implemented.

Conclusions Our findings provide evidence that the introduction of restrictive firearm legislation effectively reduced the rates of firearm suicide and homicide. The decline in firearm-related deaths seems to have been mediated by the legal restriction of firearm availability. Restrictive firearm legislation should be an integral part of national suicide prevention programmes in countries with high firearm suicide rates.

Declaration of interest None.

A recent study (World Health Organization, 2006) has pointed to the possibility of preventing disease by controlling environmental factors; it has been estimated that more than 20% of total suicides in North America and Europe can be attributed to environmental factors, one of which is access to firearms. Suicide prevention measures include the training of primary healthcare personnel, school-based programmes, improvement of the availability of telephone hotlines and crisis centres, implementation of guidelines for the media’s portrayal of suicide, and the restriction of access to methods and means of suicide (Bertolote, 2004). One of the methods of reducing access to suicidal means is the restriction of firearm availability through more stringent firearm legislation (Mann et al, 2005). Generally, studies on the impact of firearm availability and legislation on suicide rates can be divided into three categories: cross-sectional comparisons of firearm availability and suicide rates in different countries; cross-sectional correlations of the stringency of firearm control and suicide rates in the USA; and quasi-experimental studies examining the impact of changes in firearm legislation on suicide rates (Brent, 2001).

The availability of firearms in homes has been shown to be a risk factor for suicide (Kellermann et al, 1992) and homicide (Dahlberg et al, 2004). Furthermore, peer-reviewed publications provide vast evidence to support the view that changes in firearm legislation have influenced the rate of firearm suicides in the USA (Ludwig & Cook, 2000), Canada (Caron, 2004), Australia (Ozanne-Smith et al, 2004) and New Zealand (Beautrais et al, 2006), whereas there is only some evidence from continental Europe and Britain to suggest the same (Hawton et al, 1998; Haw et al, 2004). Other researchers have shown that different state firearm regulations across the USA have influenced suicide (Conner & Zhong, 2003) and homicide rates (Rosengart et al, 2005).

METHOD

We examined suicide and homicide rates in Austria recorded before and after a new firearm law came into effect. We expanded a longitudinal approach from a recent study which evaluated the association between changes in firearm availability and suicide rates in the USA (Miller et al, 2006) and accounted for the availability of firearms. Since the accuracy of survey data and other proxy variables of firearm availability have recently been questioned (Miller et al, 2004), the number of officially issued firearm licences was used to assess the availability of firearms.

Austrian firearm legislation

The Austrian firearms law was adapted in July 1997 after the European Council Directive 91/477/EEC on the control of the acquisition and possession of weapons was implemented in European Union member states. The transposition of the European directive into Austrian national law brought changes concerning the acquisition of firearms over 60 cm in length, which had previously enjoyed a more liberal regime. Moreover, the new law abolished the previous possibility of obtaining a firearm without specifying a reason. In addition to the European Directive, other restrictions were implemented in Austria. The legal criteria for obtaining a category B weapon (handguns, semi-automatic firearms or repeating firearms) for the first time now include psychological testing, being at least 21 years of age, and background checks. Furthermore, the new legislation also specifies safe firearm storage regulations and a 3-day ‘cooling-off’ waiting period for category C and D weapons (including long firearms with smooth bore and rifled barrels) (Commission of the European Communities, 2000).

Data collection

Records for each year between 1985 and 2005 in Austria were examined. The starting year was set at 1985 because of recording restrictions at the Ministry of the Interior, where data on firearm licences were obtained. Data on the number of suicides per year and the number of firearm homicides per year (with external cause-of-death codes according to ICD–9 and ICD–10), as well as data on the size of the general population and on unemployment rates, were obtained from Statistics Austria.
Details on these data have been described elsewhere (Eitersdorfer et al., 2006). The figures on alcohol consumption per capita were obtained from the Austrian Alcohol Coordination and Information Centre (Uhl et al., 2006).

**Statistical analysis**

To assess the effect of the enactment of the 1997 firearm law, a regression of the number of firearm licences per 100,000 inhabitants per year was performed with an autoregressive error model of first order to compare time trends before and after enactment. Similarly, Poisson regressions were performed to compare time trends of firearm suicides, of firearm suicides as a percentage of total suicides and of the total number of homicides, before and after the law was enacted (using SAS/STAT version 8 for Windows). Because of underdispersion, we allowed the variance estimate in both models to depend on an underdispersion factor estimated from the data. Data on total suicides indicated overdispersion, hence total suicides were modelled with negative binomial distribution. The regression model included linear time trends allowing for a change point in 1998. Changing population sizes were taken into account by including the respective changes in the model. To distinguish between the effects of the new legislation and other factors known to influence rates of suicide and homicide, unemployment rates and average alcohol consumption per capita were included in the analysis as covariates. In the calculation of the regression model of homicide rates, the ratio of young men in the population was also included as a covariate. All parameter estimates are reported with 95% confidence intervals. The analysis is based on figures for firearm licences, suicide, firearm suicide, firearm homicide and population sizes in the period 1985–2005. The two-tailed significance level was set at 5%.

**RESULTS**

During the observation period 6071 firearm suicides were counted. Of these, 95.1% were by men and 4.9% by women. When firearm suicides by men are split into age groups, 3.3% of all firearm suicides were by men aged under 19 years, 62.9% by men aged 20–64, and 28.8% by men aged 64 or older. Of all suicides, 0.2% were by females aged 19 years or younger, 3.8% by women aged 20–64 and 0.9% by women aged 64 or older.

Figure 1 shows the trend of the firearm suicide rate, prior to and after enactment of the legislation. We observed no significant time trend in the total number of firearm suicides between 1985 and 1997 ($\chi^2=0.04$, d.f.=18, $P=0.84$). During this period the mean firearm suicide rate was 3.96 per 100,000 (s.d.=0.19). In the period 1998–2005, we observed a significant negative trend ($\chi^2=88.0$, $P<0.0001$) with a steady decline in the firearm suicide rate of 4.7% each year (Table 1). The firearm suicide rate reached a low of 2.67 per 100,000 in the year 2005. The change in the time trend after 1998 is significant ($\chi^2=46.0$, $P<0.0001$) even when adjusted for unemployment and alcohol consumption ($\chi^2=19.9$, $P<0.0001$).

When the firearm suicide rates were analysed by age group and gender, the only significant trend changes after the legislation reform could be found in the group of women aged 20–64 years ($\chi^2=9.9$, $P=0.0016$), in men aged 20–64 years ($\chi^2=81.7$, $P<0.0001$) and in men 65 years old or older ($\chi^2=6.4, P=0.01$). The group of persons aged under 19 years and women more than 64 years old showed no significant trend change in suicide by firearms after enactment of the law. When all ages were pooled for each gender, the trend of firearm suicides after enactment changed significantly among both women and men (women: $\chi^2=12.5, P=0.0004$; men: $\chi^2=28.9, P<0.0001$).

We observed a significant negative time trend in the total number of suicides before 1998 ($\chi^2=131.0$, d.f.=18, $P<0.0001$), as well as in the period 1998–2005 ($\chi^2=23.4, P<0.0001$). The overall suicide rate decreased from 27.6 per 100,000 in 1985 to 16.7 per 100,000 in 2005. The change in the trend was not significant ($\chi^2=0.3, P=0.590$), even when adjusted for unemployment and alcohol consumption ($\chi^2=0.03, P=0.870$).

During the examined pre-legislation period, a significant positive time trend in firearm suicides as a percentage of total suicides was observed ($\chi^2=68.5$, d.f.=18, $P<0.0001$). During this period the percentage of firearm suicides increased from 14.3% in 1985 to 19.3% in 1997. Conversely, during the post-implementation period, a significant negative time trend was observed ($\chi^2=30.5, P<0.0001$), with a decrease to 16.8% in 2005 (Fig. 2). The growth factor of the percentage of firearm suicides per year, derived from Poisson regression analysis, was +2.6% prior to and −2.9% following the change in legislation (Table 1). Consequently, with 1998 set as offset, the change in trend was significant ($\chi^2=53.5, P<0.0001$). Figure 3 shows the course of suicide methods in the examined period. After the implementation of the firearm law, no increase in other methods was observed.

Figure 4 shows the time trend of the total number of firearm homicides, indicating no significant trend before 1998 ($\chi^2=0.04$, d.f.=18, $P=0.840$) with a mean of 0.39 (s.d.=0.08) firearm homicides per 100,000. The growth factor during this period was a modest +0.2% per year. A negative time trend was observed in the post-1998 period ($\chi^2=23.6, P<0.0001$); the growth factor was −2.3% per year. Firearm homicides reached a 20-year low of 0.16 per 100,000 in the year 2005. The change in the firearm homicide trend pre- and post-legislation was significant ($\chi^2=14.3, P=0.0002$). When adjusted for unemployment, alcohol consumption and the proportion of young men in the population, the trend change remained significant ($\chi^2=3.9, P=0.049$) (Table 1).

Finally, we observed a significant positive time trend in firearm licences per 100,000 before 1998 (Fig. 5) with an estimated increase of 140 per year ($t=6.85, P<0.0001$) and a significant negative time trend after the law with an estimated decrease of 125 per year ($t=-3.52,$
IMPACT OF FIREARM LEGISLATION ON SUICIDE

Table 1  Parameter estimates derived from the Poisson regression model.

<table>
<thead>
<tr>
<th>Model 1 (unadjusted)</th>
<th>Model 2 (adjusted)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Growth factor (95% CI)</td>
</tr>
<tr>
<td>Percentage of firearm suicides among all suicides</td>
<td>1.026 (1.019 to 1.032)</td>
</tr>
<tr>
<td>Firearm suicide rate</td>
<td>1.000 (0.994 to 1.010)</td>
</tr>
<tr>
<td>Total suicide rate</td>
<td>0.975 (0.971 to 0.979)</td>
</tr>
<tr>
<td>Firearm homicide rate</td>
<td>1.002 (0.977 to 1.029)</td>
</tr>
<tr>
<td>Percentage of firearm suicides among all suicides</td>
<td>1.026 (1.016 to 1.038)</td>
</tr>
<tr>
<td>Firearm suicide rate</td>
<td>1.00 (0.968 to 1.013)</td>
</tr>
<tr>
<td>Total suicide rate</td>
<td>0.975 (0.969 to 0.982)</td>
</tr>
<tr>
<td>Firearm homicide rate</td>
<td>0.994 (0.894 to 1.10)</td>
</tr>
</tbody>
</table>

1. A growth factor of e.g. 1.05 indicates an increase in the suicide rate of 5% per year.
2. Adjusted for unemployment, per capita alcohol consumption.
3. Adjusted for unemployment, per capita alcohol consumption and proportion of young men in the population.

\(P=0.0026\). Hence, the change in the trend of firearm licence rates is also significant \((t=−5.28, P<0.0001)\). Significant positive autocorrelation \((P=0.0015)\) was observed (Table 1).

DISCUSSION

More stringent firearm legislation has been suggested as an evidence-based suicide prevention strategy (Mann et al., 2005). We examined the effects of a firearm legislation reform in Austria over a period of 12 years prior to and 8 years following its enactment. The results show that the firearm suicide rate decreased among women aged 20–64 years, men aged 20–64 years and men aged 65 years or older; firearm suicides as a percentage of total suicides decreased; the firearm homicide rate decreased; and the overall firearm licence rate decreased after enactment of the new law. These results hold true even when adjusting for common confounders of suicide rates such as unemployment and average alcohol consumption per capita as well as the proportion of young men in the population.

The observed decline in the number of firearm suicides among some age groups after enactment of more stringent legislation is in congruence with previous studies. In the USA, the Brady Handgun Violence Prevention Act has been shown to have reduced suicide rates among people aged 55 years and over (Ludwig & Cook, 2000). In Canada, Leenaars & Lester (1997) observed a decrease in firearm suicides, but this effect was not apparent for those over 65 years old. In Australia a decline in the firearm suicide rate, especially among younger men, was observed in metropolitan and provincial cities after legal restrictions were introduced in 1992 (Cantor & Slater, 1995). The decline in
firearm suicides in Australia was further accelerated after the enactment of a more stringent firearm law in 1996 (Goldney, 2006). Data from New Zealand showed that the mean firearm suicide rate decreased after assessment of firearm reliability and licence tests were introduced in 1992 (Beautrais et al., 2006). Besides this decline in firearm suicides, a decrease in firearm suicides as a percentage of total suicides was also observed, similar to the decrease observed in our findings. The decrease in the percentage of firearm suicides has been discussed as ruling out the possibility that the decline in firearm-related suicides is due to changes in overall suicide rates (Beautrais et al., 2006).

A series of studies is also available from Canada, where the criminal code was amended in 1977 (Bill C-51). Some studies reported a decrease in firearm suicide rates after enactment of the reform (Rich et al., 1990; Carrington & Moyer, 1994; Leenaars et al., 2003). Although Rich et al. (1990) suggested possible switching effects from firearm suicides to suicides by jumping, comparing 5-year periods before and after the reform, a subsequent study (Carrington & Moyer, 1994) with a longer observation period found this trend to be insignificant. The findings on the effects of legal restriction in Canada have been replicated by a further study, after two additional firearm laws (Bills C-17 and C-68) were enacted in 1991 and 1993; subsequent to this legislation, firearm suicides, the percentage of firearm suicides and firearm homicides further decreased (Bridges, 2004). However, when re-examining these Canadian data, Leenaars et al. (2003) found evidence of switching for men but not women. Since firearm suicide rates among women accounted for a small proportion of all firearm suicides in our study, switching effects for each gender were not calculated.

The potential switch to other suicide methods is a complex phenomenon which is often misinterpreted. Since firearm suicides (as a percentage of all suicides) decline after an effective law, the proportion of other methods of suicide logically increases. But this does not explain a possible switching phenomenon – one has to consider the course of suicide rates. Accordingly, a switching phenomenon would be reflected in an increase in absolute suicides by all other methods (or some of the methods) after the enactment of a more stringent firearm law, which was not observed in our work.

Besides cross-sectional studies on the association of firearm availability and firearm suicides in the USA (Dahberg et al., 2004; Miller et al., 2004) and Austria (Etzersdorfer et al., 2006), longitudinal studies have shown that firearm suicides are associated with household firearm ownership (Miller et al., 2002; Haw et al., 2004). Our results support and supplement these findings, additionally demonstrating that a change in firearm legislation has an impact on firearm availability. A change in firearm legislation simultaneously alters the number of firearm licences and the firearm suicide rate, and this emphasises the hypothesised association between both variables.

In a cross-sectional study we have recently shown that there is a trend towards higher firearm suicide rates in counties with higher firearm licensing rates (Etzersdorfer et al., 2006). These findings have been described as representing an urgent public health issue (Leenaars, 2006). Along with the results of the longitudinal analysis reported here, there is even more evidence to support the view that a reduction in firearm suicides and firearm homicides can be achieved by restriction of firearm availability through implementation of more stringent firearms legislation.

Limitations of the study
Because of the limiting ecological design of this study, it must be noted that the decreasing firearm homicide and suicide trends we observed – although having significantly changed in and after the year of the legal reform – may be attributable to socio-economic or other factors. The influence of well-known confounders such as consumption of alcohol per capita and unemployment did not significantly diminish the positive effects of the new law as observed in our study. This has also been found in earlier multivariate models, where the association of firearm availability and firearm suicide rates has not been ruled out (Ludwig & Cook, 2000; Leenaars et al., 2003; Miller et al., 2006).

Since the accuracy of survey data and other proxy variables of firearm availability are questionable (Miller et al., 2004), the number of officially issued firearm licences was used to assess the availability of firearms. We acknowledge that this measure is an underestimation of firearm availability, but it is the only available measure for Austria. Another limitation of our results is that aggregated data analyses allow no direct conclusion about individuals. Miller et al. (2002) discussed the possibility that individuals who die by suicide may not share the characteristics (level of income, available firearm) of the group from which they were drawn. However, this does not limit the growing evidence that more stringent firearm legislation reduces firearm deaths in the overall population.

Finally, it could be argued that the decrease in firearm suicide rates observed in our study might be a randomly found effect of an overall decline in suicide rate and not due to the change in legislation. This is found to be improbable when the decline in post-legislation firearm homicides is also taken into consideration. We could show that, along with a decline in the number
of firearm suicides, the number of firearm homicides also decreased after enactment of the new legislation. The firearm homicide rate trend had previously been steady, but it decreased after enactment, whereas the overall suicide rate was in continuous decline both before and after the legal reform. We argue that the effect of changes in firearm laws can be observed in two different measures: firearm suicides and firearm homicides. Restrictive firearm legislation has previously been shown to reduce firearm suicides as well as homicides in cross-sectional (Conner & Zhong, 2003; Rosengart et al., 2003) and longitudinal studies (Loftin et al., 1991; Leenaars & Lester, 1994; Leenaars et al., 2003; Bridges, 2004).

**Implications of the study**

Suicide prevention plans encounter resistance when they demand more stringent firearm laws, because national firearm restrictions inevitably affect many people, including those who are not at risk of suicide. Yet it should also be remembered that those who are at risk will be more effectively protected by such laws. Some researchers state that suicide prevention strategies are based on two different approaches, which should be balanced – namely, the restriction of the means of suicide and the prevention of mental disorders (Gunnell & Lewis, 2005). They rightly point to the fact that restriction of means does not address the root cause of the problem. However, until the effective prevention of mental disorders realistically becomes our long-term aim, we should recall that the restriction of means very probably does prevent suicide in the short term, thereby increasing the likelihood of people who are suicidal receiving professional help in time. Despite its limitations, this study provides evidence from a European country that firearm suicides and homicides may indeed be prevented through legal restriction of the availability of firearms. Therefore, we recommend that further steps be taken in Austria and in other countries to reduce the availability of firearms.

**REFERENCES**


Caudate volume in offspring of patients with schizophrenia

RAJAPRABHA KARAN RAJARETHINAM, AMEET UPA DHYAYA, PON TSOU, MARGIE UPA DHYAYA and MATCHERI S. KESHAVAN

Summary  Caudate nuclei are smaller in drug-naïve people with schizophrenia but larger in antipsychotic-treated patients. In this magnetic resonance imaging study we found volume reduction of right and left caudate by 8.9 and 8.1% respectively in 50 offspring without psychosis of patients with schizophrenia compared with 53 age- and gender-matched controls, providing new evidence that caudate volume reduction may be a trait-related abnormality in schizophrenia.

Declaration of interest  None.

The caudate nucleus is a major target area for the subcortical dopamine projection system, which is implicated in the pathophysiology of schizophrenia. Caudate and other basal ganglia nuclei help regulate and organise the information flow between frontal lobes and the rest of the brain and play a major part in higher cognitive functions and movement (Middleton & Strick, 1994). Disruptions in this system or lesions of the basal ganglia result in movement disorders and behavioural problems similar to schizophrenia (Heckers, 1997).

Previous studies have shown enlarged caudate in patients treated with dopamine-blocking antipsychotics (Jernigan et al., 1991); however, a reduction has been reported among drug-naïve patients (e.g. Keshavan et al., 1998) but it is not known whether this reduction precedes the illness. At-risk studies have shown brain abnormalities and behaviour deviations supporting neurodevelopmental pathology prior to psychosis (Lawrie et al., 2001; Rajarethinam et al., 2004; Job et al., 2005; Keshavan et al., 2005). We predicted that individuals at risk would have a smaller caudate nucleus than those with no family history of mental illness.

METHOD

Fifty young people (22 males and 28 females, mean age 15.4 years, s.d.=3.6) with at least one parent with schizophrenia and 53 healthy comparison participants (27 males and 28 females, mean age 16.3 years, s.d.=4.4) with no family history of mental illness participated from an ongoing study at the University of Pittsburgh. Other findings from this study have been reported elsewhere (Rajarethinam et al., 2004; Keshavan et al., 2005). The parental diagnosis in the high-risk group was ascertained with the Structured Clinical Interview for DSM-IV (SCID; Spitzer, 1992) and consensus clinical diagnosis. Comparison participants were similar in age, gender and socio-economic and geographical background but with no family history of mental illness. In the high-risk group 26 (57%) had Axis I psychopathology (SCID; Spitzer, 1992) and consensus clinician diagnosis. Seven participants were being treated with stimulants and four with antidepressants, but none with antipsychotics. The University of Pittsburgh institutional review board approved the study. All participants provided written consent; those under 18 years provided informed assent also. Participants aged 15 years or younger were evaluated with the Schedule for Affective Disorders and Schizophrenia for School-Age Children (K–SADS–PL version; Kaufman et al., 1999) and those over 15 were evaluated using the SCID. Anyone with a lifetime history of psychosis, significant neurological or medical illness, current substance use disorder, or any contraindication for magnetic resonance imaging (MRI) was excluded from participating in either group.

For MRI, 124 T1-weighted 1.5 mm coronal slices, without interslice gap, were obtained using a 1.5 T GE scanner (General Electric, Milwaukee, Wisconsin, USA) with three-dimensional spoiled gradient recall acquisition, matrix 256 x 256 x 192, field of view 24 cm, repetition time 25 ms, and time to echo 5 ms. Scans were reviewed to exclude structural abnormalities. The total brain volume was measured by a semi-automatic method using BRAINS2 software (Magnotta et al., 2002) and manually edited for accuracy. The interrater reliability between the four raters ranged from 0.98 to 0.99.

Caudate volumes were measured using the artificial neural network application, a semi-automated tracing method in BRAINS2, with manual editing for accuracy by two trained raters (P.T. and M.U.). Measurement included the head and body of the caudate but not the tail. The test–retest reliability for each rater was established using a set of three scans: (P.T.) right caudate r=0.98, left caudate r=0.99; (M.U.) right caudate r=0.99, left caudate r=0.95. The interrater reliability (intraclass r=0.94 for both right and left caudate) was established using a set of nine scans. Differences in age and intracranial volume were examined by two-tailed unpaired t-tests. Multivariate analysis based on the general linear model was conducted with group status (at-risk v. control) and intracranial volume as predictor variables and caudate volumes as dependent variables.

RESULTS

Age and gender distribution were not significantly different between the groups. The mean total brain volume of the high-risk group (1311.25 cm3, s.d.=130.7) was significantly smaller than that of the comparison group (1391.21 cm3, s.d.=136.3, F=9.223, P<0.003). Analysis of covariance using brain volume as a covariate revealed that the right caudate was significantly smaller in the high-risk group (F=4.014, P<0.05) and the left caudate showed a trend for reduction (F=2.92, P=0.091) (Fig. 1). The right and left caudate nuclei were smaller in the high-risk group by 8.9 and 8.1% respectively (right caudate: control mean volume 3.58 cm3, s.d.=0.51; high-risk mean volume 3.26 cm3, s.d.=0.45; left caudate: control mean volume 3.57 cm3, s.d.=0.51, high-risk mean volume 3.28 cm3, s.d.=0.49). Psychopathology or medication status did not have any significant effect on caudate volumes.

DISCUSSION

Our findings indicate that abnormalities in the caudate nucleus may be seen in young relatives of patients with schizophrenia. Individuals at risk for schizophrenia exhibit
behavioural problems and brain abnormalities, suggesting that some form of the pathological process may begin before the onset of symptoms (Keshavan et al., 2005). However, not all of these at-risk individuals would develop schizophrenia; therefore, the observed neuroanatomical alterations may reflect a measure of familial risk or susceptibility. The conversion to psychosis may result from an interaction between such susceptibility factors and unknown environmental influences or developmental/maturational changes that may involve this system.

Our observations are consistent with studies of basal ganglia function in individuals at risk for schizophrenia. Adult first-degree relatives of schizophrenia patients made more errors on an antisaccade task than a comparison group, suggesting a dysfunction of dorsolateral prefrontal cortex, caudate nucleus, or both (Clementz et al., 1994). Similarly, functional MRI studies have shown decreased activation of the caudate with antisaccade tasks in unaffected relatives (Raemaekers et al., 2006). These findings suggest that an alteration of the structural and functional integrity of corticostriatal neural networks may represent familial or premorbid risk of schizophrenia. It is conceivable that this network may have a role in other neurocognitive deficits such as attentional impairments found in at-risk individuals (Keshavan et al., 2005). In addition, in the context of conflicting research regarding the increase and decrease in caudate volume in relation to treatment or drug-naive status, our data clearly support the hypothesis that volume reduction rather than enlargement of the caudate nucleus is associated with the pathophysiology of schizophrenia. In contrast, Lawrie et al. (2001) reported no difference in the caudate volumes in at-risk relatives (not offspring), some of whom were symptomatic.

To our knowledge, few studies have examined basal ganglia in asymptomatic, untreated, adolescent offspring who are at genetic risk of schizophrenia. High-risk studies enable investigation of neuropathology without the confounds of state-related illness manifestations and medication effects. The neuroanatomical specificity of the observed findings is unclear, but caudate volume reductions might be part of an abnormal corticostriatal network; we and others have found prefrontal and temporal cortical volume deficits in this population (Rajarethnam et al., 2004; Job et al., 2005). The precise mechanisms underlying caudate volume reduction are unclear, and may involve either a failure of normal development or an excessive pruning (Keshavan et al., 2005).

These findings are intriguing, but must be considered preliminary, need replication and may not be generalisable to non-familial forms of schizophrenia. Although the difference was modest, type I errors are unlikely as the sample was large. Our findings support the notion that smaller caudate is a marker of genetic susceptibility, but it is not known whether this abnormality is present at birth or becomes evident during childhood and adolescence. Prospective studies in high-risk individuals suggest that about 10–15% develop schizophrenia, although up to 40% develop schizophrenia-spectrum psychopathology (Erlenmeyer-Kimling et al., 1997). Follow-up of these individuals will help elucidate the role of the caudate in premorbid vulnerability to and later progression into schizophrenia.

ACKNOWLEDGEMENTS

Support from National Institute of Mental Health grants MH 64023, 01180 (M.S.K.), NARSAD independent investigator award (M.S.K.) and GCRC grant M01 RR00056. We thank Diana Dworakowski, MS, and Debra Montrose, PhD, for help with recruitment and assessment, and Jeffrey Nutche, BS, for image processing.

REFERENCES


Familial liability to schizophrenia and premorbid adjustment

MURIEL WALSHE, MARK TAYLOR, KATJA SCHULZE, ELVIRA BRAMON, SOPHIA FRANGOU, DANIEL STAHL, EUGENIA KRAVARIŢI, EILEEN DALY, PAUL FEARON, ROBIN M. MURRAY and COLM MCDONALD

Summary We assessed premorbid functioning during childhood and adolescence in 50 people with schizophrenia from multiply affected families, 39 of their unaffected siblings, 69 people with schizophrenia with no family history of psychosis, 67 of their unaffected siblings and 83 controls. People with schizophrenia had poorer premorbid social and academic adjustment and exhibited a decline between childhood and adolescence compared with controls. Unaffected siblings from multiply affected families also had poor academic functioning in adolescence, with a decline between childhood and adolescence. This may represent a familial (presumed genetic) effect.

Declaration of interest None. E.B. and C.M. supported by the Wellcome Trust.

Cognitive and social deficits pre-date the onset of schizophrenia (Isohanni et al., 2005; Cannon et al., 2006) but it is unclear whether they reflect genetic liability. We examined the early social and academic functioning of individuals from the Maudsley Family Study to determine whether premorbid impairments in these functional domains are related to familial liability for schizophrenia. We hypothesised that individuals with a higher presumed genetic liability (i.e. those from multiply affected families) would display more prominent social and academic impairment than their counterparts from non-affected families.

METHODS

The recruitment and clinical assessments of the sample are described elsewhere (McDonald et al., 2006). Briefly, multiply affected families were defined as having two or more first- and/or second-degree relatives with schizophrenia or another psychotic disorder, and singly affected families were those in which the index individual had no known family history of psychosis as far as their third-degree relatives. None of the control sample had a personal or family history of psychotic illness. All participants were White/Caucasian, aged 18–50 years and gave informed consent for their mother to be interviewed. The study had ethics approval.

Fifty people with 'familial' schizophrenia (37 male, 13 female; mean age 32 years, s.d.=6.1), 39 of their unaffected siblings (14 male, 25 female; mean age 34 years, s.d.=7.8), 69 people with 'non-familial' (52 male, 17 female; mean age 31 years, s.d.=6.4), 67 of their unaffected siblings (34 male, 33 female; mean age 35, s.d.=7.6) and 83 controls (42 male, 41 female; mean age 31 years, s.d.=7.1) were recruited. Patients fulfilled DSM–IV (American Psychiatric Association, 1994) criteria for schizophrenia (n=112), schizoaffective disorder (n=6) or psychotic disorder 'not otherwise specified' (n=1). Eleven (2 'non-familial' and 9 'familial') unaffected siblings had had an earlier DSM–IV Axis I non-psychotic psychiatric disorder, predominantly major depressive disorder.

A modified Premorbid Social Adjustment (PSA) scale (Cannon-Spoor et al., 1982) was used to examine childhood and adolescent functioning (Foerster et al., 1991; Hollis, 2003). The PSA scale assessed five areas (socialisation, peer relations, academic achievement, school adaptation and hobbies) over two consecutive time periods: 5–11 years (childhood) and 12–16 years (adolescence). The PSA scale was then simplified into two categories: social adjustment (socialisation, peer relations and hobbies); and academic adjustment (academic achievement and school adaptation) (Allen et al., 2005). Higher scores indicated poorer functioning. Any developmental deterioration was calculated as a 'change score' by subtracting childhood adjustment from adolescent adjustment.

The scale was administered to the mothers by face-to-face interview (64%) or using a self-report questionnaire (36%). Reliability was established by asking 21 mothers who had completed face-to-face interviews to complete a self-report version of the PSA scale at a later time (on average 4 years). The scores for both time-points were highly consistent (correlation coefficient=0.80).

Multivariate analysis was carried out using STATA version 9.0 with clustered robust standard errors to account for the non-independence of individuals within families and for possible violations of normality and equal variance assumptions. Multiple linear regression was used to compare premorbid adjustment and change scores (dependent variables) of each patient and sibling group (independent variables) with the control group, controlling for age and gender. Scores were log-transformed to normalise the distributions. All tests were two-tailed using a 0.05 level of significance.

RESULTS

The groups did not significantly differ in measures of parental social class (x²=3.5, d.f.=2, P=0.07) or sibship size (F(2, 150)=0.4, P=0.65) but there were significant group differences for age (F(4, 151)=2.8, P=0.03) and gender (x²=22.2, d.f.=4, P=0.002), which were controlled for in subsequent analyses. Siblings of people with non-familial psychosis were older than the control group and there was an excess of males in both patient groups.

Premorbid function scores are presented in Table 1. Compared with controls, both groups with schizophrenia had significantly worse social and academic function in childhood and adolescence, both of which deteriorated over time. The deterioration in social functioning only reached statistical significance for people with 'familial' schizophrenia. In a post hoc analysis directly comparing 'familial' and 'non-familial' schizophrenia, no significant difference was found in either premorbid social functioning or deterioration over time.

Neither unaffected sibling group differed significantly from controls in their social functioning during childhood or...
adolescence (Table 1). However, siblings of people with familial schizophrenia demonstrated significantly worse academic functioning than controls during adolescence, and had a deterioration in academic functioning between childhood and adolescence. Post hoc analysis demonstrated that siblings of people with familial schizophrenia also had a significantly greater decline in academic functioning when compared directly with ‘non-familial’ siblings ($B = -0.31$, $P = 0.02$, 95% CI $-0.56$ to $-0.06$). Analyses were repeated excluding those 11 unaffected siblings with a history of non-psychotic psychiatric disorders but this made no difference to the results.

**DISCUSSION**

To our knowledge this is the first study demonstrating that unaffected siblings of people with familial schizophrenia have poor academic functioning during adolescence and deterioration in academic performance between childhood and adolescence compared with controls. This finding, coupled with its absence in siblings of people with non-familial schizophrenia suggests that academic problems may be related to genetic risk for schizophrenia. This is consistent with reports that adult relatives of people with schizophrenia underperform on cognitive tests compared with controls (Snitz et al., 2006).

This study has some methodological limitations. Separating families on the basis of family history of psychosis runs the risk that some families may be misclassified. People with non-familial schizophrenia may not represent illness phenocopies but multiply affected families are presumed more likely to carry a greater genetic susceptibility load than those families with only one member affected. This is supported by studies which have found more prominent neurobiological deviations among unaffected relatives from more densely affected families (McDonald et al., 2006). It is also possible that recall bias was introduced by the retrospective assessment of behavioural functioning during childhood and adolescence, i.e. that mothers from multiply affected families were more likely to recall negative events in their children. However, maternal ratings across both patient groups (familial and non-familial) were very similar, arguing against such recall bias operating in multiply affected families.

Our finding that people who go on to develop schizophrenia have abnormal premorbid social and academic functioning in childhood and adolescence is in accordance with previous research (e.g. Isahanni et al., 2000), and other studies (Allen et al., 2005) have suggested that academic and social impairment accelerates as people who later develop schizophrenia move from childhood to adolescence.

**REFERENCES**


**Table I** Premorbid Social Adjustment (PSA) scale scores

<table>
<thead>
<tr>
<th></th>
<th>Familial schizophrenia</th>
<th>Sibling – familial schizophrenia</th>
<th>Non-familial schizophrenia</th>
<th>Sibling – non-familial schizophrenia</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=50)</td>
<td>(n=39)</td>
<td>(n=69)</td>
<td>(n=67)</td>
<td>(n=83)</td>
</tr>
<tr>
<td>Childhood social adjustment</td>
<td>4.08 (1.16)*</td>
<td>3.43 (0.64)</td>
<td>4.33 (1.35)**</td>
<td>3.60 (0.97)</td>
<td>3.52 (0.80)</td>
</tr>
<tr>
<td>Adolescent social adjustment</td>
<td>4.36 (1.26)**</td>
<td>3.51 (0.90)</td>
<td>4.43 (1.47)**</td>
<td>3.39 (0.72)</td>
<td>3.47 (0.75)</td>
</tr>
<tr>
<td>Change score for social adjustment</td>
<td>0.28 (0.73)*</td>
<td>0.05 (0.70)</td>
<td>0.10 (0.99)</td>
<td>-0.21 (0.69)</td>
<td>-0.05 (0.62)</td>
</tr>
<tr>
<td>Childhood academic adjustment</td>
<td>3.08 (1.10)**</td>
<td>2.51 (0.68)</td>
<td>3.07 (1.14)**</td>
<td>2.58 (0.86)</td>
<td>2.36 (0.58)</td>
</tr>
<tr>
<td>Adolescent academic adjustment</td>
<td>3.46 (1.42)**</td>
<td>2.84 (0.96)**</td>
<td>3.49 (1.32)**</td>
<td>2.60 (0.87)</td>
<td>2.41 (0.66)</td>
</tr>
<tr>
<td>Change score for academic adjustment</td>
<td>0.38 (0.97)*</td>
<td>0.33 (0.74)**</td>
<td>0.42 (0.99)**</td>
<td>0.01 (0.71)</td>
<td>0.04 (0.44)</td>
</tr>
</tbody>
</table>

*P < 0.05, **P < 0.01, ***P < 0.001 vs controls in linear regression analysis.
Insight, psychopathology and global functioning in schizophrenia in urban Malawi

NIALL CRUMLISH, PRINCE SAMALANI, ANTHONY SEFASI, ANTHONY KINSSELLA, EADBHARD O’CALLAGHAN and HARRIS CHILAILE

**Summary** Insight, psychopathology and functioning are related in schizophrenia, but it is unclear whether insight relates independently to functioning after controlling for psychopathology. Equally, any such relationship may vary culturally. We investigated the relationship between insight, psychopathology and functioning in 60 patients with schizophrenia in Mzuzu, a town in Malawi. After controlling for psychopathology, functioning was associated with the ‘symptom relabelling’ dimension of insight ($P=0.01$). This preliminary finding suggests that symptom-focused psychoeducation might be appropriate for African patients with schizophrenia.

**Declaration of interest** None. Study funded by the Stanley Medical Research Institute.

Insight is a core concept in psychosis, and its relationship to psychopathology and functioning in schizophrenia continues to be examined (Amador & David, 2004). Insight correlates closely with symptom severity (Mintz et al, 2003), but it is unclear whether insight has an association with functional outcome independent of its association with psychotic symptoms.

Most insight research has been carried out in the West, but concepts of mental illness vary across cultures (Saravanan et al, 2004) and the findings of Western insight studies do not automatically apply elsewhere. Although a literature is emerging in Asia (Kim et al, 1997; Tharyan & Saravanan, 2000), the sole African study that examined insight related it only to adherence (Adewuya et al, 2006). To address this deficit, we investigated the relationship of the dimensions of insight to psychopathology and particularly functioning among 60 Malawians with schizophrenia.

**METHOD**

The study centre was the St John of God Community Mental Health Service in Mzuzu. Mzuzu is the largest town in northern Malawi, with a population of over 100,000. Tumbuka is spoken by all Mzuzu residents, and English by most. Christianity, the main religion, coexists with traditional spiritual beliefs. These beliefs influence perceptions of mental illness and of individual symptoms. The traditional explanation for mental illness is ulowi, bewitchment, whereas auditory hallucinations are typically interpreted as the voices of deceased ancestors. Traditional healers, sing’anga, are frequently consulted. The sample comprised the first 60 people with schizophrenia, schizophreniform disorder or schizoaffective disorder recruited to a randomised controlled trial of carer education. We received ethical approval from the National Health Sciences Research Committee, Lilongwe, and obtained written informed consent from participants.

We diagnosed participants using the Structured Clinical Interview for DSM–IV–TR (SCID; First et al, 2002). During this interview we assessed illness duration and the type of treatment sought at onset, and rated functioning with the Global Assessment of Functioning scale (GAF; SCID Axis V). Insight was rated using the Schedule for Assessment of Insight (SAI; David, 1990). The SAI rates three dimensions of insight: treatment adherence (SAI–TA), recognition of illness (SAI–RI) and symptom relabelling (SAI–SR). Symptom relabelling involves the recognition of a psychotic symptom and the understanding that it is a pathological event. The sub-scale totals are summed for a total insight score. To measure psychopathology, we used the Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984) and the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1983).

Two researchers collaborated with four staff members from St John of God Centre on translating instruments into Tumbuka. After independent translations, a consensus meeting was held at which definitive translations were agreed. Adequate interrater reliability ($\kappa>0.8$) was achieved on all measures, post-translation.

We used bivariate correlations to identify relationships between insight, psychopathology and functioning. We entered unadjusted correlates of SAPS, SANS and GAF as independent variables into hierarchical stepwise regression models with SAPS, SANS and GAF as dependent variables. We added total insight to each model before replacing it with each insight dimension in turn. For each dependent variable, the best regression model was that with the largest $R^2$ value.

**RESULTS**

Clinical and demographic characteristics of the sample are shown in Table 1. Of those with a diagnosis other than schizophrenia ($n=6$), 4 had schizophreniform disorder and 2 had schizoaffective disorder.

Scores on SAI–SR were correlated with SAI–TA ($r=0.32$, $P=0.01$). SAI–RI and SAI–SR scores were highly correlated ($r=0.62$, $P<0.001$). SAI–RI and SAI–TA scores were related at a trend level ($r=0.24$, $P=0.06$).

In bivariate correlations, SAPS total score was inversely associated with SAI total score ($r=-0.39$, $P=0.002$), SAPS–SAI–TA ($r=-0.46$, $P<0.001$) and SAI–TA ($r=-0.31$, $P=0.02$). After stepwise regression, correlates of SAPS total score were SAI–SR ($R^2$ change=0.15, $P=0.002$) and duration of illness ($R^2$ change=0.06, $P=0.04$). Longer illness correlated with more severe symptoms. The SANS total score was correlated with SAI–TA ($r=-0.31$, $P=0.02$), and SAI–TA was the only independent variable in the regression model predicting SANS total ($R^2$ change=0.11, $P=0.01$).

In bivariate correlations, GAF correlated positively with total insight ($r=0.47$, $P<0.001$), SAI–SR ($r=0.48$, $P<0.001$), SAI–TA ($r=0.46$, $P<0.001$) and SAI–RI ($r=0.31$, $P=0.02$). The best regression model explained 64% of GAF variance.
It comprised SANS total (R² change = 0.33, P < 0.001), SANS total (R² change = 0.23, P < 0.001), SAI-SR (R² change = 0.05, P = 0.01) and lifetime cannabis misuse (R² change = -0.04, P = 0.01).

**DISCUSSION**

This was a study of the relationship between insight, psychopathology and functioning in an urban Malawian population among whom traditional beliefs were widely held, as indicated by the proportion of patients initially seeking traditional treatments. Low scores on recognition of illness and symptom relabelling may reflect participants’ attribution of illness and individual symptoms to *silowei* or ancestors. Our principal finding was that insight correlated positively with global functioning, independent of confounders. This finding is in keeping with most Western studies (Pini et al., 2001; Lysaker et al., 2007). In finding a positive correlation between symptom relabelling and functioning, we differ with Mutsatsa et al. (2006), but we also differed in methodology. Theirs was a first-episode sample, whereas ours was a prevalence sample, and they measured social functioning, whereas we measured global functioning.

Two explanations offer themselves for the relationship between relabelling and functioning. First, the ability to relabel symptoms may be related to improved cognitive performance, which is itself associated with improved functioning (Morgan & David, 2004). Alternatively, the functional impairment caused by psychotic symptoms may relate to the meaning given to the symptoms as well as to the symptoms themselves. Although our study design cannot show causation, one interpretation could be that psychotic symptoms are functionally harmless regardless of their perceived origin, but it is more harmful to believe that an ancestor is communicating with you than to know that you are experiencing a hallucination. Understanding the origin of psychotic symptoms may ameliorate functional harm, even when the symptoms themselves persist. This preliminary finding suggests that symptom-focused psychoeducation might be appropriate for this population.

The Scale for the Assessment of Negative Symptoms (SANS) is a well-validated instrument used to measure global negative symptoms in psychiatric patients. It includes several subscales, each assessing different dimensions of negative symptoms. The Scale for the Assessment of Illness in Schizophrenia and Related Disorders (SAI) is a structured clinical interview designed to assess insight into illness in schizophrenia and related disorders. It consists of two components: the SAI-SR, which assesses symptom relabelling, and the SAI-TA, which assesses treatment adherence. The Global Assessment of Functioning (GAF) is a scale used to assess overall levels of functioning in psychiatric patients, ranging from 1 to 100, with higher scores indicating better functioning.

---

**Table 1 Characteristics of the sample (n=60)**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Score</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years: mean (s.d.)</td>
<td>33.7 (9.8)</td>
<td>20–65</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>33 (55)</td>
<td></td>
</tr>
<tr>
<td>Years of education completed: mean (s.d.)</td>
<td>9.1 (3.0)</td>
<td>3–16</td>
</tr>
<tr>
<td>Illness duration, years: mean (s.d.)</td>
<td>8.6 (5.9)</td>
<td>0.2–26.0</td>
</tr>
<tr>
<td>Traditional healer consulted at onset, n (%)</td>
<td>23 (38)</td>
<td></td>
</tr>
<tr>
<td>Diagnosis of schizophrenia, n (%)</td>
<td>54 (90)</td>
<td></td>
</tr>
<tr>
<td>Lifetime cannabis misuse, n (%)</td>
<td>12 (20.0)</td>
<td></td>
</tr>
<tr>
<td>Lifetime alcohol misuse, n (%)</td>
<td>10 (17)</td>
<td></td>
</tr>
<tr>
<td>SANS total: mean (s.d.)</td>
<td>3.1 (4.1)</td>
<td>0–14</td>
</tr>
<tr>
<td>SANS total: mean (s.d.)</td>
<td>4.2 (4.7)</td>
<td>0–18</td>
</tr>
<tr>
<td>SAI total: mean (s.d.)</td>
<td>6.8 (4.1)</td>
<td>0–14</td>
</tr>
<tr>
<td>SAI–TA: mean (s.d.)</td>
<td>3.2 (1.1)</td>
<td>0–4</td>
</tr>
<tr>
<td>SAI–RI: mean (s.d.)</td>
<td>2.3 (2.3)</td>
<td>0–6</td>
</tr>
<tr>
<td>SAI–SR: mean (s.d.)</td>
<td>1.4 (1.5)</td>
<td>0–4</td>
</tr>
<tr>
<td>GAF: mean (s.d.)</td>
<td>58.9 (14.5)</td>
<td>24–80</td>
</tr>
</tbody>
</table>

GAF, Global Assessment of Functioning; SAI, Schedule for Assessment of Insight (RI, recognition of illness; SR, symptom relabelling; TA, treatment adherence); SANS, Scale for the Assessment of Negative Symptoms; SAPS, Scale for the Assessment of Positive Symptoms.

**REFERENCES**


Correspondence

EDITED BY KIRIAKOS XENITIDIS and COLIN CAMPBELL

Contents  ▪ Omega-3 fatty acid for recurrent self-harm: unanswered questions ▪
Callous—unemotional traits and autistic psychopathy ▪ Diagnostic stability: clinical v. research ▪ Limitations of cognitive—behavioural therapy for sleep disorders in older adults

Omega-3 fatty acid for recurrent self-harm: unanswered questions

The study by Hallahan et al (2007) has clinically important implications but before accepting the findings as valid we wish to raise a few points regarding some of the methodological and analytical aspects.

Of the 392 patients initially assessed for eligibility, only 39 (10%) completed the study, a large number (343) having been excluded for various reasons. Although this rigorous selection procedure might have enhanced the internal validity of the findings, we are concerned that the generalisability of the findings in the real-world clinical situation (i.e. external validity) might have been compromised.

Certain sample characteristics merit attention. Apart from mentioning that participants had had at least one lifetime self-harm episode in addition to the index episode, the report does not provide any data on the number, frequency, severity and recency of self-harm episodes. These data are important to characterise the sample and to ensure that they did not differ between the two groups. For example, the risk profile of a 60-year-old patient with two self-harm episodes 10 years apart would be very different from that of a 20-year-old with the previous episode only 10 days prior to the index episode. Furthermore, in patients with borderline and other personality disorders, suicidality and impulsivity can vary drastically over time, even in a single day. Instruments rated every 4 or 6 weeks might not capture the ‘real’ picture. Finally, significantly more participants in the placebo group were single or divorced compared with the active drug group. In view of this significant difference, marital status should have been included in the logistic regression and other analyses.

For analysis of suicidality scores the two groups were compared after categorical classification of values (no suicidal ideation v. presence of any suicidal ideation) to obtain a statistically significant difference. For all other variables of interest mean scores were compared. When the mean suicidality scores were compared the difference was not statistically significant. Indeed, it is interesting to note that the proportion of self-harm episodes was actually higher during the study period in the patients on active drug (7/22, 38.2%) compared with those in the placebo group (7/27, 25.9%), although the difference was not statistically significant.

Finally, it is not clear what the findings really mean in terms of decrease in ‘surrogate markers of suicidal behaviour’. Hallahan et al discuss the findings in terms of improved mood and well-being, but the logistic regression analysis showed that depression and other psychological measures did not have any effect on the suicidality score. Other surrogate markers such as impulsivity and aggression scores were not significantly different between the two groups.


D. Basu Postgraduate Institute of Medical Education and Research, Chandigarh, India. Email: db_sm2002@yahoo.com
P. K. Barnwal Postgraduate Institute of Medical Education and Research, Chandigarh, India doi: 10.1192/bjp.191.3.264

Authors’ reply: We thank Basu & Barnwal for their comments. As regards exclusion of so many patients, we stress that easily the biggest reason for exclusion was that the episode of self-harm was the patient’s first. We make it clear why we chose recurrent self-harm rather than all patients with self-harm. The other exclusion criteria seem reasonable (regular fish consumption, etc.) and we see no reason why the findings are not applicable to ‘real-world’ patients. We knew that with such a small population subgroup analysis would be of dubious validity, therefore further defining the groups (e.g. according to recency of other self-harm episodes) was redundant. We certainly could have excluded those patients whose other episode(s) of self-harm were remote from the current one, but we chose not to.

We agree that more measuring points would have been desirable, especially in this capricious sample. This was a resource issue rather than a methodological one. We note the point regarding marital status being different between the two groups but re-analysis of the data controlling for this did not materially affect the results. It was agreed at study outset that in the absence of sufficient power to analyse actual differences in recurrent self-harm we would use the suicidal ideation sub-scale of the OAS–M. One either has suicidal ideation or not (whereas one can have ‘some’ depressed mood) and it seems appropriate to use a categorical measure here.

We suggest using ‘potential marker’ for ‘surrogate marker’ and confess we used the latter word loosely. There was quite good correlation (r=0.5) between measures of depression and the OAS–M suicidality sub-scale score. None the less logistic regression suggested that changes in suicidality were independent of depression scores, which indicates that factors additional to affect drive suicidal ideation. We agree that these findings could be clinically important. However, our findings can be regarded as no more than pilot data, owing to the small sample size. As fish oils are not patentable products, a larger study (with enough power to investigate actual reductions in self-harm) is unlikely to come from industry. Therefore we are continuing to seek funding for such a study.

M. Garland St Ita’s Hospital, Portranee, County Dublin, Ireland. Email: mgarland@ireland.com
B. Hallahan Department of Psychiatry, Beaumont Hospital and the Royal College of Surgeons in Ireland
J. R. Hibbeln Laboratory of Membrane Biochemistry and Biophysics, National Institute on Alcohol Abuse and Alcoholism, Rockville, Maryland, USA
J. M. Davis Institute of Psychiatry, University of Illinois at Chicago, Illinois, USA doi: 10.1192/bjp.191.3.264a
Callosal—unemotional traits and autistic psychopathy

Viding et al (2007) made no reference to autistic psychopathy (Asperger, 1944) nor did any of the other papers in Supplement 49 on assessment risk and outcome in severe personality disorder. The severe unpunitive conduct and aggression problems were well recognised by Asperger (1944) and overlap with what Viding et al (2007) describe as 'more severe, aggressive, and stable pattern of antisocial behaviour and a specific neurocognitive profile indicative of defects in affect processing'. This is precisely what children (and adults) with autistic psychopathy and antisocial behaviour demonstrate (Fitzgerald, 2001, 2003).


M. F. Fitzgerald Department of Child and Adolescent Psychiatry, Trinity College, Dublin, Ireland. Email: Fitz@ol.ie
doi: 10.1192/bjp.191.3.265

Diagnostic stability: clinical v. research

Baca-Garcia et al (2007) highlight some of the important issues related to current nosological systems but other issues need consideration. They voice their concern that with such a high degree of diagnostic instability, the validity of epidemiological, clinical and pharmacological research is questionable. However, in most studies appropriate diagnostic schedules and interviews are used for assessment of patients and a high degree of diagnostic stability has been shown for patients assessed in this manner (Tsuang et al, 1981; Schimmelmann et al, 2005).

Baca-Garcia et al (2007) did not discuss factors such as the level of qualification and number of years of experience in psychiatry of the evaluators, whether the patients were evaluated by the same or different assessors at each visit, the place (i.e. in-patient, out-patient, emergency setting) of first contact, the mean duration of contact, etc., which can influence diagnostic stability. It is also not clear whether at each follow-up proper diagnostic evaluations of patients were performed before diagnosis was recorded.

Furthermore, diagnosis was recorded using ICD–9 codes, but clinicians were using the ICD–10 classification system and this might have lead to errors in conversions and reconversions. Although Baca-Garcia et al reported that clinicians entered one or two diagnoses at the time of evaluation, they have not presented any data regarding comorbidity. Furthermore, when we compare the 'diagnosis received in at least 76% of evaluations' the diagnostic stability in the emergency setting was more than in the out-patient setting for all disorders except eating disorders. This perhaps reflects the likelihood of the evaluators recording the previous diagnosis rather than doing a complete diagnostic evaluation in the emergency setting.

Baca-Garcia et al raise issues which are common in day-to-day practice and highlight the fact that the proper evaluation of the patient requires use of appropriate diagnostic schedules and obtaining information from all possible sources. It is inappropriate to conclude from the study that our diagnostic systems and all research based on this nosological system are flawed.


P. P. Das Postgraduate Institute of Medical Education and Research, Chandigarh, India. Email: drparthapratimdas@yahoo.com

S. Grover Postgraduate Institute of Medical Education and Research, Chandigarh, India
doi: 10.1192/bjp.191.3.265b

Authors’ reply: Asperger’s use of the term psychopathy refers to personality disorder/psychopathology rather than to psychopathy as defined by current criteria. Recent research carried out with colleagues indicates that although there are individuals who have the neurocognitive profile associated with both autistic-spectrum disorders and psychopathy, most individuals with autistic-spectrum disorders (even those with antisocial behaviour) do not show neurocognitive deficits characteristic of psychopathy (Rogers et al, 2006). More importantly, a case review of 177 cases originally diagnosed by Asperger found no raised incidence of criminal offences compared with rates in the general population (Hippler & Klicpera, 2003). It is clear that there are individuals with Asperger’s syndrome/autistic-spectrum disorder who commit crimes (Baron-Cohen, 1988; Scragg & Shah, 1994). However, Asperger’s psychopathy does not equal psychopathy as defined by current practice.


E. M. Viding Department of Psychology, University College London, London, UK. Email: e.viding@ucl.ac.uk
doi: 10.1192/bjp.191.3.265a

Authors’ reply: Our article reports on diagnoses of real patients in the real world and hence variability ranges and the diagnostic process may be affected by factors such as psychiatrist or practice characteristics.

Regarding the question of whether full assessments were performed at each visit, we believe that practitioners tend not to update diagnoses if there is no salient clinical change. We hypothesised that clinicians would be less likely to change diagnoses, biasing the data against our reported finding.

Perhaps the most compelling point is that not all diagnoses were unstable. Thus, it is more likely that our findings reflect inconsistencies in our nosological
system rather than clinician or practice characteristics, or setting effects. For example, some disorders may not always begin with the features required for diagnosis (e.g. mania in bipolar disorder) and therefore diagnostic instability may reflect the time required to consolidate the diagnosis (Baca-Garcia et al., 2007).

Our nosological system is in constant evolution, with major revisions each 15 years. Unfortunately, administrative procedures change more slowly than psychiatrists. Recording from one ICD system to another may affect the validity of diagnoses but not stability, since any error in the conversion of diagnostic codes would likely be constant, given the use of computerised algorithms.

Diagnoses in pharmacological and clinical studies have good internal validity (appropriate diagnostic schedules and interviews). In general, follow-up periods are short and selection bias is likely since participants are selected from specific programmes or units, often based on meeting specific entry criteria. Of note, Peralta et al. (2007) recently reported that the National Hospital Discharge Register was the most reliable means of screening for psychotic and bipolar disorder and was much better than the Composite International Diagnostic Interview (CIDI). They concluded that multiple information sources are key to accurate diagnoses. Studies such as ours, where patients are followed over long periods and across several settings, are closer to this approach than clinical trials based on diagnostic schedules and interviews performed in a research unit over a short period or large cross-sectional epidemiological studies based on a single assessment.


E. Baca-Garcia Department of Psychiatry, Fundación Jimenez Diaz University Hospital, Autonomous University of Madrid, Spain. Email: ebacgar2@yahoo.es

M. M. Perez-Rodriguez Department of Psychiatry, Ramón y Cajal University Hospital, Madrid, Spain

M. A. Oquendo Department of Neuroscience at the New York State Psychiatric Institute and Columbia University, New York, USA. doi: 10.1192/bjp.191.3.265c

Limitations of cognitive–behavioural therapy for sleep disorders in older adults

When the possible side-effects of hypnotics are considered, there is an argument for alternative treatments of sleep disorders in older adults. Sivertsen & Nordhus (2007) emphasised the role of cognitive–behavioural therapy (CBT) in the management of sleep disorders in this population. However, there are also limitations to this approach.

Mental health practitioners or physicians with formal sleep medicine training currently deliver CBT, but they are few in number and could not cater for all that need therapy (Wetzler & Winslow, 2006). This could be the main reason for the prescribing of hypnotics for older adults despite knowledge of their side-effect profile and potential for misuse. Therefore, more workshops are needed for training of mental health professionals in CBT so that they can incorporate these techniques in their routine care of older adults.

There are no clear guidelines about the optimum number and duration of treatment sessions for sleep disorders, particularly for the elderly. It is also unclear how long CBT continues to be effective. Moreover, CBT refers to a number of non-pharmacological treatments for insomnia, but which are the most effective needs more research. There is insufficient evidence to recommend sleep hygiene education, imagery training and cognitive therapy as single therapies or as additions to other specific approaches (Morgenthaler et al., 2006).

Research groups are also working on other effective non-pharmacological interventions for older adults such as acupressure (Chen et al., 1999). Exercise (Montgomery & Dennis, 2004), although not appropriate for all in this population, may also help in inducing sleep. Nevertheless, Sivertsen & Nordhus gave a new insight into this neglected area and provided an impetus for more studies in the elderly.


Authors’ reply: Dr Prakash calls for more training workshops to improve implementation of cognitive–behavioural therapy (CBT) for older adults with sleep disorder. Although we agree that there are too few sleep specialists, we believe that the key to more effective implementation is to provide the same training for other health professionals, including primary care nurses. Although there is no consensus on which component should be included in CBT for insomnia, our experience is that sleep restriction and stimulus control are both crucial for improving sleep in this age group. These components can easily be adapted for use by most health professionals.

In Norway, the Norwegian Medical Association has started to offer training workshops on CBT for insomnia for its members and the Norwegian Psychological Association will soon follow this important initiative.

However, we share Dr Prakash’s concern that there is still insufficient research on how to optimise the treatment and there is clearly a need for studies to determine which component works best and for whom.

B. Sivertsen Department of Clinical Psychology, University of Bergen, Bergen, Norway. Email: Borge.Sivertsen@psyk.uib.no

L. H. Nordhus Department of Clinical Psychology, University of Bergen, Norway. doi: 10.1192/bjp.191.3.266a

O. Prakash Geriatric Clinic and Services, Department of Psychiatry, National Institute of Mental Health and Neurosciences, Bangalore 560029, India. Email: op560029.india.Email: op@nimhans.kar.nic.in doi: 10.1192/bjp.191.3.266


Department of Clinical Department of Clinical


One hundred years ago

Asylum reports – Royal Edinburgh Asylum

Dr. Clouston, after a period of critical hesitation, announces in his Report his entire acceptance of the microbe theory of general paralysis evolved by Dr. Ford Robertson and Dr. M’Rae. Since he has had every opportunity of inspecting the work of these two pathologists on the spot, his adherence to the new belief will be taken as a valuable support to it, for many men who have neither time nor opportunity for examination themselves will be quite content to rely on Dr. Clouston’s deliberate judgment. What a vista the theory opens up! Dr. Clouston with all justice claims credit for the work carried on for some years now by the Scottish Asylums Laboratory. He states that 20 per cent. of his pauper and 10 per cent. of his private admissions are traceable to alcohol, and he asserts that education is the best remedy for the evil. He recommends that children should be taught more of the effects of alcohol as a branch of knowledge that will help in future life. One has only to see the immense good done by voluntary bodies working among the young, such as Bands of Hope, to feel sure that the recommendation of a routine instruction in this matter is absolutely sound. The total of general paralytics admitted is very heavy – 64 in 428, but the fact that females number 38 to 26 of the males must be a record.

REFERENCE

Journal of Mental Science, January 1907, 211.

Researched by Henry Rollin, Emeritus Consultant Psychiatrist, Horton Hospital, Epsom, Surrey
doi: 10.1192/bjp.191.3.267
Choosing ten books that have most influenced my practice is an odd challenge. I wasn’t one of those medical students who wanted to do psychiatry since he was 12, or read most of Freud or Jung. In fact throughout medical school, and for several years as a junior doctor, I thought psychiatry and psychiatrists were pretty weird. It was only in my general practice vocational training year that I realised how much of medicine concerned the psychological and began my psychiatric training. This influence came not through books that I read but through the patients whom I saw.

So what is this review for? Is it to recommend some good reading, help others improve their practice, show how erudite I am, reveal something of myself or all of the above? I’m not quite sure. I read book reviews out of curiosity because they show me how people think. A good review tells us as much about the reviewer as the book. That’s why periodicals with dull names and even duller designs, like the London Review of Books, survive, despite Arts Council subsidies. However, as I imagine it was for those who have gone before me, selecting these ten books was a challenge. It requires a degree of introspection that might be fascinating for me but risks being pretty dull for you. Some works influence our practice directly by their wise words on cognition, psychopathology, sociology or statistics, while others change us as people. I will give equal weight to both.

Sex and sexuality
I suppose no better place to start is with James Baldwin’s Giovanni’s Room. If ever there was a tale of ambivalence, depression and longing, this is it; a young American in Paris struggling against his passionate feelings for an Italian who is free and unashamed of his sexual desires. Such ambivalence was common when the book appeared in Cold War America but even today fundamentalist believers in the Christian, Jewish and Muslim faiths seem to be more exercised by homosexuality than climate change, civil and military violence, poverty or the threat of global epidemics. Our apocryphal visitor from Mars can only watch and wonder. I stumbled upon this novel on the top shelf of a bookshop in a tiny, 1970s New Zealand town that boasted more sheep than people. Courageously published in 1956 by a Black American, it became the impetus for my first faltering steps towards openness and insight. Facing the disapproval of family and society, not through brave rebellion but because there was no choice, has enabled me to see the world from the outside – if one can ever really do that. It has made me forever alert to hypocrisy and rejection and has been invaluable in my understanding of the fear and struggle in the lives of people who consult me. I am grateful that I found the courage to face the prim shop assistant and buy it.

What can we know?
No psychiatrist’s consideration of ten books that influenced them can fail to mention Karl Jaspers’ General Psychopathology. However, I include it not because I was mesmerised by my training by its lucid description of the phenomenology of mental illness; I wasn’t. I discovered the real gems later, where Jaspers focuses on meaning and belief. His painstaking approach to the spiritual in the context of the psychological is profound and he manages to pull it off without the usual romantic or post-modern notions that so often cloud the subject.

Pretty interesting, if you’re interested. And may I remind any trainee, who at this point thinks I’m getting a bit dull, that Jaspers also wrote some liberal (for its time) stuff about masturbation and even speculated about the psychological origins of orgiastic cults. Excellent bedtime reading.

In similar vein Kant’s philosophy, particularly that found in the Critique of Pure Reason and the Critique of Practical Reason (Wood, 2001), has profoundly affected the way I think. His painstakingly obsessive digging down through what and how we can perceive and know has influenced me over the years. In particular, I find his ideas on morality, means and universal ends complex and yet fascinating. It is interesting to see how his concepts of morality, as rational precepts that can arise through reflection in any person, provided they have universal application, might be a springboard for 20th-century post-modern thought, although Kant would be amazed to see where it has led.

Meaning
I practise cognitive–behavioural therapy but in every patient I encounter, no matter how smart or sophisticated, I see a longing for meaning and purpose. Although cognition is inevitably the way we make sense of our perceptions, cognitive therapy often disappoints when it is rigidly applied and fails to grasp our patients’ struggle to find meaning. Some years ago Viktor Frankl’s Man’s Search for Meaning had a powerful effect on me. Although I do not subscribe to the school of existential psychotherapy that arose from this young psychiatrist’s reflection on his imprisonment in Auschwitz, it was clear to me that man’s search for meaning in the most deprived and cruel conditions imaginable reveals much about survival of the human spirit. It doesn’t give me answers to patients’ questions but it makes me see where they’re asking them.

Hard on its heels in the yearning for meaning stakes comes George Eliot’s much neglected last novel Daniel Deronda. This story of 19th-century English attitudes to Jews is both instructive and moving. Eliot draws a memorable portrait of Mordecai the Jew (a somewhat secondary character) who sees in Deronda (the main character, who doesn’t know he is Jewish for most of the novel) the fulfilment of his spiritual desire, the ultimate soulmate with whom...
he can discuss meaning and fulfilment. Yearning for the ‘one’ who will explain and fulfil is likely to disappoint in the end. However, reading Deronda offers an antidote to the pervasive reductionism of psychiatry that, although greatly increasing our knowledge, be it of neuroscience or epidemiology, is ultimately sterile stuff when we encounter people in distress.

**Epidemiology**

Like many academics, I don’t tend to read books about my own subject, epidemiology, as papers are always more salient and topical. One writer in particular, however, who is endlessly entertaining, as well as informative, is the statistician David Steiner – and it is no small task to make statistics funny. His book written with Geoffrey Norman, *Biosistics: The Bare Essentials*, is a classic. Like many epidemiologists I actually like statistics and, like most nerds, can even read books about it on holiday. So, although this will not generally appear in the *Observer’s* list of books for summertime reading, it will keep you guessing, laughs at our academic pomposity and is a great resource for epidemiologists in the making. It combines humour, humility and scholarship in equal measure and is well worth going back to again and again, no matter how sophisticated statistically we think we have become.

**Realism**

Sometimes when we look back it seems that everything is serendipity. We comprehend as we hear birdsong on the sunny morning of a funeral of a loved one that the world is indifferent to our fears, hopes and plans. At other times, however, like the celebrity who has started to believe he really does matter, our lives seem part of a lofty Hegelian roll of history. Thinking like that inevitably drew me to the stuff of realism, and fads followed, first for Thomas Hardy but later for Emile Zola, Gustave Flaubert and George Eliot. There is one of this kind, however, that stands out perhaps from the rest, at least for a psychiatrist, with its chilling description of that malady which has done its best to devastate so many families – alcoholism. *L’Assommoir* by Emile Zola relates the life of Gervaise who, abandoned by her lover and father of her two children, marries a man who descends into alcoholism before Gervaise descends into the same state herself. It reveals the playful, hideous face of alcohol as it charms and harms its way into people’s lives and, even worse, those who are close to them. Zola was criticised for his stereotypes of feckless working-class drunks. However, his depiction of medical treatment of alcoholism in 19th-century Paris and the dissipation wreaked by alcohol among the poor and those about to be poor is a must for aspiring psychiatrists who, as they look under metaphorical stones, will find alcohol addiction everywhere.

**After theory**

Like every junior doctor who wants to specialise in the discipline of his current post, I am in danger of overemphasising something read recently. However, I can’t resist including Terry Eagleton’s *After Theory* in my ten best. Eagleton’s fluency and wit, as well as his ability to stand back and look at the (often rather silly) invasion of cultural theory into almost every academic subject (including psychiatry), makes this book a page turner. Instead of investing effort into things like a ‘structuralist reading of Popeye the Sailorman’ Eagleton urges academics and others towards a return to essential truths about such things as love, morality and death. There is a lot to read and ponder on in this short book, particularly for psychiatrists interested in the mental health of particular groups variously defined by their race, age or sexuality.

**Self-help**

I want to include a book that was prominent on the bookshelves of my father’s generation but is still irrepressible today. In fact, it was recently revived by BBC Radio 4’s ‘Book of the Week’. Dale Carnegie’s *How to Win Friends and Influence People* deserves a second look. If you can put aside its focus of getting on in business, it offers insights into the human mind that have not been surpassed by the abundance of self-help books we see in bookshops now. That way we think determines our happiness and fulfillment is an age-old idea contained in many philosophical and theological systems that have been rediscovered and expanded in cognitive–behavioural therapy today. Carnegie emphasises the self-absorbed nature of all of us but most particularly those with what we now call common mental disorder. He shows how concern for others, resilience in times of stress and exposure to what we fear (be they thoughts or things) are the keys to overcoming it. It calls for less of ‘me, me, me’ and more of ‘you, you, you’, an approach that, as mental health professionals, we are prone to regard as trivial in contrast to rich, narrative-based explanations for mental distress. In fact, I am always intrigued by my patients’ explanations for their depressive or anxiety states (and often their psychoses). This or that event in childhood, or a more recent loss or trauma are all confidently offered up as things that must be considered as causes for their depression or anxiety and put right in therapy. For an epidemiologist who can rarely put cause A together with effect B with much confidence or for very long (Davey Smith & Ebrahim, 2002), these leaps of faith are impressive. Given time, however, patients begin to see that the myriad of influences they endure each day makes a lost cause of concluding cause and effect and that concentrating on behaviour and cognition right now might be more profitable.

**A passion**

I cannot end without including a book on one of my passions. Like most otherwise lapsed New Zealanders, I still swim. Although exercise is a surprisingly neglected approach to disordered mental states (presumably because there is so much survival value in laziness), things are changing. I am fortunate enough to chair the trial steering committee for an ongoing, randomised trial of exercise for depression that is taking place in south-west England. My own experience is that exercise elevates mood and that swimming is the most sublime way to achieve that buzz. It also provides an unequalled time for reflection and meditation, which is not all about suspended animation or sitting in a corner. Perhaps thinking is more accurate or incisive when the cardiovascular system is at its most exuberant. No doubt someone somewhere is researching that. You can find no better description of swimming at the heart of a novel than in Jamie O’Nell’s *At Swim Two Boys*. Here swimming is the symbol for a passionate struggle of politics, religion, class and sexuality set in the time of the Easter uprising in Dublin in 1916. So, sometimes, when I am leaping about in wild surf that is determined to drown me or gliding through the limpid blue of my local gym’s swimming pool, I see a point in Being.
REFERENCES


Michael King  Department of Mental Health Sciences, Faculty of Biomedicine, University College London, Hampstead Campus, Rowland Hill Street, London NW3 2PF, UK. Email: m.king@medsch.ucl.ac.uk
The invocation of metaphor here explicitly refers to Szasz’s work on the concept of mental illness in the 1960s, specifically the charge that the concept is illegitimate: a metaphor or myth. Influential though this charge may have been, for example in shifting the terminology in the diagnostic manuals from mental illness to mental disorder, one can ask whether this issue of terminology matters much, compared with the clinical phenomena, the services and the science. For this reason, a book which targets this 40-year-old problem invented by Szasz starts somewhat at a disadvantage. The notion of metaphor is no doubt important and interesting, but at the same time it is also somewhat specialist and esoteric. Its value as a key to turn the great locks of problems in clinical practice, the science of psychopathology and the sociology of psychiatry is doubtful – and I’m inclined to think it bends and breaks, unable to withstand the forces. So to this reviewer’s mind this book on the metaphor of mental illness starts in the wrong place and uses the wrong tool. Nevertheless, it is a valuable update on several themes of conceptual analysis that run through the philosophy of psychiatry.

This book covers various issues concerning the concept of mental illness and its relation to general medical illness. It is a philosophical analysis, not a clinical or scientific one. The author considers the so-called likeness argument, which has various forms in the literature, according to which mental illnesses or disorders really are such because they are sufficiently similar to agreed physical illnesses. He criticises this form of argument on the grounds that the likenesses invoked typically recommend themselves as a result of the assimilation to the physical case, rather than being independent reasons for it. This criticism seems to work better for hypothesised likenesses, such as causation by disease, than evident ones such as distress and impairment of functioning. In any case, the author concludes that the categorisation of psychiatric conditions as illnesses is a matter of metaphor: ‘an imaginative shift into the illness category’.

Mood and Anxiety Disorders in Women
Edited by David J. Castle, Jayashri Kulkarni & Kathryn M. Abel
Cambridge University Press. 2006. 290pp. £27.99 (pb)
ISBN 0521547539

Interest in the area of women’s mental health has been slowly gathering pace. Women not surprisingly differ from men in terms of the epidemiology and pathophysiology of certain mental health problems, the treatments they respond to, the services they require and the issues they face. Women are almost twice as likely as men to suffer from depressive and anxiety disorders. They are less likely than men to misuse alcohol and other substances but, when they do, the impact on the family is profound. Failure to address gender-specific differences in mental health not only burdens women themselves but also families, society in general and the mental health of future generations. In this book, a multinational group of authors crystallises work in this area to create an invaluable resource for all those involved in women’s mental health.

The contributors consider mood, anxiety and related disorders from a broad biopsychosocial perspective, charting gender differences and gender-specific issues through life from before puberty to old age. The volume’s range is wide, covering not only anxiety, depression and bipolar disorder, but also childhood sexual abuse, domestic violence, gender-specific vulnerabilities to personality disorders, substance misuse, premenstrual syndrome, pregnancy, the post-partum period and the menopause. The authors appropriately round off the volume’s excellent collection by challenging clinicians’ a priori assumptions that women’s mood disorders in old age represent ‘the inevitable decline of dementia’, making instead a plea to redress that imbalance by challenging the view that a woman ‘has had her innings’.

Despite its attractive cover, the book is not quite coffee-table material. Most chapters are beautifully written while remaining rich in research information, but in some chapters, heavy biological, pharmacological and statistical terms might frustrate the efforts of the well-informed non-medical reader.
A small quibble (and only that) is that despite the comprehensive summaries throughout, I missed a satisfying concluding chapter which might have drawn together the excellent material of the preceding chapters. The volume ends abruptly following the chapter on old age, and, as a reader, I felt the need for a eulogy.

Nevertheless, I unequivocally recommend this book. It makes an ambitious contribution to our understanding of gender disparity within the field of women’s mental well-being, effectively collating current disparate information into a coherent integrative overview. The result is a collection of meaty essays which should comprehensively satisfy the appetite for an enlightened and broadened perspective.

Jessica Heron  Department of Primary Care and General Practice, University of Birmingham, Edgbaston B15 2TT, UK.
Email: j.c.heron@bham.ac.uk
doi: 10.1192/bjp.191.3.272a

Understanding Nicotine and Tobacco Addiction
Edited by Gregory Bock & Jamie Goode.
John Wiley & Sons. 2006. 284pp. £80.00 (hb).
ISBN 0470016574

This book is an edited collection of talks given by major researchers in nicotine addiction at a symposium funded by Novartis. Nicotine is an unusual addictive drug because it is not all that hedonic and it is perfectly possible to smoke and perform complex tasks, like driving a car, providing you do not crash while lighting up. It is much more difficult to get rats to self-administer nicotine than it is cocaine, for example. However, humans find it really difficult to stop using cigarettes and the majority of attempts to quit end in failure within a few weeks, even with maximum therapy. How can something so non-descript in its effects get such a grip on us? Unfortunately, no one contributing to this seminar will tell you, but readers will get partial answers to these questions.

The 15 chapters in this book are diverse. Some of them, for example the one on nicotinic acetylcholine receptor functions in the central nervous system, are essentially papers that give the results of one or a series of closely interrelated experiments. I always struggle with such work to understand where it fits in to the picture of smoking we see. Fortunately, many of the chapters are followed by the edited transcript of a discussion, in which, sometimes, clinical researchers try to grapple with the basic science and apply it to humans. In this case, however, the comments are left to the basic scientists alone, which means that less of an integrative perspective is offered. Nevertheless, the questions and comments do put the findings into a somewhat broader context. At times, these discussions are inadvertently amusing.

Chapters such as ‘Defining and assessing nicotine dependence in humans’ do take an integrative approach and probably offer insights that could not be gained by reading the journals. The author draws on his own recently published theory of addiction, which is not specific to tobacco, to examine how the DSM–IV criteria, and other widely used measures of dependence, apply to smoking, and offers new insights both on smoking and the concept of dependence in general. The fact that these chapters follow one another show the reader that the text leaps around without any linking and does not offer a coherent account of the phenomena of nicotine addiction and smoking.

This is an expensive book aimed at the nicotine researcher. Anyone who has attended conferences of the Society for Research in Nicotine and Tobacco will have heard many of these talks and had more fun than they will reading this book. However, it does summarise some aspects of the rather disparate approaches taken to understanding this most widespread of lethal addictions. Bringing these into the same symposium is one thing, integrating their insights to explain the tobacco epidemic is quite another.

Paul Aveyard  Department of Primary Care and General Practice, University of Birmingham, Edgbaston B15 2TT, UK.
Email: p.n.averyard@bham.ac.uk
doi: 10.1192/bjp.191.3.272

Mental Health Issues in the Media: An Introduction for Health Professionals

This is a timely book given the 10 years since Otto Wahl’s Media Madness and Greg Philo’s Media and Mental Distress. Both were landmark publications in tracing the evolution of influential media representations of psychiatric illness in the USA and UK respectively. We continue to battle the same stereotypes but Morris illustrates several victories where media makers have retreated to regroup. It is contemporary in the objects of its gaze, if not in outlook, and should be recommended reading for students and trainees who may need assistance in seeing the wood from the trees.

Quite rightly, he includes a chapter on literature – from trend-setting classics to the Harry Potter phenomenon. The breadth of UK television and internet examples
Psychiatry in the Scientific Image


In the dialectical problem of whether the disorders of the mind are basically biological or social, we are always being swayed one way or the other. You can tell which way a book with the title Psychiatry in the Scientific Image is going to jump. Its positioning is complex, however.

According to the author this book is deeply reactionary, a qualified defence of the medical model which shows psychiatry as a branch of medicine dedicated to uncovering the neurological basis of disease entities. This has intuitive appeal, given that we are animals with a biology including a brain that is the foundation of mental life. Also according to the author, the book is the first on psychiatry from within analytical philosophy of science. The result, therefore, is a deeply reactionary book at the cutting edge of philosophy of science. This is a finely balanced and subtle position that is not easy to summarise. For example, the deep conservatism has the brain as fundamental to psychology and psychiatry, and yet, recent philosophy of science envisages many levels of causal explanation, among which it is difficult to say which is fundamental. Tension is relieved here – the medical model vindicated – with the thought that psychological abnormalities can be traced to specific causal factors that are realised in brain tissue. The brain is fundamental in the sense that it realises everything. (A social science analogue is to have itself as fundamental in the sense that everything – including biomedical science – is a social practice.) This view of the brain as fundamental belongs with an up-to-date suitably broad understanding of neuroscience: that it draws on the cognitive and social sciences as well as molecular biology.

The book tackles three sets of questions about mental disorder: concept; explanation; and classification. It is weakest on the first topic, apparently taking for granted the fact of mental disorder, while cursorily dismissing social science critiques of the medical model. This untroubled view belongs generally with the avowedly ‘realist’ approach of the book, which wants to get on with tracking facts and causes, not worrying about concepts and construct validity. The book is strongest on multi-level causes and the lack of viability of reductionism, although there is also some tension here. Discussion of classification in the last part of the book rehearses the aspiration that nosology should track causal histories of conditions, not operationalised, observational criteria. However, the problem of reconciling controversial and shifting complex, multi-level, causal models of psychiatric conditions with a simple and relatively stable classificatory system for clinical and research purposes is, in this reviewer’s opinion, neither sufficiently recognised in the book nor resolved.

Peter Byrne
University College London, Gower St, London WC1E 6BT, UK. Email: p.byrne@ucl.ac.uk
doi: 10.1192/bjp.191.3.272a
After considering the definitions of addiction, West rethinks how we appraise behavioural theories and asks us to use ‘observation of nature’ as a starting point. He argues for a Popperian stance, that a theory is wrong if a counter-example exists, then sets the scene for his own theory with an erudite review of the literature.

The reader is walked through the evolution of his theory before it is introduced as ‘PRIME theory’ (plans, responses, impulses/inhibitory forces, motives and evaluations); a hierarchical representation of the motivational system which serves well as a template for human behaviour despite the unwieldy schematics.

West blends PRIME with chaos theory in a Pythagorean attempt to understand behaviour in mathematical terms. He invites us to think of the motivational system as an epigenetic landscape with hills and valleys (Chreods) through which a ball (time) travels resulting in a number of potential future outcomes depending on its course. He acknowledges the metaphor in applying chaos theory to addictive behaviour and this book joins a growing discourse on the subject. The concept is user friendly and explains why addictions manifest so differently despite often similar underlying pathologies. A strength is that it allows for such variance, but as a result the theory becomes too inclusive for rigorous testing. My limited understanding of chaos theory left me wondering whether human factors such as the capacity for mentalisation might influence its relevance to psychiatric disorders.

In citing economic and neurophysiological theories, the author’s intention is that other disciplines could add to the work. It would also be interesting to hear how PRIME interfaces with cultural and psycho-dynamic constructs. In the closing chapter, West applies his theory to addiction and suggests approaches for intervention that I would have liked to hear more about, such as how one might detect an imminent Chreod bifurcation.

In the end I felt that, in addition to PRIME theory, the book had introduced a valuable representation of what could be called good psychiatric formulation. West encourages us to think differently about people and populations with substance use problems, and I now find myself wondering how my clients’ epigenetic landscapes might be looking. The work is grounded in common sense and goes a long way towards explaining what the author calls the big observations (‘observation of people in their natural habitat or uncontested scientific evidence’), and it adroitly deals with the challenges inherent in postulating any unified theory of human behaviour. He leaves the way clear for future research and is ready to pass on the baton in the collective endeavour of incremental science.

Amanda Redvers Capital and Coast District Health Board, Rptide House, 61–63 Thorndon Quay, Wellington, New Zealand.
Email: Amanda.redvers@ccdhb.org.nz
doi: 10.1192/bjp.191.3.273a

Aviation Mental Health: Psychological Implications for Air Transportation
Edited by Robert Bor & Todd Hubbard.
Ashgate Publishing. 2006. 376 pp. £65.00 (hb). ISBN 0754643719

If you are looking for a book to read on a long haul flight, this is not the one! A random selection of some of the key issues discussed in Aviation Mental Health reveals why. Pilot suicide by aircraft and the nature of language used in airport announcements ‘this is your last and final call’ being two topics that may set off a train of thought that is not modified even by the fascinating fact that air travel is 18 times safer than staying at home.

Aviation mental health is a topic that impinges on many aspects of medical practice, from the management of flying phobias to severe in-flight medical emergencies such as acute psychotic episodes. This is perhaps the first textbook to cover the whole range of aviation mental health from selection and management of flight and cabin crew through to the management of the psychological consequences of flying and crashing.

With such a wide target audience it has perhaps been difficult in this first edition to balance the content between specialist and generalist information. The chapters range from quite technical multi-author submissions on sleep and mental performance with general applications through to single-author chapters on highly specialised topics such as psychological aspects of astronaut selection. Overall, however, the content is well balanced with an appropriate level of theory and advice on practical management.

The style of the book, as with many multi-author collections, lacks coherence. Perhaps the editors will be able to address this in a second edition? The layout of the chapters within the book is confusing. Part 1, ‘psychological issues of flight and cabin crew’, deals with issues relating to passengers, whereas Part 2, ‘psychological processes amongst passengers and crew’, does not. The third section, ‘related themes in aviation’ has the appearance of a standby line of passengers who couldn’t be fitted into one of the previous two sections, the content ranging from occupational factors in pilot mental health through to aviation psychology in South Africa.

Overall this is a useful practical guide to an important area of occupational mental health which, despite the inevitable teething troubles of a first edition, is well worth reading.

Roger Bloor Harplands Hospital, Stoke-on-Trent ST4 6TH, UK. Email: pco01@keele.ac.uk
doi: 10.1192/bjp.191.3.274

Eating Disorders in Children and Adolescents
Edited by Tony Jaffa & Brett McDermott.

This international and multi-authored volume is aimed at practitioners and researchers in the field of eating disorders in children and adolescents. The book is
pleasingly presented and generally well-written, although it was a little surprising to find that some of the chapters were written by authors not working with children and adolescents. The introduction consists of a truly fascinating historical and developmental review, including reference to Norton’s conclusion in 1694 that the condition ‘is due to a malfunction of the brain . . . ’ – clearly a man ahead of his time.

In section two, entitled ‘scientific underpinnings’, there are learned chapters on the regulation of food intake and body weight, the development of weight and shape concerns, and the relation of dieting to eating pathology. The chapter on physical and cognitive changes is uneven in that the section on cognitive changes neglects many fascinating new findings. The chapter on genetics is too technical for the average practitioner or researcher but that on epidemiology would have benefited from some

editorial assistance. The neuroimaging chapter is sound but would have benefited from more focus on the findings in childhood and adolescence.

Section three focuses on ‘abnormal states’, with useful contributions on anorexia nervosa, eating disorders in boys, atypical eating problems, disability and chronic illness, and bingeing and bulimia nervosa, as well as chapters on comorbidity, and trauma and obesity.

The section on evidence-based care has useful reviews of acute and chronic medical complications, individual and family psychotherapies, models of service delivery, and psychopharmacology, albeit the latter being rather too slanted toward eating disorders in adults.

The final section, entitled public health perspectives, offers interesting discussions on primary and secondary prevention, although frustratingly makes no mention of the potential of targeting children in middle childhood (e.g. 6- to 10-year-olds) who are possibly more likely to benefit than adolescents, in whom unhealthy attitudes may have already developed. The chapter on outcome reminds us of how poor the prognosis is for eating disorders.

This volume does not focus on the subtleties and complexities of clinical practice and those wishing to enhance their clinical skills will need to look elsewhere. However, there is much of interest and value for those who wish to gain an overview of current knowledge of the many problems in this population, without having to delve into a voluminous text.

Bryan Lask  St George’s University of London, Department of Psychiatry, Jenner Wing, Cranmer Terrace, London SW17 0RE, UK.
Email: b.lask@sghms.ac.uk
doi: 10.1192/bjp.191.3.274a
**From the Editor’s desk**

**PETER TYRER**

---

**SWIMMING UNDER WATER**

Michael King (pp. 268–270), in his intriguing list of ten books, includes Jamie O’Neill’s novel *At Swim Two Boys* and invokes the pleasure of swimming as it ‘provides an unequalled time for reflection and meditation.’ I agree, but set my sights rather lower. I love swimming, but having more respiratory reserve than efficient stroke action, prefer gliding in the deep to thrashing on the surface. This serenity of silent propulsion is an excellent environment to get everything into proportion. This issue of the *Journal* is a reflective one and while I could not pretend that reading it should simulate exactly all the sensations of swimming underwater, I hope it helps to restore the balance we all have to maintain in psychiatric practice. It includes the first of the articles in our Reappraisal section, in which subjects that have long provoked argument and debate are reviewed. Wilson & Nutt (pp. 195–197) show that there is now much more to the treatment of insomnia than the traditional benzodiazepines and I am glad to note the word ‘addiction’ is not mentioned once. Goodwin & Geddes (pp. 189–191) challenge the view that schizophrenia should be at the centre of our secondary care system in psychiatry. In making out the case for bipolar disorder as an alternative, backed up by Harrison & Critchley (pp. 192–194), it might be worth reminding you now that Oxford swings and is bidding to become the Bipolar Capital of the World, so there could be a little partiality here. And, as we are reminded, bipolar disorder and schizophrenia are part of the same spectrum recently remoulded roughly from the neat split imposed by that former colossus, Emil Kraepelin (Craddock & Owen, 2005).

But it is certainly right to question the attention given to what has become a ubiquitous diagnosis in anyone who has the misfortune to stay for more than a few days in a psychiatric unit in a developed country, and whose attribution is now so closely linked to hospital care. My new specialist registrar came to me brightly burnished for his job a year ago but has now ruefully remarked, ‘I thought I was coming here to learn how to treat psychiatric patients, but now realise that I had to learn how to treat beds’. The real problem for those trying to make psychiatric care more cost-effective is that the relentless preoccupation of services with insightless psychosis prevents us looking at other matters of great import. So anything we can do that can break this pattern is invaluable, and measures such as ‘symptom relabelling’ (Crumlish et al, pp. 262–263) or other forms of relearning (Cooke et al, pp. 234–237) to improve insight would be excellent if they worked, and might allow some time to look at other issues such as the influence of personality abnormality and outcome of schizophrenia, completely virgin territory at present (Newton-Howes et al, 2007), or some clinical developments following from the many demonstrated neuroanatomical abnormalities such as cortical sulcal thickness (Goghai et al, pp. 229–233) that now compete for our souls. Anything, indeed anything, that takes our minds off the Premier Antipsychotic Drug League where millions of pounds in transfer fees are spent in moving a drug up a notch or two in an increasingly irrelevant table (Rosenheck, pp. 238–245), would be welcome, could prevent early death (Joukamaa et al, 2006) and allow more attention to be paid to the risk of suicide (Hawton et al, 2005). And, almost by default, the care of significant affective and anxiety problems has become lost to psychiatry and transferred to other technology in the form of self-help and computer-controlled therapy. We will return to this subject in a later issue but it must be of some solace to the humans who traditionally deliver this care that the ‘pure’ form of self-help has some deficiencies ( Rapee et al, pp. 246–252) and that patients appreciate at least a little guidance before reading the instructions and swimming under the cure-all waters (Khan et al, pp. 206–211).

**STIGMA AND THE MEDIA**

While in reflective mode it is worth reminding ourselves that in our well-intentioned attempts to protect our patients from the perils of outrageous fortune we may sometimes stigmatised them. If we are to make schizophrenia no more discriminatory than diabetes (Lee et al, 2005), we have to expose patients and problems of mental illness more openly to the media and not clothe them awkwardly in ill-fitting garments derived from the Data Protection Act. The Royal College of Psychiatrists has just reported a success with a new publications venture (Persaud, 2007), a book written with the help of many College experts that includes a case history at the beginning of every chapter. It is edited by Raj Persaud and briefly enjoyed 6th spot in the UK best-selling books list before *Harry Potter and the Deathly Hallows* rudely pushed it aside. Some people smile ruefully when I tell them about this success and regard it as all slightly unnecessarily. I don’t agree. If we can get the world reading about mental ill health in all its guises and can show how it can be resolved in language that everyone understands, we are half-way there. So watch out, diabetes.


