Psychiatry in pictures:

Psychiatry in pictures


Highlights of this issue:

Highlights of this issue
Sukhwinder S. Shergill

EDITORIALS:

Possibilities for the prevention and treatment of cognitive impairment and dementia
David Burke, Ian Hickie, Michael Breakspear, and Jürgen Götz

Prison mental health inreach services
Julie Steel, Graham Thornicroft, Luke Birmingham, Charlie Brooker, Alice Mills, Mari Harty, and Jenny Shaw

REVIEW ARTICLES:

Chinese herbal medicine for schizophrenia: Cochrane systematic review of randomised trials
John Rathbone, Lan Zhang, Mingming Zhang, Jun Xia, Xiehe Liu, Yanchun Yang, and Clive E. Adams

SPECIAL ARTICLES:

Sir Martin Roth FRS: An appreciation
Alan Kerr and David Kay
PAPERS:

Mental health of US Gulf War veterans 10 years after the war
Rosemary Toomey, Han K. Kang, Joel Karlinsky, Dewleen G. Baker, Jennifer J. Vasterling, Renee Alpern, Domenic J. Reda, William G. Henderson, Frances M. Murphy, and Seth A. Eisen

Common mental disorders and the built environment in Santiago, Chile
Ricardo Araya, Alan Montgomery, Graciela Rojas, Rosemarie Fritsch, Jaime Solis, Andres Signorelli, and Glyn Lewis

Cross-national prevalence and correlates of adult attention–deficit hyperactivity disorder

Serotonergic function in children with attention – deficit hyperactivity disorder: Relationship to later antisocial personality disorder
Janine D. Flory, Jeffrey H. Newcorn, Carlin Miller, Seth HartY, and Jeffrey M. Halperin

Personality disorder and impaired functioning from adolescence to adulthood
Andre W E. Skodol, Jeffrey G. Johnson, Patricia Cohen, Joel R. Sneed, and Thomas N. Crawford

Cognitive remediation therapy in schizophrenia: Randomised controlled trial
Til Wykes, Clare Reeder, Sabine Landau, Brian Everitt, Martin Knapp, Anita Patel, and Renee Romeo

Target groups for the prevention of late – life anxiety
Filip Smit, Hannie Comijs, Robert Schoevers, Pim Cuijpers, Dorly Deeg, and Aartjan Beekman

All–cause mortality among recipients of electroconvulsive therapy: Register-based cohort study
Trine Munk-Olsen, Thomas Munk Laursen, Poul Videbech, Preben Bo Mortensen, and Raben Rosenberg

Dementia Screening Questionnaire for Individuals with Intellectual Disabilities
Shoumitro Deb, Monika Hare, Lindsay Prior, and Sabyasachi Bhaumik
SHORT REPORTS:

Foetal brain development in offspring of women with psychosis
Mary C. Clarke, Mary Cannon, Matthe W Hogg, Maureen N. Marks, Sue Conroy, Susan J. Pawlby, Anne Greenough, and Kypros Nicolaides

Zuclopenthixol in adults with intellectual disabilities and aggressive behaviours:
Discontinuation study
Frank Haessler, Thomas Glaser, Manfred Beneke, Akos F. Pap, Ralf Bodenschatz, Olaf Reis the Zuclopenthixol Disruptive Behaviour Study Group

Correspondence:

Smoke-free mental health units
D. Arnone and L. J. Simmons

Smoke-free mental health units
A. Williams

X-chromosome abnormality and schizophrenia
M. Otter

Authors' reply
S. Van Rijn, A. Aleman, H. Swaab, and R. S. Kohn

One hundred years ago:

Havelock Ellis and Henry Rollin

Book reviews:

Social Inclusion of People With Mental Illness
David O'Flynn

Cognitive Therapy of Schizophrenia
Rebecca Rollinson
Cognition + Addiction
Duncan Raistrick

From the Editor's desk:

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

BY SUKHWINDER S. SHERGILL

APPRECIATION AND HOPE

Everybody is inspired by a good story and the special article - an appreciation – about Sir Martin Roth (Kerr & Kay, pp. 375-378) definitely falls into this category, as well as providing an informative historical perspective on the prominent figures and changes in British psychiatry during the past century. One of his main contributions was to increase research into dementia, and the editorial by Burke and colleagues (pp.371-372) highlights the promise of advances in dementia research. They review more recent data suggesting that there is considerable potential for generating new synapses, neurones and cortical networks at all stages of brain development and conclude that there is room for cautious optimism in applying these approaches to cognitive impairment and dementia.

BABY BRAINS, COGNITIVE REMEDIATION AND LEARNING DISABILITY

There are subtle structural changes in the brains of patients with schizophrenia. Clarke et al (pp.445-446) investigated whether these were evident in the brains of offspring, where the mother had a psychotic illness. Their prenatal ultrasound study showed no significant differences in brain measures between offspring of ill mothers and control participants, but a non-significant trend for increased lateral ventricular size in the former group. Cognitive deficits are common in schizophrenia and they are correlated with patients’ functional outcome. Cognitive remediation therapy has been shown to result in improvements in working memory in schizophrenia, with associated improvement in social functioning and at little additional financial cost (Wykes et al, pp. 421-427). It is difficult to diagnose dementia in learning disability; Deb et al (pp.440-444) describe the development and psychometric properties of a dementia screening questionnaire and conclude that this can now be used as a valid, reliable, observer-rated screen for dementia in this population. Antipsychotic treatment is often used to reduce aggressive symptoms in disturbed patients. Haessler et al (pp.447-448) demonstrate that discontinuation of zuclopenthixol treatment in patients with learning disability and a history of prior aggressive behaviour can result in increased aggression.

DEPRESSION AND ECT

Electroconvulsive therapy (ECT) is an effective treatment for depression, with rapid effects relative to pharmacological interventions. However, the public perception is that of a dangerous intervention; Munk-Olsen et al (pp.435-439) examined the mortality associated with ECT, using a register-based cohort study. They found lower overall mortality from natural causes in the patients who received ECT but a higher suicide rate, especially within the first week of treatment. They emphasise the importance of monitoring suicide risk, both during and at termination of treatment.

PERSONALITY DISORDER, ADHD, SEROTONIN AND STABILITY

Disturbance of central nervous system serotonergic function has been associated with aggressive behaviour in adults; this relationship is less clear in childhood and adolescence. A prospective study of children with attention-deficit hyperactivity disorder (ADHD) showed that those with lower childhood serotonergic function were more likely to develop antisocial personality disorder as adults at 9-year follow-up (Flory et al, pp. 410-414). The importance of ADHD in adulthood has been demonstrated by Fayad et al (pp.402-409), who report a prevalence of 2-4% in adults surveyed as part of an international epidemiological study. It was more prevalent in higher-income countries, often coexisted with other disorders and caused considerable disability. However, the longitudinal course of personality disorder, particularly its stability and the consequences for functioning, remain unclear. Skodol et al (pp.415-420) found that individuals with persistent personality disorder over their lifetime demonstrated significant impairment in adulthood. However, adult-onset disorders were also associated with similar impairment. Interestingly, only 25% of patients with a personality disorder at the age of 22 retained that diagnosis at the age of 33 years, and those in remission of their symptoms at 33 years demonstrated relatively little impairment.

DANGEROUS PERSONALITY DISORDER

The supplement to the journal includes several well articulated views regarding the often contentious category of dangerous and severe personality disorder. The background to the development of this category, and the opportunities offered by it, are particularly relevant. The articles by Maden (pp.s8-s11) and Mullen (pp.s3-s7) reflect on the fact that politicians and civil servants invented dangerous and severe personality disorder in 1999 in response to public outcry, but that the initial cynicism has been tempered by the knowledge that this programme of work will have tangible benefits for some patients with personality disorder.
Psychiatry in pictures
EDITED BY ALLAN BEVERIDGE

Looking across the estuary from Tor Point (1980–1995). Ronald Quixano Henriques (b. 1949). Picture selection and text by Dr Glenn Roberts

Ron is a prolific artist of strikingly and richly colourful images. Commenting on this picture he says, ‘It’s like a judge or a doctor looking over his estate, across the valleys, far from the crowds, he’s reassimilated his riches’. Ron has spent nearly all his adult life in large psychiatric hospitals and residential care, despite which he has cultivated and sustained a passionate creative life. The succession of vivid and distinctive images he has produced from the early 1980s he attributes to being in very frightening circumstances and ‘ultraperceptive’. He comments that during his time at Digby, ‘every time I tried to paint voices came on the radio trying to distract me from being clever – I know they are not real. For 12 years I’ve been trying to adjust and express my experiences, trying to put myself back into the world . . . new paintings I do will not be circumstantial’. Recently Ron has moved from residential care for a more independent but still supported lifestyle. He is gradually fulfilling his aim of becoming more ‘bohemian’ such that he can sit in cafes and watch the world go by, and he has enthusiastically resumed painting.

Another of Ron’s paintings has appeared on the cover of the Gaskell publication Enabling Recovery: The Principles and Practice of Rehabilitation Psychiatry (2006), for in many ways Ron’s life illustrates some of the complexities and paradoxical possibilities of recovery. These and other paintings are available as high-quality reproductions from his cousin at http://www.gallery36.co.uk; all profits go to Ron.
Possibilities for the prevention and treatment of cognitive impairment and dementia

DAVID BURKE, IAN HICKIE, MICHAEL BREAKSPEAR and JURGEN GOTZ

Summary  The human brain has a remarkable capacity for plasticity, but does it have the capacity for repair and/or regeneration? On the basis of controversial new evidence we speculate that the answer may be 'yes', and suggest that clinicians should therefore approach cognitive impairment and dementia with a new, cautious optimism.

Declaration of interest  None.

NEURAL PLASTICITY, ANGIOGENESIS AND NEUROGENESIS

It is widely accepted that physical activity, learning and social factors exert alterations in gene expression, giving rise to changes in patterns of neural connectivity and functionality throughout life (Kandel, 1998). These changes are achieved through mechanisms of neural plasticity, synaptogenesis, angiogenesis and possibly neurogenesis. The evidence for neurogenesis in the adult human brain, however, is controversial (Bhardwaj et al, 2006). A number of studies have demonstrated neurogenesis in the healthy adult human brain, in the hippocampus (Eriksson et al, 1998; Draganski et al, 2004) and in the olfactory bulb (Bedard & Parent, 2004). Studies have also demonstrated neurogenesis in the hippocampus of patients with Alzheimer’s disease (Jin et al, 2003), in the subependymal layer adjacent to the ventricles in patients with Huntington’s disease and Alzheimer’s disease (Curtiss et al, 2003), and around areas of cerebral cortical infarction in younger adults with stroke (Jin et al, 2006).

Neural plasticity, synaptogenesis and neurogenesis require parallel angiogenesis. New vessels develop in response to tissue demands, mediated principally by vascular endothelial growth factor, which responds to local factors such as inflammation, blood pressure, oxygen saturation, lipid levels, insulin levels and tissue perfusion (Fam et al, 2003). Many vascular risk factors may therefore modify and promote these processes of neural plasticity, synaptogenesis, angiogenesis and neurogenesis.

CARDIOVASCULAR AND CEREBROVASCULAR DISEASE AND COGNITIVE IMPAIRMENT

The respective associations between cardiovascular disease, cerebrovascular disease and cognitive impairment are well known. The risk factors for cardiovascular disease – hypertension, diabetes, obesity, smoking, low levels of high-density lipoprotein (HDL), high levels of low-density lipoprotein (LDL), high concentrations of fibrinogen and of homocysteine, and alcohol misuse – are also risk factors for cerebrovascular disease. Additional risk factors for cerebrovascular disease include cardiac arrhythmia, carotid atheroma, hypotension, transient ischaemic attacks, coronary artery bypass grafts, angio-plasty, ischaemic heart disease and metabolic syndrome. These can all then be considered to be risk factors for cognitive impairment and most of the dementias (for review, see O’Brien et al, 2003).

MECHANISMS OF NEURO-VASCULAR DAMAGE AND REPAIR IN THE BRAIN

Vascular risk factors lead directly or indirectly to oxidative stress and a cascade of inflammatory events that result in vascular damage in the brain, compromising neural activity and hence causing cognitive impairment (Yaffe et al, 2005). Oxidative stress may occur peripherally in response to obesity, smoking, alcohol, inactivity, atherosclerosis, hyperlipidaemia and psychosocial stress, and centrally in response to hypertension, diabetes, hyperhomocysteinaemia, hypoperfusion, protein aggregation in Alzheimer’s disease and ischaemia (McEwen, 2002). Oxidative stress then leads to inflammation, and this in turn results in a loss of endothelial wall integrity, further compromising perfusion and leading to increased surrounding cell damage and loss. It would therefore seem reasonable to speculate that repair of cell damage in the brain caused by oxidative stress, inflammation and vascular damage can be expected if the conditions promoting the latter events are treated or prevented, and the potential for angiogenesis, neural plasticity, synaptogenesis and neurogenesis is maximised.

POSSIBILITIES FOR TREATMENT AND PREVENTION

Exercise has been shown through observational studies to be associated with enhanced reaction time and a variety of cognitive executive control processes, retrospectively, cross-sectionally, prospectively and by meta-analysis; and observational studies suggest the cognitive benefits of exercise are achievable in young and old individuals with and without pre-existing cognitive impairment (Larson et al, 2006).

Similarly, structured formal learning has been implicated as a way of enhancing targeted cognitive abilities in a sustained manner, including verbal episodic memory, reasoning and speed of information processing (Ball et al, 2002). Additionally, complex environments that stimulate problem-based learning promote structural and functional neuronal changes, and older people may respond by recruiting neural circuitry in a fashion that is different from younger individuals (Grady et al, 2003).

Social engagement is associated with positive effects on cognition in humans, and similar positive effects have been observed in relation to supportive psychotherapy and problem-solving therapy, social relations and social support, social ties and marital status, and living arrangements and social network indices (Helmer et al, 1999; Alexopoulos et al, 2003). The biological mechanism is proposed to be neural plasticity (the cognitive reserve hypothesis), neurogenesis and vasculogenesis (the vascular hypothesis) and cortisol regulation (the stress hypothesis) (Frattigioni et al, 2004).

Dietary regulation and supplementation could also be reasonably expected to play a part in providing the chemical substrates necessary to improve neurovascular
function. Increased HDL and decreased LDL concentrations and marine omega-3 polyunsaturated fatty acid consumption are associated with better cardiovascular and cognitive function (Kalminj et al., 2004). Reduced energy intake with nutritional maintenance may suppress oxidative stress, stabilise calcium homoeostasis, induce neurotrophic factors and may reduce the β-amyloid deposition associated with Alzheimer’s disease (Patel et al., 2005).

There is also speculation that intake of antioxidant compounds in red wine, dark chocolate, curcumin, some fruits, grains and vegetables, vitamin E and vitamin C may improve cardiovascular function (Engelhart et al., 2002).

Medical interventions including cessation of smoking, treatment of depression, control of hypertension, folate acid plus vitamin B₁₂ supplementation sufficient to reduce raised homocysteine levels and melatonin may provide reduction of risk for cardiovascular, cerebrovascular and depressive illness (Hickie et al., 2005). Although the limited benefits of cholinesterase inhibitors and N-methyl-D-aspartate receptor antagonists in dementia are generally acknowledged (Görz et al., 2006), there is ongoing controversy with regards to the role of other pharmacological agents such as non-steroidal anti-inflammatory drugs, statins and hormone replacement therapy (Rosenberg, 2005).

CONCLUSION

Recent advances in the neurosciences suggest that young, old and impaired human brains may be able to respond to the demands of activity, experience and environmental factors by creating new functional synapses, neurons and networks through the intimately related processes of angiogenesis, neural plasticity, synaptogenesis and neurogenesis. These advances are particularly exciting in relation to the convergence of evidence regarding the contribution of vascular risk factors, genes, diet, physical activity, cognitive activity, psychological functioning and social functioning to the aetiology of acquired cognitive impairment and dementia. Taken together, these findings open the door to an array of possible new directions in the treatment and prevention of cognitive impairment and dementia through interventions that promote mental health, lifelong education, functional intimate relationships and social engagement, and that target healthy eating, dietary supplementation, exercise and effective cardiovascular treatment (when needed). In our opinion, where the prior paradigm of dementia as an inevitably progressive neurodegenerative disease was often a cause for clinical pessimism and inaction, there is now an emerging evidence base for a more optimistic, proactive approach to cognitive impairment and dementia.

ACKNOWLEDGEMENTS

The authors thank Gavin Andrews for advice on an earlier version of this paper.

REFERENCES


Prison mental health inreach services

JULIE STEEL, GRAHAM THORNicroFT, LUKE BIRMINGHAM, CHARLIE BROOKER, ALICE MILLS, MARI HARTY and JENNY SHAW

Summary

Prison mental health inreach teams have been established nationwide in England and Wales over the past 3 years to identify and treat mental disorders among prisoners. This paper summarises the policy content and what has been achieved thus far, and poses challenges that these teams face if they are to become a clear and effective component in the overall system of forensic mental healthcare.

As many as nine out of every ten prisoners in the UK display evidence of one or more mental disorders (Singleton et al, 1998). Despite this, detection of mental illness on reception to prison has been found to be ineffective, with many prisoners' mental disorders left both undetected and untreated (Birmingham, 2003). Better and more accessible services need to be provided to mentally ill prisoners. This is not a new problem (Gunn et al, 1978). The standard of prison healthcare has been of concern since the earliest reports on prison welfare, with frequent campaigns for the National Health Service (NHS) to take responsibility for prison healthcare from the Home Office (Royal College of Psychiatrists, 2007). This was the main recommendation of Patient or Prisoner, published in 1996, which highlighted the shortcomings in the prison healthcare system; it also argued for equivalence, namely that 'prisoners are entitled to the same level of healthcare as that provided in society at large' (H M Inspectorate of Prisons, 1996). Recommendations made in The Future Organisation of Prison Health Care (H M Prison Service & NHS Executive Working Group, 1999) were accepted by the government, which led to the Department of Health and the Home Office sharing responsibility for prison health. In 2003 it was announced that responsibility for the provision of healthcare would be completely transferred from the Home Office to the Department of Health from April 2006.

IMPACT AND POTENTIAL ADVANTAGES OF PRISON MENTAL HEALTH INREACH TEAMS

At the same time as inreach teams have been introduced, there has been a reduction of 18% in prison suicides for 2004–5 (Howard League for Penal Reform, 2006). It is not clear whether this is due in part to the new inreach teams, as a series of concurrent factors are likely to have contributed to this finding. These include risk-reduction initiatives within prisons, such as the Safer Locals strategy and the implementation of the Assessment, Care in Custody and Teamwork (ACCT) programme. Another probable factor is the 'dilution' effect seen in the USA whereby a rising imprisonment rate means that on average a less unwell or disabled population is sentenced or on remand, and because a larger proportion of the prison population serves long sentences it tends to be more clinically stable (Gore, 1999). In this respect the pattern of imprisonment in the UK is progressively changing to resemble American trends.

Mentally disordered offenders in prison could be managed through the same channels as those in the community, if inreach teams were to form part of a joined-up approach to care in which there were functioning crisis teams and assertive outreach teams in the custodial environment. Secure hospital care could therefore be arranged within the course of fixed sentences through transfers under sections 47 and 48 of the Mental Health Act 1983 (Department of Health, 2006). This would enable the more appropriate use of scarce and valuable secure beds.

The REMIT AND CHALLENGES OF INREACH TEAMS

Prison inreach teams were intended to be the main vehicle for improvements in mental health services for prisoners, especially those with severe and enduring mental illness. In fact, forms of such teams have existed for several decades at some prisons, for example Belmarsh and Pentonville, and were provided by non-forensic specialists. The current mental health inreach teams are different in that they are intended to provide care to all prisons in England and Wales. The original intention was stated in this way:

‘For those persons judged to have the greatest need, the NHS will fund the establishment of multi-disciplinary teams, similar to community mental health teams (CMIHTs) offering to prisoners the same sort of specialised care they would have if they were in the community’ (Department of Health & Home Office, 2001).

The key point is that, upon joining the NHS, these new inreach teams should bring the mainstream NHS framework to apply equally to prisoners.

Despite nationwide inreach teams being a relatively new initiative, the challenges to such services are already clear. There are already signs of ‘mission creep’. The original intention was to restrict inreach services to treating people with severe and enduring mental illness, but already national policy has been broadened to include all those in prison with any mental disorder (Brooker et al, 2005). Prisoners often present a complicated clinical picture as they frequently have complex and comorbid problems. Are the general mental health staff in such teams, who do not necessarily have any forensic training, sufficiently expert to provide effective care? In fact a perverse incentive may now operate, in that inreach teams are less likely to want create referrals for themselves. The role of inreach services in relation to people with personality disorders is not yet clear. Now that the evidence base for effective interventions for personality disorder is growing, meeting the treatment needs of people who frequently present with personality rather than illness-driven problems has to be addressed in practice throughout the prison establishment. Should this fall within the remit of an inreach team, be provided in specialist personality disorder units, or should there be a combination of the two? Does the general psychiatric in-patient
sector have the capacity to accept transfers of people identified in prison as requiring hospital assessment and treatment? Are inreach teams effective for both sentenced and remand prisoners, and can such teams operate rapidly enough to connect the latter successfully, given high turn-around rates and unpredictable court decisions and release dates? To date all these questions about the remit of inreach teams remain unanswered.

Evidence of treatment models that have been found to be effective in the community, such as community mental health and assertive outreach teams, cannot be directly applied to the prison population because issues of criminality can complicate the picture (Brooker et al., 2002). Constraints within the prison environment such as security issues, information sharing and treating prisoners without their consent have an impact on the translation of community-based treatments into secure settings. Conflicting views on the balance between care and control within a prison environment may also affect the outcome of using these treatment models in prison.

Drug and alcohol misuse and dependency need to be a core focus of such clinical interventions in prison. The greatest health issue (and the real solution to suicide risk) is to address the substance misuse issues of prisoners (Gore, 1999). Yet paradoxically there is relatively little evidence for effective interventions for people with ‘dual diagnosis’, i.e. concurrent substance misuse and severe mental illness. Such patients are often excluded from studies of the general adult psychiatric population, and so caution is required in extrapolating findings from the general adult services to the prison population (Brooker et al., 2002). Drug and alcohol treatment services in prisons, using the Counselling, Assessment, Referral, Advice and Throughcare (CARAT) system are already well established. Through a more formal collaboration between inreach and CARAT services, some form of dual diagnosis service could be implemented. Drug-free wings might be a therapeutic setting in which to treat prisoners with such comorbidity.

Clinical experience to date suggests that inreach services are operating using limited and idiosyncratic models of care. The average team size for example, is three members of staff. Official guidance has been deliberately non-prescriptive, and innovative commissioning by primary care trusts will therefore be required to sustain the initial momentum to deliver an equivalent standard of care nationwide.

CONCLUSION
Giving the NHS direct responsibility to commission mental healthcare for prisoners allows us to reconsider what services should be provided on the basis of equity and effectiveness. Should home treatment and crisis response teams be as available to prisoners as to everyone else? Should assertive outreach teams, and specialist drug and alcohol treatment teams, similarly supplement generic inreach teams by taking on patients who need such intensive treatment and who happen to be in prison? In other words, should prisoners receive care that is either identical or equivalent to the care that they would receive if they were in the community? Should we continue to insist that prisoners cannot be treated without or against their consent? How best can people with mental illness be assisted to engage with community services after release from prison? In fact, inreach teams are only one element in a complex and rapidly changing landscape, including new arrangements for care pathways (Department of Health & National Institute for Mental Health in England, 2005), treatment of women in prison, and policy changes to expedite transfers to hospital under sections 47 and 48 of the Mental Health Act in less than 1 week by 2008 (Royal College of Psychiatrists, 2007). This new national policy in England has therefore prompted a wholesale renaissance in the treatment of mentally ill prisoners in recent years: the next challenge is to assess the impact of these changes in practice.

DECLARATION OF INTEREST
None.

REFERENCES


Sir Martin Roth FRS

An appreciation

ALAN KERR and DAVID KAY

Since the mid-19th century central Europe had been the cradle of conceptual thinking in clinical psychiatry. Following the turbulence and persecution in the 1920s and 1930s some of a later generation of mid-European psychiatrists emigrated to Britain, bringing with them the traditions of meticulous and detailed observation, a broad clinical perspective and fresh ways of looking at problems. Among those individuals making significant contributions were Willy M ayer-Gross, Erich Guttmann, Erwin Stengel, Felix Post, F. Kraupl Taylor, Max Hamilton and Martin Roth. Roth, who was born in Budapest in 1917, the last survivor of this extraordinarily gifted group, died on 26 September 2006 at the age of 88.

In this Appreciation, bearing in mind that Martin Roth was not given to talking freely about his own experiences and feelings, we have used some of his own words, spoken on those rare occasions when he revealed more of himself than usual, as he did on the occasion of the celebration on his 80th birthday or when interviewed by one of us (A.K.).

EARLY CAREER

Martin Roth's decision to embark on a career in medicine arose out of 'the necessity to qualify at something I could make a living at'. He had considered becoming a musician at the age of 13 or 14 but felt he couldn't afford to take the risk. Nevertheless, when a medical student, his talent enabled him to play the organ at weddings and the piano for ballet classes, providing much-needed financial assistance at the time. He continued to play the piano throughout his life, Bach giving him particular pleasure. His entry to St Mary's Hospital in 1936 had been delayed for 3 months by the need for surgery. In 1942 he had to have a much more serious operation. The diagnosis was one of regional ileitis.

'I expected to have a relatively short life and this [experience] has influenced me very much.' At St Mary's he had little interest in the prevailing rugby culture, but did row for the hospital.

While a house physician, Roth was struck by the story of Mary Walker, an unassuming physician at a London County Council hospital. Dr Walker had gone to a lecture on myasthenia gravis and had heard that the symptoms bore a close resemblance to curare poisoning. She later asked what the antidote for curare poisoning was, and on being told it was physostigmine gave the drug to some patients with the condition; after an injection, their symptoms vanished. She presented these cases at the Royal Society of Medicine,

'but never really gained full recognition for her amazing discovery. But that inspired me. I viewed this as a basis for going into the study of the neurological foundations of mental activity.'

Martin Roth decided to take a post in neurology at Maida Vale Hospital, where he worked for the outstanding neurologist Russell (later Lord) Brain, who became his friend.

'He [Brain] had a great interest in psychiatry and psychoanalysis which he kept quiet. He was a man of mighty intellect and remarkable gifts as a personality.'

The decision to work in neurology was made in the hope

'that those activities of the highest level of nervous activity would be presenting disturbances or diseases at the cortical level. The interest of British neurology at the time, with some exceptions, stopped at the neck. Anybody who dabbled above that level was virtually a charlatan.'

After two and a half years Roth became disenchanted with neurology; however, he felt he had learnt a great deal, and commented, 'I had done my own neuropathology in a case of anosognosia'.

Martin Roth also recalled:

'I worked in London during the bombing attacks with flying missiles, incendiary bombs and V2 rockets. I took my Membership of the Royal College of Physicians in 1944 when explosions could be heard all around us. Our home near Tower Bridge was bombed and my mother and father moved up to near Manchester. A lot of my possessions were destroyed.'

Having been impressed by the scientific integrity and precision of Elliot Slater's writings, Roth was introduced at his own request to Slater (later to become an influential editor of the British Journal of Psychiatry; Roth was to be a member of this journal's editorial board for 40 years), which proved to be the start of a close, lifelong personal and professional relationship. Slater asked him to go to the Maudsley Hospital:

'When I arrived there I found [Professor Sir Aubrey] Lewis remarkable for his ability to pick out gaps in knowledge and read the relevant literature. At first I was impressed but later I found he poured jars of cold water on people, some of whom gave splendid presentations. I was also finding difficulty with the regime, and I clashed with Lewis. I felt unhappy and it was clear that I had no future there. I left after two years and three months.'

Martin Roth then moved to the Crichton Royal Hospital in Dumfries, a mental hospital with a fine reputation in psychiatric research, where he met Willy M ayer-Gross, whose immense knowledge, enthusiasm and warm-heartedness he found infectious. M ayer-Gross ignited his interest in old age psychiatry by giving him the task of reviewing developments in the field for a chapter in an American book. So began his interest in the clinical area in which his greatest achievements were to be made.

Also while at the Crichton, when only 36, he was invited by M ayer-Gross to become
co-author, with Elliot Slater and himself, of what became a highly influential postgraduate textbook in British psychiatry. The first edition of Clinical Psychiatry was published in 1954 and the book was later translated into five languages. Roth himself wrote that it:

'was intended to foster precise observation and disciplined inference in clinical practice. In research, it recommended an empirical approach [to] the modern classification of psychiatric disorders. . . . The historical and developmental dimension and the concepts of empathy and understanding developed by Jaspers were also treated as essential for the study of neurosis and personality disorder in particular but important for reaching a diagnosis in all forms of disorder.'

This book played a considerable part both in repopositioning biological psychiatry at the forefront of clinical practice in the UK and in the dramatic paradigm shift in America in the late 1960s and early 1970s away from the predominance of psychoanalysis and social psychiatry, to give centre stage once more to issues of psychopathology, diagnosis and classification, culminating in the publication of DSM-III. This sudden and remarkable change was described by the eminent American psychiatrist Gerald Klerman in a 70th birthday tribute to Martin Roth as the 'neo-Kraepelinian revival', the essential tenet of which was that psychiatry is the specialty of medicine concerned with mental disorders, with the scientific understanding of those disorders and the treatment of individuals suffering from them, utilising a categorical approach. Although this was the approach he espoused, Martin Roth was at pains to emphasise that he did not wish to eliminate dynamic concepts in the study of illness. Those who worked with him clinically were aware of his subtle understanding of the interaction between psychological, social and biological factors in the individual patient.

OLD AGE PSYCHIATRY AND OTHER RESEARCH INTERESTS

After leaving the Crichton Royal, Roth became Director of Clinical Research at Graylingwell Hospital, a large county mental hospital in Sussex, where he carried out his ground-breaking work on the clinical syndromes and outcomes of the mental disorders of old age. In 1958 he was invited by the World Health Organization – the start of a long association with this body – to act as rapporteur and consultant to organise the first Expert Committee on Mental Health of the Aged. This resulted in a working paper that drew attention to the paucity of epidemiological evidence and dearth of research in the whole field.

Roth had moved in 1956, at the age of 39, to take up the Chair of Psychological Medicine in Newcastle upon Tyne. This period, which lasted 21 years, proved enormously fruitful. On arrival, he reorganised the teaching programme and had to answer ward consultations (done in a white coat) because

'in an English teaching hospital the professor who does not respond to his consultant colleagues is invisible. He gets no support and I think this is right, so I was often on the wards.'

Indeed, Roth did not shirk the duties of a clinical consultant, with weekly ward rounds, out-patient clinics and case conferences, where he was an inspiring teacher. He established units for child psychiatry, neurosis and psychogeriatrics, and provided postgraduate training for clinical psychologists within his department.

Martin Roth also had to raise research funds and quickly obtained the support of the Medical Research Council, which set up a research group in Newcastle that carried out clinical and outcome studies in affective disorders and epidemiological work in old age psychiatry. The affective work led to a prolonged dispute regarding the classification of depressive illnesses with Robert Kendell, then Reader at the Institute of Psychiatry (and later Professor of Psychiatry in Edinburgh), which remained unresolved to this day. Roth also instigated important clinico-pathological studies with Dr Garry Blessed, consultant psychogeriatrician, and Sir Bernard Tomlinson, then pathologist at the Newcastle General Hospital, which established the pathological distinctness of the clinical syndromes in his classification of old age disorders.

In 1977, at the age of 58, he moved to the Foundation Chair in Psychiatry at Cambridge, where he had to set up a new department but continued his work on affective and anxiety disorders and jointly edited the Handbook of Anxiety, in five volumes (1988–1992). This period was notable also for the publication of CAMDEX, The Cambridge Mental Disorders of the Elderly Examination. However, his most important achievement was to assemble a strong international team, including Claude Wischik and supported by Sir Aaron Klug, to investigate the molecular pathology of Alzheimer's disease.

ROTH AND THE COLLEGE

Another enduring contribution to his specialty was his key role in the establishment of the Royal College of Psychiatrists. Although initially he was one of a group of academic psychiatrists who favoured psychiatry becoming a faculty subsumed within the Royal College of Physicians, following a meeting with its President (by a curious quirk of fate, Lord Brain), an abrupt change of course was decided upon and the mainly university group threw in their lot with their colleagues, drawn predominantly from the mental hospitals, who had fought a long and vigorous campaign for an independent Royal College of Psychiatrists.

Although he was a late candidate to stand for the position of first President of the College, Martin Roth was elected by a comfortable margin. Under his leadership psychiatrists in training were given formal involvement and membership on key College bodies, an acknowledgement of the commitment he had given to them at the rumbustious inaugural meeting of the College concerning the need for a truly democratic structure. With Kenneth Ransley and William Trethowan he helped establish respectively the model Approval exercise and MRCPsych Examination, with inevitable and substantial benefits to education and training and therefore to clinical standards in psychiatry. However, it was very much Roth's personal achievement during his time in office (1971–1975), acting largely on his own, to find a permanent home for the College. It was from Lord Goodman that he learnt that 17 Belgrave Square was available. The then huge sum of £750 000 had to be borrowed and substantial amounts of interest paid.

'We didn't choose to go to a fashionable place but we couldn't get any other.'

With additional help from generous sponsors, the premises in one of the most attractive squares in London were secured.

Martin Roth was a dominant personality who possessed charm and wit, an exceptional intellect and breadth and depth of knowledge of the literature within both psychiatry and medicine. He possessed an unusual capacity for addressing conceptual and clinical conundrums in a novel and imaginative way. As Thomas Barnes put
pathologist J.A.N. Corsellis, demonstrated a quantitative relationship between cognitive level and extent of brain damage, whether of the type found in Alzheimer's disease or arteriosclerotic dementia. These findings placed his separation of dementia proper from functional disorders on a secure basis, led to the birth of the specialty of psychogeriatrics, and transformed the care of the elderly. Subsequently, employing Roth's definitions, a small population-based study, the forerunner of much larger studies in Europe and the USA, allowed preliminary estimates to be made of prevalence, which showed that the rate of dementia increased progressively with age after 60.

Among the functional disorders Roth was fascinated by the non-affective paranoid-hallucinatory psychoses of late onset, found mainly in women, but without evidence of dementia or brain damage, that he called 'late paraphrenia'. The choice of this term was intended to leave on one side its relation to schizophrenia, but it was open to criticism and the condition now cannot be identified by any rubric in DSM or ICD. This is to be regretted, since its status is still controversial.

At Cambridge he returned to the topic of dementia and his team produced a stream of papers on the molecular pathology of the protein tau, found in an abnormal insoluble form in the Alzheimer neurofibrillary tangles - the neocortical density of which is strongly correlated with dementia symptoms - whereas in normal ageing tau remains mostly in solution. Although these studies have not yet led to a specific treatment they constitute a notable scientific achievement.

Martin Roth's studies of affective and anxiety states were undertaken in a climate in which the view of Sir Aubrey Lewis that there was no essential difference between endogenous (or psychotic) and reactive (or neurotic) disorders was predominant. With his colleague Roger Garside, senior lecturer in clinical psychology, Roth used the multivariate statistical programs becoming available on computers to look for categories of illness and distinct patient groups. In a series of papers he found evidence that these groups differed in respect of symptoms, outcome and response to treatment. However, replication by others did not always produce the same results, and the dichotomy implicit between 'endogenous' and 'reactive' was not always fulfilled. We have looked at the latest editions of the major classificatory systems for an indication of the current status of this dichotomy. In ICD-10 neurotic and stress-related disorders (section F40-48), often with coexistent depression, include phobic-anxiety and other anxiety disorders, obsessive-compulsive disorder, and reactions to severe stress and adjustment disorders. The classification of depressive disorders is based on severity, but endogenous (or melancholic) depression finds a place as an option to state whether or not a 'somatic syndrome' is present. In DSM-IV the main type of depression is major depressive disorder, but the presence of 'melancholic features' - which are identical to those of endogenous depression - can be specified. It appears that although classification has become more complex and nomenclature has changed, the concept of endogenous depression as a distinct non-neurotic disorder survives.

In Aubrey Lewis's classification, depressive and anxiety disorders form a single illness within the affective disorders, distinguished only by severity. Roth showed that after separating cases into anxiety states and depressive illness on the basis of the patients' predominant mood, these groups differed from each other in their clinical features, premorbid personalities, response to treatment and longer-term outcome. The fact that both DSM-IV and ICD-10 allocate affective (mood) disorders and anxiety states to separate sections seems to imply implicit agreement with Roth's view that these disorders are distinct. Both these classifications treat manifestations of anxiety such as panic disorder, agoraphobia, social phobias and generalised anxiety as categorical entities, but Roth objected that although this approach had stimulated research, it ignored the traditional European concept of anxiety disorders as manifestations of personality disorder (and therefore as continuous with normality). He
concluded that both categorical (syndrome) and dimensional (personality diagnosis) features were necessary to complete the clinical picture of affective disorders, and for prediction of outcome.

For his scientific achievements in psychiatry, particularly in the field of old age psychiatry, Martin Roth became one of only three psychiatrists (of whom the earliest was Sigmund Freud) to be elected Fellow of the Royal Society.

SELECTED BIBLIOGRAPHY


Chinese herbal medicine for schizophrenia

Cochrane systematic review of randomised trials

JOHN RATHBONE, LAN ZHANG, MINGMING ZHANG, JUN XIA, XIEHE LIU, YANCHUN YANG and CLIVE E. ADAMS

Background  Chinese herbal medicine has been used to treat millions of people with schizophrenia for thousands of years.

Aims  To evaluate Chinese herbal medicine as a treatment for schizophrenia.

Method  A systematic review of randomised controlled trials (RCTs).

Results  Seven trials were included. Most studies evaluated Chinese herbal medicine in combination with Western antipsychotic drugs; in these trials results tended to favour combination treatment compared with antipsychotic alone (Clinical Global Impression 'not improved/worse' n=123, RR=0.19, 95% CI 0.1–0.6, NNT=6, 95% CI 5–11; n=109, Brief Psychiatric Rating Scale 'not improved/worse' RR=0.78, 95% CI 0.5–1.2; n=109, Scale for the Assessment of Negative Symptoms 'not improved/worse' RR=0.87, 95% CI 0.7–1.2; n=109, Scale for the Assessment of Positive Symptoms 'not improved/worse' RR=0.69, 95% CI 0.5–1.0, NNT=6 95% CI 4–162). Medium-term study attrition was significantly less for people allocated the herbal/antipsychotic mix (n=897, four RCTs, RR=0.34, 95% CI 0.2–0.7, NNT=23, 95% CI 18–43).

Conclusions  Results suggest that combining Chinese herbal medicine with antipsychotics is beneficial.

Declarai on of interest  None.

theory. Nevertheless, because of the enormous population of China, even if herbal medicines are given to only a small proportion of the estimated 13 million Chinese people with schizophrenia, these treatment approaches could still be some of the most prevalent used for this illness.

METHOD

Full details of all methods used and the pre-defined inclusion criteria are published elsewhere (Rathbone et al, 2005). Randomised controlled trials were included if participants had schizophrenia, schizophreniaform psychosis or a schizophrenia-like illness, diagnosed by any criteria. Interventions included Chinese herbal medicines (plant, animal or mineral) given in any dosage or combination, with or without a basis in traditional Chinese medical theory, compared with any other approach.

Studies were identified from searches of the Cochrane Schizophrenia Group's register of trials, which incorporates regular searches of BIOSIS Inside, CENTRAL, CINAHL, EMBASE, MEDLINE and PsycINFO; the hand-searching of relevant journals and conference proceedings and searches of several grey literature sources. Additional databases searched included the Traditional Chinese Medical Literature Analysis and Retrieval System, the Chinese Biomedical Database, the China National Knowledge Infrastructure database and the Allied and Complementary Medicine Database (AMED). Full details of the English and Mandarin phrases used are reported elsewhere (Rathbone et al, 2005).

Data were not utilised from studies in which more than 50% of participants in any group were lost to follow-up (this does not include the outcome of 'leaving the study early'). In studies with a less than 50% withdrawal rate people leaving the study early were considered to have had the negative outcome, except for the event of adverse effects and death. For binary outcomes, the fixed-effects relative risk and its 95% confidence interval were calculated. The numbers needed to treat/harm (NNT/NNH) were also calculated. An estimate of the weighted mean difference (WMD) between groups and its 95% confidence interval were calculated for continuous data. Data were not pooled if standard deviations were too wide, suggesting considerable skew (Altman & Bland, 1996). Heterogeneity between studies was
assessed by inspecting the relevant graph; this was supplemented using the I² statistic (Higgins et al., 2003). If inconsistency was high (≥ 75%), the data were not pooled but were presented separately and the reasons for heterogeneity investigated.

Citations were inspected independently by at least two reviewers. The reliability of the data extraction was checked using a methodology was used in three studies, all of which used Ginkgo biloba extract (EGb761) combined with antipsychotics (NNT = 1.2) when compared with haloperidol, as were data from the Scale for the Assessment of Negative Symptoms (SANS) (n = 109, RR = 0.87, 95% CI 0.7 to 1.2). However, the Scale for the Assessment of Positive Symptoms (SAPS) did slightly favour the herbal medicine plus haloperidol combination (n = 109, RR = 0.69, 95% CI 0.5 to 1.0; NNT = 6, 95% CI 4 to 162). Continuous short-term BPRS data - Meng et al. (1996), unknown antipsychotic; Zhu et al. (1996), chlorpromazine - significantly favoured the herbal medicine plus antipsychotic combination (Fig. 2(c)), but data were heterogeneous (I² = 81%).

RESULTS

Electronic searches resulted in over 640 citations but most clearly did not meet the inclusion criteria. Full copies of only 14 studies were obtained, of which we could include 7 (Table 1). Of those we excluded, three were not randomised (Cao & Wang, 2000; Gong et al., 2000; Rong, 2001), three did not report usable data (Zhao et al., 1997; Wang, 1998b; Han et al., 2002) and one study did not use Chinese herbal medicine (Zhen & Feng, 1992).

We identified 16 citations dating from 1987 to 2002 for the seven included studies. Overall, descriptions of studies were poorly reported. Two trials were available in both Chinese and English (Luo et al., 1997; Zhang et al., 2001a, 2001b, 2001c) in Chinese only (Meng et al., 1996; Zhu et al., 1996; Chen et al., 1997; Zhang et al., 1997) and in English only (Zhao et al., 1997). All seven included studies were conducted in China and were described as being randomised, but none gave a description of the allocation procedure. Double-blind methodology was used in three studies, all of which used Ginkgo biloba extract (EGb761) combined with antipsychotics. All trials included in this review contained a moderate risk of bias (category B; Alderson et al., 2004). Trials ranged in sample size from 40 to 545 participants and lasted from 20 days to 6 months. Only one study (Zhang et al., 1997) attempted to allocate treatment according to traditional Chinese medicine syndrome differentiation. The other six studies employed Western diagnoses of schizophrenia with no further differentiation into the traditional Chinese syndromes, and six used operational diagnostic criteria. Three studies included people with chronic schizophrenia (mean duration 17 years), three did not report participants' history of illness and one study involved mostly people at first admission to hospital.

Herbal medicine alone

v. chlorpromazine

Only one study (Zhang et al., 1987) gave the treatment group herbal medicines without the addition of an antipsychotic. Over a 20-day period, global state outcome ‘not improved/worse’ significantly favoured the control group receiving chlorpromazine (n = 90; RR = 1.88, 95% CI 1.2 to 2.9, NNT = -4, 95% CI 2 to 14). No participant left the study early.

Herbal medicine plus antipsychotics v. antipsychotics alone

Herbal medicines given according to traditional Chinese medicine syndrome differentiation - in only one study (Zhang et al., 1997), using dang gui cheng qi tang or xiao yao yao san - when combined with antipsychotic medication (unspecified) scored significantly lower for the outcome of global state ‘not improved/worse’ than the control group given unspecified antipsychotics (NNT = 6, 95% CI 5 to 11; Fig. 2(a)). Further global state data from the Clinical Global Impression (CGI) scale - Meng et al. (1996), unknown antipsychotic; Zhu et al. (1996), chlorpromazine - also favoured the herbal medicine plus antipsychotic group (Fig. 2(b)).

Zhang et al. (2001) found Brief Psychiatric Rating Scale (BPRS) scores dichotomised to ‘not improved/worse’ were equivocal (n = 109, RR = 0.78, 95% CI 0.5 to 1.2) when Ginkgo biloba plus haloperidol were compared with haloperidol, as were data from the Scale for the Assessment of Negative Symptoms (SANS) (n = 109, RR = 0.87, 95% CI 0.7 to 1.2). However, the Scale for the Assessment of Positive Symptoms (SAPS) did slightly favour the herbal medicine plus haloperidol combination (n = 109, RR = 0.69, 95% CI 0.5 to 1.0; NNT = 6, 95% CI 4 to 162). Continuous short-term BPRS data - Meng et al. (1996), unknown antipsychotic; Zhu et al. (1996), chlorpromazine - significantly favoured the herbal medicine plus antipsychotic combination (Fig. 2(c)), but data were heterogeneous (I² = 81%). Medium-term BPRS data (Fig. 2(c)) also favoured the herbal medicine plus antipsychotic combination: Luo et al. (1997), antipsychotics clozapine, chlorpromazine, sulpiride, perphenazine and haloperidol; and Zhang et al. (2001), haloperidol (n = 621, WMD = -4.17, 95% CI -5.5 to -2.8). Medium-term SANS scores (Fig. 2(d))
### Table 1: Characteristics of included studies

<table>
<thead>
<tr>
<th>Date of study</th>
<th>First author</th>
<th>Number of publications</th>
<th>Methods</th>
<th>Participants</th>
<th>Interventions</th>
<th>Control group</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1987</td>
<td>Zhang</td>
<td>2</td>
<td>N K</td>
<td>H 2.9</td>
<td>FE 90 16 ± 51</td>
<td>NK</td>
<td>DGCQT 50 ml b.i.d. (n = 45)</td>
</tr>
<tr>
<td></td>
<td>Meng</td>
<td>1</td>
<td>Yes</td>
<td>H 8</td>
<td>N K 40 18 ± 60</td>
<td>M, F</td>
<td>Ginkgo biloba² 240 mg/day + antipsychotics (n = 23)</td>
</tr>
<tr>
<td>1996</td>
<td>Zhu</td>
<td>1</td>
<td>OL</td>
<td>H 4.3</td>
<td>NK 67 17 ± 53</td>
<td>F</td>
<td>Hirudo &amp; Rheum palmatum² + chlorpromazine, 300 mg/day (n = 32)</td>
</tr>
<tr>
<td>1996</td>
<td>Chen</td>
<td>2</td>
<td>SB</td>
<td>H 25.1</td>
<td>CI 120 27 ± 58</td>
<td>NK</td>
<td>Xingshen + antipsychotics³, 300 ± 700 mg/day (n = 60)</td>
</tr>
<tr>
<td>1997</td>
<td>Luo</td>
<td>3</td>
<td>Yes</td>
<td>H 16</td>
<td>NK 545 18 ± 60</td>
<td>M, F</td>
<td>Ginkgo biloba², 120 mg t.d.s. + antipsychotics (n = 315)</td>
</tr>
<tr>
<td>1997</td>
<td>Zhang</td>
<td>1</td>
<td>OL, C</td>
<td>H, CH 12</td>
<td>CI 123 Mean ± 32</td>
<td>M, F</td>
<td>DGCQT or XYS, 200 ± 300 ml/day + antipsychotics³ (n = 66)</td>
</tr>
<tr>
<td>2001</td>
<td>Zhang</td>
<td>6</td>
<td>Yes</td>
<td>H 12</td>
<td>CI 109 27 ± 61</td>
<td>M, F</td>
<td>Ginkgo biloba², 360 mg + haloperidol 0.25 mg/kg per day (n = 56)</td>
</tr>
</tbody>
</table>

C, community; CI, chronic illness; DGCQT, dang gui cheng qi tang; F, female; H, hospital; M, male; NK, not known; OL, open label; SB, single-blind; XYS, xiao yao san.

1. Standardised extract of Ginkgo biloba (EGb761).
2. No further details on medicine and/or dosage.
3. Chlorpromazine, clozapine, sulpiride.
4. Clozapine, chlorpromazine, sulpiride, perphenazine and haloperidol.
Fig. 2  Comparison of herbal medicine + antipsychotic v. antipsychotic (BPRS, Brief Psychiatric Rating Scale; NNT, number needed to treat; RR, relative risk; SANS, Scale for the Assessment of Negative Symptoms; WMD, weighted mean difference).
significantly favoured the herbal medicine plus antipsychotic group.

Adverse events are associated with antipsychotic medication, and combining herbal medicines with chlorpromazine (Zhu et al., 1996) did not mitigate extrapyramidal adverse effects, with both groups being equivocal. Constipation, however, was significantly lower in the herbal plus antipsychotic combination group (0/32) despite patients receiving the constipating antipsychotic chlorpromazine (n = 67; RR = 0.03, 95% CI 0.0 to 0.5; NNH = 2, 95% CI 2 to 4); the comparison group (chlorpromazine alone) fared less well (19/35). Medium-term studies found significantly fewer patients leaving the study early (Fig. 2(e)) in the herbal plus antipsychotic group (n = 897, four RCTs, RR = 0.34, 95% CI 0.2 to 0.7; NNH = 23, 95% CI 18 to 43).

### Sensitivity analysis: Ginkgo biloba alone or plus antipsychotics v. antipsychotics

Studies of *Ginkgo biloba* were tested in a sensitivity analysis by comparing them with the original pooled data (*Ginkgo biloba* data pooled with other herbs). Effect sizes for CGI and BPRS scores were increased for *Ginkgo biloba* when analysed separately, although these differences were not statistically significant.

### DISCUSSION

Six of the seven studies evaluated the use of Chinese herbs for schizophrenia rather than traditional Chinese herbal medicine for schizophrenia, i.e. treatment was allocated according to a diagnosis of schizophrenia without further differentiation according to traditional Chinese methodology. Study sizes were generally small and pooled data were typically derived from one or two studies. All outcomes, therefore, were underpowered. The one study that incorporated traditional Chinese medical theory did show significant improvement in global state but was limited by lack of blinding. There were no descriptions of allocation concealment and no assurances that blinding was maintained. The type of antipsychotic used and the dosages were often poorly reported, although three studies used the same herbal intervention - *Ginkgo biloba* (EGb761). The remainder, however, used different herbal medicines, and unfortunately all three *Ginkgo biloba* studies used different antipsychotic medications.

#### Herbal medicine alone v. antipsychotics

Global state measured as ‘not improved/worse’ favoured the chlorpromazine group (NNT = 67; RR = 0.03, 95% CI 0.0 to 0.5; NNH = 2, 95% CI 2 to 4); the comparison group (chlorpromazine alone) fared less well (19/35). Medium-term studies found significantly fewer patients leaving the study early (Fig. 2(e)) in the herbal plus antipsychotic group (n = 897, four RCTs, RR = 0.34, 95% CI 0.2 to 0.7; NNH = 23, 95% CI 18 to 43).

#### Herbal medicine plus antipsychotics v. antipsychotics

The herbal medicine group receiving either *dang gui cheng qi tang* or *xiao yao san* plus antipsychotics were significantly less likely to have an outcome of ‘no change or worse’ compared with participants receiving only antipsychotics, measured using the Clinical Global Impression scale (NNT = 6, 95% CI 5 to 11). This could be an important finding and does fit with the CGI continuous scores. These results are broadly encouraging and suggest that combining herbal medicines with antipsychotics might be beneficial, although results are only based on two small studies (total n = 103). These vaguely positive finding also apply to mental state outcomes. The dichotomised BPRS and SANS measures reported by Zhang et al. (2001); n = 109) were equivocal, but SAPS scores again showed borderline significance in favour of the herbal medicine plus antipsychotic combination. Medium-term continuous SANS data, however, provided more robust results, with three studies (n = 741) favouring the herbal plus antipsychotic combination group. The experimental group had, on average, nine points less on this scale than those allocated to antipsychotic drugs alone. In our opinion, in this group of chronically unwell people such an average difference would be noticeable and clinically meaningful. Further supporting this improvement, both short-term and medium-term BPRS scores were significantly better for those receiving herbal medicines plus antipsychotics compared with those receiving antipsychotic drugs alone, although there was heterogeneity in these results. The latter might have been due to the use of different antipsychotic drugs between trials.

Adverse effect Treatment Emergent Signs and Symptoms scores were reported by Zhang et al. (2001), but standard deviations were wide and no conclusion can be made with confidence. Only one study (Zhu et al., 1996; n = 67) reported extrapyramidal symptoms, and these were not significantly different between groups. In one trial in which both groups were given chlorpromazine, constipation was significantly more frequent in the control group (NNH = 2). In this trial the herb used was a purgative used also in Western medicine - *Rhizoma rhizoma palmatum* (rhubarb).

Numbers of participants leaving the study early in the short term were similar for both groups. Medium-term data showed significantly fewer left early in the herbal medicine plus antipsychotic group compared with people receiving only antipsychotics (n = 897; 2% vs. 7%). In this context of these studies, the addition of herbal medicine did not worsen treatment compliance and there is the suggestion that the addition of the herbal medicine made it easier for participants to take standard antipsychotics.

We did a post hoc sensitivity analysis for the single herb *Ginkgo biloba*, used outside the traditional Chinese medicine approach within a Western model of schizophrenia. We found no evidence that this particular herb had remarkable effects.

The application of traditional Chinese herbal medicine is fundamentally interwoven with syndrome differentiation. Failure to apply syndrome differentiation may result in treatments being ineffective or even harmful. Despite this, there is some evidence that these Chinese herbal medicines, combined with antipsychotics and given in a way that is not in keeping with their normal use within traditional Chinese medicine, may be beneficial for people with schizophrenia across a range of outcomes. If these medicines are used within their usual context the positive effects could be greater. Even the gains seen in this review would still be important for the millions for whom these treatments are used. Both West and East need well-reported (Mohan et al., 2001) randomised trials that are adequately powered, blinded and of sufficient duration so we can detect meaningful
treatment effects with high levels of confidence.

REFERENCES


Correspondence: John Rathbone, Cochrane Schizophrenia Group, Academic Department of Psychiatry and Behavioural Sciences, University of Leeds, 15 Hyde Terrace, Leeds LS2 9LT, UK. Tel: +44 (0)113 343 1897; fax: +44 (0)113 3432723; email: jrrathbone@cochrane-sz.org

(First received 31 May 2006, final revision 8 September 2006, accepted 16 January 2007)
Mental health of US Gulf War veterans 10 years after the war

ROSEMARY TOOMEY, HAN K. KANG, JOEL KARLINSKY, DEWLEEN G. BAKER, JENNIFER J. VASTERLING, RENEE ALPERN, DOMENIC J. REDA, WILLIAM G. HENDERSON, FRANCES M. MURPHY and SETH A. EISEN

Background Gulf War veterans reported multiple psychological symptoms immediately after the war; the temporal course of these symptoms remains unclear.

Aims To assess the prevalence of war-era onset mental disorders in US veterans deployed to the Gulf War and in non-deployed veterans 10 years after the war.

Method Mental disorders were diagnosed using structured clinical interviews. Standard questionnaires assessed symptoms and quality of life.

Results Gulf War-era onset mental disorders were more prevalent in deployed veterans (18.1%, n=1061) compared with non-deployed veterans (8.9%, n=1128). The prevalence of depression and anxiety declined 10 years later in both groups, but remained higher in the deployed group, who also reported more symptoms and a lower quality of life than the non-deployed group. Remission of depression may be related to the presence of comorbid psychiatric disorders and level of education. Remission of anxiety was related to treatment with medication.

Conclusions Gulf War deployment was associated with an increased prevalence of mental disorders, psychological symptoms and a lower quality of life beginning during the war and persisting at a lower rate 10 years later.

Declaration of interest None. Funding detailed in Acknowledgements.

Approximately 700,000 US military personnel were deployed to the Middle East during the 1991 Persian Gulf War. Unique aspects of that deployment included a relatively large proportion of reservists and National Guard units, exposure of deployed personnel to potentially harmful natural and manufactured environmental toxins (Presidential Advisory Committee on Gulf War Veterans' Illnesses, 1996) and relatively few casualties. Epidemiological studies have demonstrated greater psychological symptoms and disorders among these personnel compared with Gulf War-era veterans who were not deployed to the Persian Gulf region (Perconte et al., 1993; Sutker et al., 1993; Southwick et al., 1995; Iowa Persian Gulf Study Group, 1997; Fukuda et al., 1998; Proctor et al., 1998; Stuart & Bliese, 1998; Ishoy et al., 1999; Unwin et al., 1999; Ismail et al., 2002). The National Health Survey of Gulf War Era Veterans and Their Families (Eisen et al., 2005) was designed to collect epidemiological data at several time points to assess the prevalence of medical and psychological conditions in both deployed and non-deployed veterans. Beginning in 1995, postal and telephone surveys were conducted of 15,000 deployed veterans and 15,000 non-deployed veterans. Those who had been deployed to the Gulf War reported more psychological symptoms, poorer functional impairment and poorer health-related quality of life compared with the non-deployed group (Kang et al., 2000). The study reported here evaluated a subset of these veterans using face-to-face psychological examinations conducted between 1998 and 2001. Based on retrospective report, we calculated the prevalence of mental disorders beginning during the deployment period, and their continued prevalence 10 years later.

METHOD

Study population and recruitment

Recruitment for the survey phase of the National Health Survey of Gulf War Era Veterans and Their Families performed in 1995 has been described elsewhere in detail (Eisen et al., 2005). Briefly, the US Department of Defense's Defense Manpower Data Center identified the entire cohort of 693,826 deployed veterans and approximately half (800,680) of the non-deployed veterans who were in military service between September 1990 and May 1991. A total of 15,000 deployed and 15,000 non-deployed veterans were solicited to participate in the study. To ensure that female, reservist and National Guard personnel were adequately represented, a stratified random sampling method was applied to each group so that a fifth of each sample were women (3000), a third were reservists (5000) and approximately a quarter were members of the National Guard (4000).

For the examination phase of the study, a list of potential participants was created by random selection from the 11,441 deployed and 9476 non-deployed veterans who participated in the 1995 study, stratified by deployment status and region of last known residence at the time of the original survey (based on home telephone area code). Potential participants were assigned to the participating Veterans Affairs (VA) medical centre closest to their home. Participating medical centres were located in Albuquerque, Baltimore, Birmingham, Boston, Cincinnati, Hines (Chicago), Houston, Miami, Minneapolis, New Orleans, New York, Portland (Oregon), Richmond, Salt Lake City, San Diego and St Louis. Recruitment packages that included an introductory letter, a detailed explanation of the purpose and nature of the study, a letter of intent form and a pre-addressed stamped return envelope were mailed to the veterans. Because of lower participation among non-deployed veterans, an additional 799 were solicited to obtain examined groups of equal size.

Signed letters of intent were returned to the Hines VA Cooperative Studies Program Coordinating Center, which forwarded them to the participating VA medical centre to which the veteran was assigned. Site personnel then contacted the participant and scheduled the examination. Travel, hotel, per diem costs and an honorarium of $200 were provided by the research project. The protocol and consent form were approved by the Hines Cooperative Studies Program Human Rights Committee and the institutional review board at each individual site and at the Brockton Veterans Affairs Medical Center. Participants gave

385
signed informed consent shortly before the start of the examination.

**Mental health assessment**

Mental health was assessed using two methods: structured clinical interviews, yielding mental disorder diagnoses; and paper-and-pencil, self-report measures of current symptoms. In the structured interviews, participants were asked about their lifetime experience of different symptoms, including the times when symptom constellations started and stopped. With these data, onset prevalence for disorders was calculated for the period ranging from 1 January 1991 to 30 July 1993, which encompasses the period from the beginning of the conflict to the date beyond which no further deployment to the Middle East occurred. Disorders with an onset during this period are referred to as Gulf War-era onset disorders. To assess the course of these mental disorders, prevalence rates were assessed for war-era onset disorders still present within 1 year of the current study, approximately 10 years after the resolution of the Gulf War. We also calculated the prevalence of disorders with onset prior to January 1991 and overall lifetime prevalence.

**Mental disorders**

Diagnoses of post-traumatic stress disorder (PTSD) were made with the Clinician Administered PTSD Scale (CAPS; Blake et al., 1995), a structured interview yielding PTSD diagnoses according to DSM-IV criteria (American Psychiatric Association, 1994). Participants first indicated their exposure to stressful life events on a standard life events checklist; a follow-up interview then assessed whether these events met criterion A (exposure to a traumatic event). This assessment was not externally validated through record review. The participant was then asked to choose the two most stressful of the events reported. The symptom interview focused on the presence of any PTSD symptoms related to the two events, over the veteran's lifetime as well as in the past month. Symptoms were assessed regardless of whether the selected events met criterion A, but the disorder was only diagnosed if both criterion A and the symptom criteria were met. Participants could receive up to two diagnoses of PTSD associated with the two events; however, cases of PTSD in the time intervals discussed above were calculated per individual. We asked whether events were related to the Gulf War. In assessing the presence of war-era onset PTSD 10 years after the war, we assured that PTSD diagnoses were linked by event (e.g. if a person had war-era onset PTSD from event 1, we checked the duration of the PTSD related to event 1 to determine whether it was present 10 years later). The remaining DSM-IV Axis I psychiatric diagnoses were made using the Composite International Diagnostic Interview (CIDI; Andrews & Peters, 1988), a computerised structured interview which yields diagnoses based on DSM-IV criteria. Diagnoses were then categorised according to the DSM-IV classification system.

**Symptoms**

Current psychiatric symptom severity was assessed using three scales. Higher scores on all three scales indicate greater symptomatology. The PTSD Checklist (Blanchard et al., 1996) was used to assess PTSD symptoms in the past month using 17 items rated on a scale of 1 to 5. We report the mean total score and the percentage of probable PTSD cases, defined by a total score of 50 or greater. The Beck Depression Inventory II (BDI-II; Beck et al., 1996) was used to assess depressive symptoms in the past 2 weeks using 21 items rated on a scale of 0 to 3. We report the mean total score and the percentage of cases in the following categories of depression severity: minimal (total score 0–13), mild (14–19), moderate (20–28) and severe (29–63). The Beck Anxiety Inventory (BAI; Beck & Steer, 1993) was used to assess anxious symptoms in the past 7 days using 21 items rated on a scale of 0 to 3. We report the mean total score and the percentage of cases in the following categories of anxiety severity: minimal (total score 0–7), mild (8–15), moderate (16–25) and severe (26–63).

**Quality of life**

Healthcare quality of life

The 36-item Short Form Health Survey (SF-36; M. H. Ormey et al., 1993; Ware et al., 1993; Kazis et al., 1998) was used to assess mental health-related quality of life in the 4 weeks preceding the evaluation. Items focused on current perception of health and normal daily functioning were rated on a Likert scale, and were summed into physical and mental component scores. Scores were standardised to a mean of 50 and standard deviation of 10; higher scores indicate better quality of life. We have previously reported on the physical component scores (Eisen et al., 2005), and report on only the mental component in this paper.

**General quality of life**

The Quality of Life Inventory (QoLI; Frisch, 1994) was used to measure general life satisfaction. It inquires about 16 areas of life, which are rated by participants on importance to their overall happiness and satisfaction with the area. The 16 areas of life in the inventory are health, self-esteem, goals and values, money, work, play, learning, creativity, helping, love, friends, children, relatives, home, neighbourhood and community.

**Combat exposure**

The Combat Exposure Scale (CES; Wolfe et al., 1998) was expanded from the original version of the scale (Gallops et al., 1981), which focuses on the presence and frequency of a range of war-zone stressors. Expanded items reflect war-zone events specific to service in the Gulf War, which in some cases extend beyond combat. Given that traditional combat in the Gulf War lasted only 5 days and did not involve all personnel, the expanded items inquired about less traditional war-zone events compared to the original items. The questionnaire was self-administered, and participants responded whether they experienced each event on the list. The responses are 'no' (rated 0), 'once or twice' (1) or 'three or more times' (2). The total score on all 33 items was used as a measure of war-zone stressor exposure.

**Training and quality control**

Research assistants received CIDI training from Dr L. N. Robins and her staff (Dr Robins is the author of the Diagnostic Interview Schedule on which the CIDI is based). Training in use of the CAPS was conducted by Dr Frank Weathers of the Brockton Psychometrics Laboratory, Brockton, M.assachusetts, USA; one of the creators of the interview, and he co-rated 20 CAPS interviews for interrater reliability. One of the authors (R. T.) supervised ongoing ratings and overall quality through weekly calls and periodic reviews.

**Statistical analyses**

Sample size calculations and participant recruitment efforts for this project were based on the predicted prevalence of disease among veterans. The achieved sample size provided 80% power to detect prevalence differences of 2.0% for PTSD (assumed
deployed veterans prevalence 2.8%, non-deployed veterans 0.8%), and a mean difference of 1.5 (s.d. = 10) for the SF-36 mental component scores. Interrater reliability was assessed with intraclass correlations for CAPS continuous variables (frequency and intensity) for symptom clusters and individual symptoms and with kappa coefficients for PTSD diagnoses.

The sampling design is a stratified random sample with unequal probabilities of selection within combinations of the strata: deployment status, gender and duty type (active service v. reserve or National Guard). Therefore, population prevalence or mean estimates for all analyses were obtained using SUDAAN software developed for the analysis of complex survey data (SUDAAN release 9.0, Research Triangle Institute, North Carolina, USA). Sample weights used in SUDAAN were based on the probability of selection combined with the probability of response. For continuous outcomes, t-tests and linear regression models compared mean responses between groups. Logistic regression models were developed for dichotomous and ordinal polytomous outcomes. The covariates considered in the multiple regression models were age, gender, ethnicity (White v. other), years of education (less than 12 years v. 12 years or more), duty type (active v. reservist/National Guard), service branch (army/marine v. navy/air force) and rank (enlisted v. officer). Candidate covariates were deleted for particular models when they caused computational problems preventing model calculation. Odds ratios, 95% confidence intervals and P values are reported for dichotomous outcomes. Comparisons of categorical data and continuous data with adjustment for covariates were based on the Wald F-statistic. For continuous data without adjustments, P values were based on the two-sample t-test.

RESULTS

Participant characteristics
Both the deployed and the non-deployed veteran groups were 78% male. The deployed group (mean age 38.9 years, s.d. = 8.8) were nearly 2 years younger than the non-deployed group (40.7 years, s.d. = 9.6; P < 0.01), were more likely to be African-American, had lower levels of education and were less likely to be married compared with the non-deployed group. The groups did not differ in terms of the percentage still on active duty (deployed, 7.8%; non-deployed, 8.5%), mean income level, 1991 unit component breakdown or 1991 military branch. Those in the deployed group were less likely to be officers compared with the non-deployed group. Differences between the groups reflect true differences between the two cohorts (Kang & Bullman, 2001).

Participation rates
Of the 1996 deployed veterans who were solicited to participate, 53% (1061) were examined; of 2883 non-deployed veterans who were solicited to participate, 39% (1128) were examined. Despite intensive efforts, 12.8% of deployed veterans and 15.2% of non-deployed veterans were not located. In addition, 34.1% of locatable deployed veterans and 45.6% of non-deployed veterans either never returned their participation letter, or an examination could not be scheduled. Because of the lower participation rates by non-deployed veterans, more people in this category (n = 799) were recruited to achieve the desired sample size of 1000 per group.

Participation bias
Historical military service data, obtained in 1991 for all solicited veterans from the US Department of Defense's M anpower Data Center, were used to evaluate participation bias on socio-demographic variables. We compared participants and non-participants in each group (deployed and non-deployed) and then used the Breslow-Day homogeneity of odds ratios test to assess the hypothesis that the odds ratios generated for the two groups were equal.

Participation bias for demographic characteristics is reported in more detail elsewhere (Eisen et al., 2005). Briefly, we found that participants were nearly 2 years older than non-participants, and that White people, women, reservists and National Guard members were significantly more likely to participate. In addition, officers and army personnel were more likely to participate than non-officer and non-army personnel, although these differences were not uniformly significant. We also calculated participation bias for mental health characteristics available from the 1995 survey. Dichotomous data were available on smoking ("Have you smoked in the past 12 months?"), drinking ("Do you drink alcohol?"); depression ("Have you experienced in the past year: difficulty getting to sleep; excessive sleepiness; awaken feeling tired; anxious, irritable, or upset; been depressed or blue; difficulty concentrating or reasoning, memory loss?"); "Have you experienced in the past 6 months: unintended loss of more than 10 pounds, unintended gain of more than 10 pounds?" and PTSD symptom severity (the PTSD Checklist). Participants did not differ from non-participants in each group on smoking, drinking, weight loss or PTSD symptom severity. Participants in each group more frequently reported the other mood symptoms; however, the degree of participation bias did not differ quantitatively between the groups. The only characteristic that yielded a statistically significant difference in the odds ratios generated by the two groups was 1995 active duty status (Eisen et al., 2005). Overall, the degree of participation bias was independent of deployment status.

Reliability of CAPS
Interrater reliability was calculated for 32 continuous CAPS symptom variables (frequency and intensity ratings for symptom clusters B, C and D, and total symptoms, current and lifetime, for two events). Excellent interrater reliability was demonstrated by high (0.90 or higher) intraclass correlations for 31 of these variables and one moderately high (0.86) intraclass correlation (event 1 cluster C intensity). The kappa coefficient was 0.77 for the diagnosis of current PTSD and 0.79 for the diagnosis of lifetime PTSD, indicating good reliability.

Prevalence of all war-related mental disorders
The prevalence of any one Gulf War-era onset mental disorder (i.e. reported initial onset between 1 January 1991 and 30 July 1993) was significantly higher among the deployed veterans than among the non-deployed veterans (Table 1). The broader categories of mood disorders and anxiety disorders were also significantly more prevalent among the former group. Specific disorders within those categories that were significantly more prevalent among the deployed veterans compared with non-deployed veterans were major depression, PTSD, panic disorder and specific phobias. Of the deployed veterans with a PTSD onset in the war-era, in 93% of cases the PTSD was related to a Gulf War event. Two less common disorders also differed significantly between groups: pain disorder was significantly more common in the deployed veteran group, and brief psychotic disorder was significantly more common among the non-deployed veterans.
### Table 1: Prevalence of mental disorders with Gulf War-era onset

<table>
<thead>
<tr>
<th>Disorder</th>
<th>Deployed %</th>
<th>Non-deployed %</th>
<th>P²</th>
<th>OR (95%CI)</th>
<th>Deployed %</th>
<th>Non-deployed %</th>
<th>P²</th>
<th>OR (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Anxiety Disorders</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTSD</td>
<td>9.4</td>
<td>2.4</td>
<td>&lt;0.0001</td>
<td>4.17 (2.36-7.38)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-PTSD Anxiety Disorders</td>
<td>6.2</td>
<td>1.1</td>
<td>&lt;0.0001</td>
<td>5.70 (2.69-12.04)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Generalised Anxiety Disorder</td>
<td>4.3</td>
<td>1.4</td>
<td>0.004</td>
<td>3.17 (1.43-6.99)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agoraphobia without panic</td>
<td>0.9</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panic Disorder</td>
<td>0.7</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specific Phobia</td>
<td>1.2</td>
<td>0.1</td>
<td>0.01</td>
<td>8.93 (1.58-50.31)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Phobia</td>
<td>1.9</td>
<td>0.8</td>
<td>0.12</td>
<td>2.32 (0.80-6.70)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obsessive-compulsive disorder</td>
<td>0.6</td>
<td>0.5</td>
<td>0.72</td>
<td>1.37 (0.24-7.93)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Substance Dependence</td>
<td>7.9</td>
<td>4.8</td>
<td>0.04</td>
<td>1.71 (1.03-2.84)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nicotine Dependence</td>
<td>3.1</td>
<td>2.2</td>
<td>0.35</td>
<td>1.43 (0.68-2.90)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alcohol Dependence</td>
<td>4.3</td>
<td>3.0</td>
<td>0.24</td>
<td>1.48 (0.77-2.87)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Illicit Substance Dependence</td>
<td>0.9</td>
<td>0.6</td>
<td>0.61</td>
<td>1.48 (0.33-6.66)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mood Disorders</td>
<td>7.5</td>
<td>4.1</td>
<td>0.02</td>
<td>1.89 (1.33-2.65)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major Depression</td>
<td>7.1</td>
<td>4.1</td>
<td>0.03</td>
<td>1.78 (1.06-2.99)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dysthymia</td>
<td>0.04</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bipolar Disorder</td>
<td>0.36</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somatic Disorders</td>
<td>1.0</td>
<td>0.3</td>
<td>0.14</td>
<td>3.40 (1.66-7.05)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Somatisation Disorder</td>
<td>0.1</td>
<td>0.01</td>
<td>0.26</td>
<td>4.93 (0.31-78.34)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conversion Disorder</td>
<td>0.03</td>
<td>0.04</td>
<td>0.94</td>
<td>0.91 (0.06-14.46)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain Disorder</td>
<td>0.9</td>
<td>0.01</td>
<td>0.001</td>
<td>68.04 (8.04-757.54)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body Dysmorphic Disorder</td>
<td>0.04</td>
<td>0.3</td>
<td>0.09</td>
<td>0.16 (0.02-1.37)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psychotic Disorders</td>
<td>0.6</td>
<td>0.9</td>
<td>0.54</td>
<td>0.65 (0.17-2.55)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>0.3</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schizoaffective Disorder</td>
<td>0.0</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delusional Disorder</td>
<td>0.04</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brief Psychotic Reaction</td>
<td>0.2</td>
<td>0.9</td>
<td>0.04</td>
<td>0.26 (0.07-0.96)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eating Disorders</td>
<td>0.03</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anorexia Nervosa</td>
<td>0.0</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bulimia Nervosa</td>
<td>0.03</td>
<td>0.0</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>One or more mental disorders</td>
<td>18.1</td>
<td>8.9</td>
<td>&lt;0.0001</td>
<td>2.28 (1.59-3.27)</td>
<td></td>
<td></td>
<td></td>
<td>2.12 (1.44-3.11)</td>
</tr>
</tbody>
</table>

**Note:**
- OR (95% CI): Odds ratio with 95% confidence interval.
- P²: P-value for the difference in prevalence between deployed and non-deployed groups.

**Prevalence of war-related mental disorders 10 years later**

War-era onset disorders that were significantly more prevalent among deployed veterans than among non-deployed veterans and had a prevalence of at least 1% in both groups were examined for their continued presence in the year prior to the examination (i.e. if the disorder began during the war, it was currently active within 1 year of our examination, approximately 10 years later). In the case of recurrent episodes of major depression, there would have been an initial episode in the war-era and an episode in the year prior to the examination, however, the episodes might have come and gone in the intervening period. We combined non-PTSD anxiety disorders into one group to yield a larger event rate for further analyses. Using our whole sample, war-era onset major depression continued to be significantly more prevalent 10 years later in the deployed
veterans group (3.2%) compared with the non-deployed group (0.8%, adjusted P=0.01) (Fig. 1). Comparing the odds ratios representing degree of remission, deployed veterans were 3.33 times more likely to be in an episode of depression 10 years later, whereas non-deployed veterans were significantly more likely to be in current remission (P=0.048). To determine whether severity of initial depression at the time of the war might have differed in the two groups, we examined the severity of depression (mild, moderate, severe) during that period and found that it did not differ between the two groups (P=0.89). We also examined the presence of comorbid psychiatric disorders with war-era onset. There was a trend for deployed veterans with depression to be more likely to have a comorbid psychiatric disorder (45.6%) than non-deployed veterans with depression (25.8%, P=0.11) and the deployed group displayed a trend for having more anxiety disorders, including PTSD (P=0.07).

War-era PTSD prevalence did not differ statistically between deployed (1.8%) and non-deployed (0.6%) groups (adjusted P=0.12) 10 years later, although the rate among deployed veterans was three times higher than in the non-deployed group (Fig. 2). Remission rates did not differ significantly between the groups (P=0.14). Prevalence of war-era onset of non-PTSD anxiety disorders differed significantly between deployed veterans (2.8%) and non-deployed veterans (1.2%; adjusted P=0.01) 10 years later, with the rate among the deployed veterans being twice that among the non-deployed group (Fig. 3). There was no significant difference in degree of remission between the groups (P=0.36).

**Use of medication for depression and anxiety**

We examined medication use at the time of assessment in veterans having war-era onset mental disorders to determine whether those with remitted disorders were being successfully treated with medication. Of the 44 veterans (36 deployed, 8 non-deployed) with war-era onset depression who still met criteria within a year of our assessment, there was no significant difference in the percentage taking any antidepressant medication (deployed 17.1%, non-deployed 33.4%; P=0.43), although nearly twice the number of non-deployed veterans were taking medication. Oft the 89 veterans with war-era onset depression whose depression had remitted, those who had been deployed (13.1%) were taking antidepressant medication at over twice the rate of non-deployed veterans (5.2%), but these rates were not statistically different (P=0.25). Overall, 14.9% of deployed veterans and 10.7% of non-deployed veterans with war-era onset depression took medication for depression at the time of their assessment. For those with any lifetime depression, 20.7% of those who were deployed and 14.2% of those who were not took medication for depression at the time of assessment.

Veterans diagnosed with PTSD and/or other anxiety disorders were examined for their use of both anti-anxiety and anti-depressant medications, as both are commonly prescribed for anxiety disorders. We combined PTSD and other anxiety disorders for this analysis because the total sample size (n=135: PTSD only, n=73; other anxiety only, n=47; both, n=15) was similar to the sample size with war-era onset depression (n=133) and we were concerned that by analysing PTSD (n=88) separately from other anxiety disorders (n=62) the cells would be too small for a meaningful analysis to be performed. Of the 78 veterans (deployed, n=61; non-deployed, n=17) with war-era onset anxiety disorders who still met criteria within 1 year of our assessment, a higher percentage of non-deployed (22.9%) than deployed (12.4%) veterans were taking medication for anxiety, but the rates were not statistically different (P=0.40). Of the 57 veterans with war-era onset anxiety whose disorders had remitted, significantly fewer deployed veterans (4.9%) took medication in comparison with non-deployed veterans (37.4%; P=0.02). Overall, 8.3% of deployed veterans and 26.6% of non-deployed veterans with war-era onset anxiety disorders were taking specific medication for anxiety at the time of assessment. For those with any lifetime anxiety disorder, 17.2% of deployed veterans and 15.4% of non-deployed veterans took medication for depression at the time of assessment. We conducted a Cochran-Mantel-Haenszel Test in SUDAAN using the entire cohort to examine if there was differential use of medication for anxiety in the two study groups. Among the deployed group there was no difference in medication usage between those with and without anxiety disorders, but among the non-deployed group those with anxiety disorders were 6.15 times more likely to take medications than those without an anxiety disorder (P=0.04).

**Prevalence of mental disorders prior to the war and lifetime prevalence**

We examined whether major depression, PTSD and non-PTSD anxiety disorders with an onset prior to the Gulf War (before 1 January 1991) differed between groups. There was no significant difference for depression (deployed 5.9%, non-deployed 7.0%; P=0.97) or PTSD (deployed 3.9%, non-deployed 4.2%; P=0.60), but the deployed group had significantly more non-PTSD anxiety disorders (deployed 12.5%, non-deployed 9.2%; P=0.02). Having had any one mental disorder with
an onset prior to the war did not differentiate the groups (deployed 25.9%, non-deployed 24.6%; P=0.13).

Lifetime prevalence of depression did not differ between the groups (deployed 21.3%, non-deployed 18.2%; P=0.22). Lifetime prevalence of PTSD (deployed 10.8%, non-deployed 6.7%, P=0.01), non-PTSD anxiety disorders (deployed 16.9%, non-deployed 11.0%; P=0.0003) and one or more mental disorders (deployed 43.6%, non-deployed 35.5%, P=0.01) were all significantly higher in the deployed group compared with non-deployed veterans.

Predictors of war-related mental disorders

Given that the deployed veterans group displayed a greater prevalence of Gulf War-era onset depression, PTSD and non-PTSD anxiety disorders compared with the non-deployed veterans, logistic regression was employed to examine predictors of these war-onset conditions. Simultaneous independent variables included deployment, pre-1991 onset of any one mental disorder (representing psychological vulnerability), combat (i.e. war-zone stressor) exposure, and demographic factors (age, gender, ethnicity, educational attainment in 1991, duty type, service branch and rank). For depression, pre-1991 mental disorder (P=0.009), war-zone stressor exposure (P=0.002) and gender (P=0.048) were significant. For PTSD, deployment was the only significant independent variable (P=0.03). For non-PTSD anxiety disorders, deployment (P=0.009), pre-1991 mental disorder (P=0.008), war-zone stressor exposure (P=0.0008), gender (P=0.0002) and duty type (P=0.0499) were all significant.

Self-report of symptoms and quality of life

Participants in the deployed veterans group self-reported more severe current symptoms of PTSD, depression and anxiety at the time of the assessment (Table 2), compared with the non-deployed group. Deployed veterans' mental component summary scores on the SF–36 were significantly lower than the non-deployed veterans’ scores (50.0 v. 53.7), reflecting a more negative self-perception of mental health-related quality of life. Deployed veterans also reported lower levels of general life satisfaction on the QoLI than did non-deployed veterans, with 24.5% of the former group reporting below-average quality of life (v. 15.8%). When we examined groups on the 16 areas assessed in the QoLI, deployed veterans reported significantly less satisfaction with the following seven areas: health (P=0.0001), self-esteem (P=0.02), goals and values (P=0.04), play (P=0.01), learning (P=0.001), love (P=0.03) and children (P=0.049).

DISCUSSION

Although there have been many studies showing greater psychological distress in Gulf War veterans, our review of the
literature as well as a recent review by Stimpson et al (2003) found that the majority of studies assessed mental disorders with questionnaires rather than with structured interviews. In addition, many studies used selective samples. In a large, representative Australian sample of Gulf War veterans, Ikin et al (2004) used the CIDI and found that post-war anxiety, affective and substance use disorders remained elevated in deployed veterans compared with non-deployed veterans a decade after the war. Our study was unique in assessing a large, national sample of US Gulf War veterans using rigorous sampling methods and structured interviews for assessing the prevalence of mental disorders. We assumed that mental disorders with a Gulf War-era onset were related to that war, and we sought to determine whether the prevalence of these disorders a decade after the war differed among deployed versus non-deployed veterans.

We found that the prevalence of war-era onset of mental disorders was significantly higher among deployed compared with non-deployed veterans; in particular, deployed veterans exhibited an increased prevalence of depression, PTSD and non-PTSD anxiety disorders, all of which had a prevalence over 1% in both groups. Ten years later, these cases of depression and non-PTSD anxiety disorders remained significantly more prevalent among deployed compared with non-deployed veterans. Post-traumatic stress disorder was over three times more prevalent among deployed veterans. As evidenced by these continued mental disorders as well as self-reported current symptoms of emotional distress, deployed veterans experienced more psychological distress and mental disorders than the non-deployed veterans both during deployment or immediately after the Gulf War, as well as 10 years later. Our results also indicate that deployment had multiple adverse effects on quality of life (health-related and non-health-related) 10 years later. However, whereas deployed veterans reported statistically significantly lower scores on the SF-36 mental component summary compared with the non-deployed group, the mean group difference of 3.9 fell short of the 4.0 group difference typically used with this measure as an index of clinical significance (Wybrick et al, 1999). Likewise, mean totals on the symptom scales did not reach the clinically impaired range, although deployed veterans were more likely to be in clinically impaired categories on these scales than non-deployed veterans.

Limitations
Our determination of the onset of mental disorders during the Gulf War era was based on the participants' retrospective report, and their recall might have been biased by their symptoms and later experiences. We did not externally validate self-report of exposure to trauma (criterion A). As suggested by Frueh et al (2005), self-report of combat trauma in veterans may be exaggerated. However, our epidemiological sample is more similar to that of Dohrenwend et al (2006) than the treatment-seeking sample of Frueh. Dohrenwend found that self-report of combat exposure correlated highly with record-based evidence of combat exposure. Our use of 'state of the art' assessment of criterion A2 helps ensure the validity of our ratings for the presence of criterion A. Although we demonstrated a lack of participation bias for demographic characteristics and selected queries regarding mood, we did not examine mental disorder diagnoses during the postal and telephone survey portions of our study. Therefore, whether increased mental disorders in the deployed group possibly reflected increased participation among deployed veterans compared with non-deployed veterans is addressed only indirectly through consideration of distress symptoms.

Causes of war-related psychological distress
The increase in psychological dysfunction in the deployed veterans could be due to multiple factors. We examined the contribution of deployment status, pre-war psychological vulnerability, combat exposure and demographic variables on disorders with Gulf War-era onsets. Only deployment predicted war-era onset of PTSD, and deployment also predicted war-era onsets of non-PTSD anxiety disorders. The war-era onsets of non-PTSD anxiety and depressive disorders were also associated with the presence of pre-war psychological vulnerability, higher levels of combat exposure and female gender.

Exposure to traumatic events is a well-recognised hazard of war, and an acknowledged precipitant of psychiatric morbidity (Goldberg et al, 1990a,b; Kulka et al, 1990) and post-war syndromes (Yam et al, 1996). However, because exposure to trauma does not uniformly lead to psychiatric morbidity (Green, 1994; Kessler et al, 1995) or increased symptom complaints (Eisen et al, 1991), individual vulnerability (Kulka et al, 1990; True et al, 1993; Kessler et al, 1995; Bromet et al, 1998; Roy et al, 1998; Shalev et al, 1998; Yehuda et al, 2000) is likely to be an important contributory factor. Consistent with prior research documenting that traumatic stress exposure is associated with disorders such as depression in addition to PTSD (Shalev et al, 1998), psychopathology in our cohort that began during the Gulf War era was not specific to PTSD but encompassed anxiety disorders more generally and depression. The development of these other disorders among deployed veterans might reflect the adverse psychological consequences of exposure to non-traumatic stressors associated with Gulf War participation, such as unexpected career and family disruption related to the rapid activation of an all-volunteer force with significant reservist and National Guard representation, uncertainty regarding anticipated chemical and biological warfare, or other deployment stressors (Nash, 2007). Our findings suggest that pre-existing mental disorders represent an individual vulnerability factor for the development of mental disorders during war deployment.

The classic way in which wars are thought to influence psychological well-being are the effects of exposure to combat-related traumatic stressors on the development of PTSD. Our interest in the more general effects of deployment across mental disorders led us to focus on onset during the interval of deployment rather than specific exposures to war-zone stressors. Thus, the war-era onset of PTSD cases does not overlap completely with traditional combat-related PTSD. Moreover, we did not externally validate self-report of combat experiences, which might have resulted in some overestimation of actual experiences (Dohrenwend et al, 2006; Frueh et al, 2005), nor did we capture combat-related PTSD cases with an onset after 30 July 1993, which might have resulted in underestimation of PTSD cases. Nevertheless, our data demonstrate that deployment to the Gulf War did contribute to a greater onset of major depression, PTSD and other anxiety disorders, with some persisting problems in these veterans 10 years later. Perhaps a broader conceptualisation of the relationship between war deployment
and psychological functioning than the link between combat stress and PTSD would better represent health consequences for our combat veterans. This expanded view might also help to explain how the experience of combat deployment could contribute to a reduced quality of life a decade after the war.

Hobfoll’s stress model of conservation of resources (Hobfoll, 1989) provides one perspective on the multiple impacts of stress. This model posits that individuals, when confronted with stress, try to minimise their net loss of resources. Stress involves a loss or potential loss of resources, including object resources (e.g. a home), conditions (e.g. a marriage, a state of peace), personal characteristics (e.g. an optimistic view of the world) and energies (e.g. time). Deployment is a condition involving loss of resources regardless of whether combat trauma was experienced. Coping itself depletes resources that may or may not offset the original loss of resources, and ‘loss spirals’ may develop when people have limited resources to offset an initial loss, or when coping reduces available resources needed to fend off future losses. For example, in relation to our findings, war stress (traumatic or non-traumatic) involves a loss of resources. Someone with optimistic tendencies has a resource to help buffer the loss of other resources related to war stress. In contrast, someone with pessimistic tendencies may require additional resources to offset their pessimism and insure protection against the development of additional depressive symptoms. For either person, energy spent on offsetting the loss of resources due to deployment may take away from investment in various domains ensuring enhanced quality of life.

Medication use

We examined the likelihood of remission of war-era onset disorders in deployed vs. non-deployed veterans. Depression was less likely to remit among deployed than non-deployed veterans, although the levels of antidepressant use in the two groups were similar. To explore possible explanations of these differential remission rates, we compared the study groups on the severity of the depression during the Gulf War era, as well as on the presence of comorbid war-era mental disorders. The severity of depression in the two groups did not differ. Depressed deployed veterans were twice as likely as depressed non-deployed veterans to have comorbid war-era onset mental disorders. The presence of these comorbid disorders could explain the reduced likelihood of remission of depression in the deployed veterans group. Trivedi et al (2006) identified predictors of remission in out-patients with major depressive disorder. Participants who were White, female, employed, or had higher levels of education or income had higher remission rates for depression, whereas longer index episodes, more concurrent psychiatric disorders, more general medical disorders and lower baseline function and quality of life were associated with lower remission rates. Thus, increased comorbidity and lower levels of education among our deployed veterans could have contributed to lower remission rates of depression. Parker et al (2000) examined 12-month outcome in 182 persons with major depression. Non-recovery at 12 months was predicted most consistently by higher baseline levels of anxiety and depression; high trait anxiety and a lifetime anxiety disorder; disordered personality function; and self-reported exposure to acute and enduring stressors at baseline assessment. The more complicated clinical presentation of multiple comorbidities associated with exposure to war stressors in our deployed veterans group might have contributed to their continued depression compared with the non-deployed group.

Anxiety disorders were equally likely to remit in the two groups, but non-deployed veterans with anxiety in remission were more likely to be taking medication than deployed veterans in remission. Indeed, in the entire cohort, non-deployed veterans with anxiety disorders were more likely to take medication for these disorders than deployed veterans with anxiety disorders. The reason for this treatment disparity is unclear and we do not know whether the deployed veterans were less likely to be prescribed medication for anxiety disorders or were less likely to take anti-anxiety medication compared with the non-deployed group.

Implications of the study

The prevalence of Gulf War-era depression, anxiety - and even PTSD to some extent - abates with time. The prevalence of all these conditions decreased among both our study groups 10 years after the Gulf War compared with the rates that were found in immediate proximity to the war. Continued depression in deployed veterans appears partially resistant to remission despite comparable levels of medication use in the two groups; however, anxiety disorders might possibly remit further in deployed veterans with greater use of medications. The presence of comorbid psychiatric disorders may make it less likely that depression will remit. We do not know the extent to which the groups might or might not have differed in their use of psychological treatments for these conditions, but the findings point to the need for adequate follow-up mental healthcare for veterans with persistent mental illnesses following major military operational deployments.

ACKNOWLEDGEMENTS

We acknowledge the support of the Cooperative Studies Program of the Department of Veterans Affairs Office of Research and Development and the work of the investigators, nurse coordinators and research assistants at the 16 participating Veterans Affairs Medical Centers (VAMC); the staff at the Hines Cooperative Studies Program Coordinating Center, Hines, Illinois, and the staff at the Cooperative Studies Program, Department of Veterans Affairs Central Office, Washington, DC; the members of the executive committee and the Data and Safety Monitoring Board; and the study project manager (Tammy Nydam, MA, Hines VAMC) and coordinators (Joyce Hunter, RN, MSN, St Louis VAMC; Leila Jackson, DrPH, VA Central Office).

We acknowledge the staff at the Brockton Psychometrics Laboratory, Brockton, Massachusetts: Lisa James and Zachary Warren, research assistants; Michael Lyons, PhD, Frank Weathers, PhD, and Marie Walbridge, PhD, consultants. Finally, we thank the participating veterans and their families for their time and effort.

REFERENCES


Ikin, J. F., Sim, M. R., Creamer, M. C., et al. (2004) Psychosocial, psychological and physical health outcomes program and Department of Preventive Medicine and Biometrics, Aurora, Colorado; Frances M. Murphy, MD, Veterans Health Administration, Department of Veterans Affairs, and Uniformed Services for the Health Sciences, Washington, DC; Seth A. Eisen, MD, MSc, Medical and Research Services, St Louis Department of Veterans Affairs Medical Center and Department of General Medical Sciences, Washington University School of Medicine, St Louis, Missouri, USA.


(First received 15 November 2005, 21 August 2006, accepted 6 November 2006)
Common mental disorders and the built environment in Santiago, Chile*

RICARDO ARAYA, ALAN MONTGOMERY, GRACIELA ROJAS, ROSEMARIE FRITSCH, JAIME SOLIS, ANDRES SIGNORELLI and GLYN LEWIS

**Background** There is growing research interest in the influence of the built environment on mental disorders.

**Aims** To estimate the variation in the prevalence of common mental disorders attributable to individuals and the built environment of geographical sectors where they live.

**Method** A sample of 3870 adults (response rate 90%) clustered in 248 geographical sectors participated in a household cross-sectional survey in Santiago, Chile. Independently rated contextual measures of the built environment were obtained. The Clinical Interview Schedule was used to estimate the prevalence of common mental disorders.

**Results** There was a significant association between the quality of the built environment of small geographical sectors and the presence of common mental disorders among its residents. The better the quality of the built environment, the lower the scores for psychiatric symptoms; however, only a small proportion of the variation in common mental disorder existed at sector level, after adjusting for individual factors.

**Conclusions** Findings from our study, using a contextual assessment of the quality of the built environment and multilevel modeling in the analysis, suggest these associations may be more marked in non-Western settings with more homogeneous geographical sectors.

**Declaration of interest** None.

There is growing interest in investigating whether contextual variables, such as those representing the built environment, can influence the prevalence of common mental disorders after accounting for individual variables (Weich, 2005). The built environment encompasses all those aspects of our habitat that are created or modified by people, such as homes, schools, parks and roads (Srinivasan et al, 2003). Up to now most studies of the built environment and mental illness have used indirect environmental measures such as individuals’ perceptions (Ellaway et al, 2001) or aggregated data (Ross, 2000; Weich et al, 2003). A few studies have assessed the built environment directly and independently (Weich et al, 2002; further details available from the authors), finding little or no variation in prevalence of common mental disorders across small or medium-sized areas (Weich, 2005).

We are unaware of any other Latin American study assessing the contextual effect of the built environment directly and using multilevel models to investigate its association with mental illness. We tested the hypothesis that contextual measures reflecting the quality of the built environment in Santiago, Chile would be associated with common mental disorders independent of individuals’ characteristics.

**METHOD**

Santiago, the capital of Chile, has a population of approximately 6 million people, representing 42% of the total Chilean population. Although Chile is considered to be a middle-income country, with a gross per capita annual income of £2600, it is one of the ten most unequal countries in the world in terms of income (World Bank, 2001). The city is geographically extended, with a mean population density of 392 persons per square kilometre. Geographically the city is neatly compartmentalised according to socio-economic groups, with wealthier people living mainly in the eastern suburbs and the poorest in the southern and northern fringes. Most of the city has basic amenities, including electricity, sanitation and drinkable water, but there are visible differences between more and less affluent sectors. Houses in the wealthier areas are bigger and of better quality. Facilities including better roads and pavements (sidewalks), green areas, overall cleanliness and abundant shops are noticeable. However, crime seems to be present in all sectors, and crime reported to police is especially common in wealthier sectors of the city because criminals target these areas and people report more incidents for various reasons, including private insurance claims.

**Sampling strategy**

A cross-sectional survey was conducted between 1996 and 1998. The sampling framework was the adult population, restricted to ages 16–64 years, living in private households in the Greater Santiago area. The sampling strategy involved a three-stage design, which included all the 35 boroughs of Greater Santiago, 248 sectors and 4300 households randomly selected with a probability proportional to the size of the sampling units. The number of households within each sector varied from 26 to 5. One person per household was chosen randomly for interview using Kish tables (Kish, 1965). Individuals from sectors with fewer than five observations were excluded from this analysis. Responses were obtained from 3870 household respondents (response rate 90%). Further details of the sampling design can be found elsewhere (Araya et al, 2001).

**Mental health, social and demographic questionnaire**

Psychiatric symptoms were assessed with the Revised Clinical Interview Schedule (CIS-R; Lewis et al, 1992), a structured and detailed psychiatric interview used extensively in primary care and community studies in Chile and elsewhere. This interview has 14 items assessing the severity of the most common psychiatric symptoms. Each item is given a score, which can then be summed to yield a total score. This continuous measure, reflecting psychiatric symptom severity, was used as our main outcome. The mean weighted K across all

---

*Freely available online through the British Journal of Psychiatry open access option.*
approximately ten small contiguous streets whose maps were prepared by the INE.

The main purpose was to try to create a score that reflected the desirability or attractiveness of a sector. Although many characteristics of a local sector are likely to affect satisfaction, the study focused on those that could be easily and reliably assessed by a walk through the area. Most items were derived from the Residential Environment Assessment Tool (REAT; Dunstan et al., 2005) used to measure area characteristics in Wales. Some additional items were included because they were thought to be important locally, such as the presence of stray dogs or bad odours. The final instrument (available from the authors) included items relating to the following characteristics, with the number of items in parentheses: public lighting (2), state of roads (6), sidewalks (4), public green areas (5), green elements on sidewalks and front gardens (4), dirtiness (1), traffic and noise (2), bad odours (1), general maintenance of properties (1), general use of the sector (4), empty sites (2), external beautification (1), presence of homeless people (1), presence of stray dogs (1), access to properties (2), balconies (1), street signs (1), public transport (4), security and safety devices (6), and a list of facilities including:

(a) essential facilities: primary and secondary schools, other type of schools, creches, primary care clinics, hospitals and private clinics;
(b) leisure facilities: public gymnasia, swimming pool, football pitch, sports club, cultural centre, library, community centre, corner shops, pharmacies, cinema, theatre, restaurants, coffee shops, bars;
(c) Other facilities: petrol station, petrol station shops and kiosks.

Some of the items required dichotomous ratings, for example the presence of lamp-posts; others had a range of possible values that could be ordered from high to low, such as the level of maintenance of front gardens. Ratings for all items were converted into a score between 0 and 1, with the value of 1 always representing either a more desirable feature (e.g. cleaner roads) or more of that particular item (e.g. more essential facilities). For example, the item for 'beauty of front garden' was initially coded 1, 2 or 3 and subsequently recoded as 0, 0.5 or 1, with 1 representing 'very beautiful' front gardens. The presence of any of the listed facilities in the sector was rated as '1' and individual scores were summed to generate three facility indices. These total scores for the facilities indices were subsequently recoded into three scales with scores ranging from 0 to 1.

Ratings were made for observations of the sector as a whole, for example the extent of street litter within the sector rather than in any particular street. Operational definitions were provided for each item, including a set of photographs to illustrate different degrees of items such as litter. The tool was initially piloted in a few neighbourhoods in Santiago of varying socio-economic status. Raters were university students of architecture and psychology trained over 2 days to ensure consistency in their understanding and ratings of items. During this training, instructions were given and photographs used to discriminate between ratings.

Two independent raters assessed each geographical sector. Each one scored all items, and after the assessment both raters met and agreed on a consensus rating, which is the one used in this study. There were ten pairs of raters in all. The walkthrough assessment was undertaken over a period of 30 days in January 2002. All assessments were conducted in the mornings in order to avoid potential differences due to timing, for instance in ratings of noise. Raters were given a map of the sector with clear boundaries and instructions to follow a predetermined route, specified in the map, when walking through the area and to cover all streets in the sector. The assessment of each sector took approximately 60 min to complete. In certain areas it was necessary to arrange for someone to accompany the raters for reasons of safety, but the raters were instructed not to talk about the neighbourhood to anyone until the ratings had been completed.

**Statistical analyses**

**Composite score for the BEAT scale**

All variables with 95% or more of respondents in one category were eliminated. Subsequently we performed factor analysis with varimax rotation of all the remaining items to assess if and how these items loaded into common factors. Cronbach's α was estimated for all the items in the scale and for the items within each one of the newly derived factors after the factor analysis. All variables were entered, including the facilities indices, into the factor analysis model and those with loadings lower than
0.4 after varimax rotation or high uniqueness values were also excluded. The scores of each item (0, 1) within a factor were added to generate a total composite score for the factor. A higher score in the factor reflected a more attractive area. We generated an alternative total score applying weights to the individual item scores according to the loadings of that particular item in the factor analysis. The total unweighted and weighted scores for each factor were compared using correlation coefficients to explore for possible differences, depending on the methods used to estimate total factor scores. Finally, for all regression models, the four factor scores were rescaled to a minimum of 0 and a maximum of 10 to facilitate comparison of effect sizes between the factors.

**Testing associations across levels**

Associations between sector (level 2) and borough (level 3) exposures and individual CIS-R scores were investigated using multilevel regression models (MLwiN version 2.02, Institute of Education, University of London, UK, and Stata Release 9), before and after adjusting for individual (level 1) predictor variables. All analyses excluded individuals from sectors with fewer than five replies and those with incomplete data for any of the individual, sector-level and borough-level variables to be included in the models. We did this because we felt that using data from only a handful of individuals might not offer a fair representation of the sector. Individuals included and excluded from the analyses were compared in terms of age, gender and socio-economic characteristics. The modelling strategy consisted of first fitting a simple variance components null model to quantify the three components of residual variation in CIS-R score as a continuous variable: borough, sectors and individuals. Analyses involving CIS-R score as a continuous outcome were based on a normally distributed multilevel model using the observed and log-transformed CIS-R scores. Estimation in all models was based on iterative generalised least squares. As the associations between individual-level variables and mental health are already well known, the primary aim of this study was to investigate the effects of the local and wider neighbourhood on mental health using the CIS-R total scores, after taking individual factors into account. Therefore the modelling strategy we adopted was to investigate sector-level predictors of CIS-R first, then to add individual and finally borough-level variables to the model and to note changes in the components of variance and coefficients for the sector-level factors. We investigated whether there were any differential associations between factor 1 and CIS-R for categories of selected individual variables by fitting appropriate interaction terms in the regression models.

**RESULTS**

**Development of the Built Environment Assessment Tool**

Twenty-seven items were removed from the original list for two reasons: 14 did not show sufficient ability to discriminate (95% or more of the answers fell into one category) and the remainder had loadings below 0.4 after varimax rotation or high uniqueness values so that they did not fit well with any of the factors. Among the items left out were several questions on the type of roads, external beautification of properties, parking on sidewalks, predominant type of properties, protection on balconies, security fences, presence of taxis, bad odours, the size of green areas on sidewalks, and presence of vagrants. Twenty-five items remained, including three representing the sum of the list of facilities. These were grouped into four main factors:

- (a) general quality of the area;
- (b) facilities, noise and traffic in the area;
- (c) public green areas;
- (d) empty sites.

These four factors all had eigenvalues over 1 and together explained 90% of the total variance (Table 1). The mean

<table>
<thead>
<tr>
<th>Items</th>
<th>Factor number</th>
<th>Factor 1 General quality</th>
<th>Factor 2 Facilities</th>
<th>Factor 3 Green areas</th>
<th>Factor 4 Empty sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion variance, %</td>
<td>50.1</td>
<td>20.3</td>
<td>12.6</td>
<td>8.2</td>
<td></td>
</tr>
<tr>
<td>Eigenvalues</td>
<td>6.2</td>
<td>2.5</td>
<td>1.5</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Loadings presented after varimax rotation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Width of sidewalks</td>
<td>1</td>
<td>0.45</td>
<td>-0.24</td>
<td>-0.18</td>
<td>0.29</td>
</tr>
<tr>
<td>General maintenance of sidewalks</td>
<td>1</td>
<td>0.69</td>
<td>-0.012</td>
<td>0.07</td>
<td>-0.04</td>
</tr>
<tr>
<td>Additional features on sidewalks</td>
<td>1</td>
<td>0.71</td>
<td>-0.02</td>
<td>0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>(wheelchairs, bicycles, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State of front gardens</td>
<td>1</td>
<td>0.70</td>
<td>0.03</td>
<td>0.03</td>
<td>-0.01</td>
</tr>
<tr>
<td>Trees on sidewalks</td>
<td>1</td>
<td>0.54</td>
<td>-0.07</td>
<td>-0.01</td>
<td>0.16</td>
</tr>
<tr>
<td>Size of trees on sidewalks</td>
<td>1</td>
<td>0.44</td>
<td>-0.17</td>
<td>-0.13</td>
<td>0.16</td>
</tr>
<tr>
<td>Green area on sidewalk</td>
<td>1</td>
<td>0.74</td>
<td>-0.15</td>
<td>-0.04</td>
<td>0.27</td>
</tr>
<tr>
<td>Dirtiness of street</td>
<td>1</td>
<td>0.63</td>
<td>-0.00</td>
<td>0.18</td>
<td>-0.08</td>
</tr>
<tr>
<td>General maintenance of properties</td>
<td>1</td>
<td>0.67</td>
<td>-0.11</td>
<td>0.17</td>
<td>-0.07</td>
</tr>
<tr>
<td>Type of properties</td>
<td>1</td>
<td>0.41</td>
<td>-0.14</td>
<td>0.08</td>
<td>0.18</td>
</tr>
<tr>
<td>Stray dogs</td>
<td>1</td>
<td>0.57</td>
<td>-0.08</td>
<td>0.19</td>
<td>0.14</td>
</tr>
<tr>
<td>Signs for orientation</td>
<td>1</td>
<td>0.41</td>
<td>-0.37</td>
<td>-0.10</td>
<td>0.05</td>
</tr>
<tr>
<td>Other public signs</td>
<td>1</td>
<td>0.67</td>
<td>-0.16</td>
<td>0.07</td>
<td>0.12</td>
</tr>
<tr>
<td>Security badges on houses</td>
<td>1</td>
<td>0.41</td>
<td>-0.29</td>
<td>-0.24</td>
<td>0.25</td>
</tr>
<tr>
<td>Guards</td>
<td>1</td>
<td>0.42</td>
<td>-0.24</td>
<td>-0.08</td>
<td>0.33</td>
</tr>
<tr>
<td>Level of traffic</td>
<td>2</td>
<td>-0.05</td>
<td>0.80</td>
<td>0.01</td>
<td>-0.17</td>
</tr>
<tr>
<td>Noise of traffic</td>
<td>2</td>
<td>0.02</td>
<td>0.71</td>
<td>-0.02</td>
<td>-0.06</td>
</tr>
<tr>
<td>Bus stop</td>
<td>2</td>
<td>0.08</td>
<td>0.54</td>
<td>-0.18</td>
<td>0.09</td>
</tr>
<tr>
<td>Essential facilities</td>
<td>2</td>
<td>0.27</td>
<td>0.41</td>
<td>-0.18</td>
<td>-0.03</td>
</tr>
<tr>
<td>Leisure facilities</td>
<td>2</td>
<td>0.12</td>
<td>0.52</td>
<td>-0.15</td>
<td>-0.01</td>
</tr>
<tr>
<td>Other facilities</td>
<td>2</td>
<td>0.14</td>
<td>0.47</td>
<td>-0.11</td>
<td>0.16</td>
</tr>
<tr>
<td>Public green areas</td>
<td>3</td>
<td>-0.11</td>
<td>-0.12</td>
<td>0.78</td>
<td>0.12</td>
</tr>
<tr>
<td>State of public green areas</td>
<td>3</td>
<td>0.28</td>
<td>-0.10</td>
<td>0.74</td>
<td>0.09</td>
</tr>
<tr>
<td>Presence of empty sites</td>
<td>4</td>
<td>0.08</td>
<td>0.03</td>
<td>0.08</td>
<td>0.80</td>
</tr>
<tr>
<td>Empty sites occupied illegally</td>
<td>4</td>
<td>0.11</td>
<td>0.07</td>
<td>0.08</td>
<td>0.82</td>
</tr>
</tbody>
</table>
Cronbach’s $\alpha$ for the items in the scale was 0.87. There were small differences between these values of $\alpha$ and those generated using only the items within each factor. Lower $\alpha$ values for the items contained in factors with lower eigenvalues are explained because the first factor with the highest eigenvalue contains the most items, thus making the greatest contribution to the overall variation in scores. Kappa coefficients for items between pairs of interviewers fluctuated from 0.69 to 0.92, with 78% of the estimated $\kappa$ coefficients above 0.85 and full agreement for 70% of the items. Simply summing items to get a total factor score assumes equal weighting of each item and that ‘non-loading’ items are not important. For this reason we compared weighted (according to eigenvalues) and unweighted scores for each of the factors. We found high correlations between these two different ways of scoring the factors (correlation coefficients for factors 1, 2, 3 and 4 were 0.99, 0.87, 0.92 and 0.88 respectively), and therefore we decided to use the simple, unweighted scores in all analyses. Correlation coefficients between the four sector-level factors are presented in Table 2.

Characteristics of the sample surveyed
A total of 3870 interviews were completed to give a response rate of 90%. These 3870 individuals clustered into 248 sectors and 35 boroughs. A total of 488 individuals (12.6%) were excluded because of missing data or because they lived in a sector with fewer than five respondents, leaving 3382 observations from 210 sectors within 31 boroughs for analysis. Excluded individuals were no different to those included in terms of age ($P=0.56$), gender ($P=0.17$) or marital status ($P=0.59$), but had lower median income (in Chilean pesos, CLP62 500 v. CLP100 000, $P<0.0001$), were less likely to be educated to university level (20 v. 36%, $P<0.001$), were more likely to live in very poor or poor quality housing (22 v. 15%, $P<0.001$), had fewer supportive individuals (3.7 v. 4.2, $P=0.01$) and have lower alcohol consumption (1.5 v. 1.8, $P=0.005$). Excluded individuals also had higher mean CIS–R scores (8.5 v. 7.2, $P<0.001$). There was no evidence that excluded sectors were any different from those included in terms of the four factors generated from the BEAT scores or violent incidents reported to police. The number of individuals per sector in the final data-set ranged from 5 to 26 and the number of sectors per borough from 2 to 26. Characteristics of the individuals, sectors and boroughs are presented in Table 3.

Variance components null model for common mental disorder
Mean CIS–R score for the total sample was 7.19 (s.d. = 8.00, range 0–49). We estimated that approximately 5.6% (95% CI 1.8–9.4) of the residual variation in total CIS–R score lies at the borough level, 3.8% (95% CI 1.8–5.7) at the sector level and 90.6% (95% CI 86.2–95.1) at the individual level (Table 4). In view of the non-normal distribution of CIS–R scores we also undertook all multilevel modelling using log-transformed scores and the results in all these models were almost identical to those presented here (further details available from the authors).

Effect of including individual, sector and borough characteristics in the model
Sector-, individual- and borough-level fixed effects were added to the null model in a cumulative manner (Table 4). Inclusion of the sector-level exposures reduced the total residual variance. The estimated percentage of residual variation at borough level decreased to 0.13%, whereas the percentage residual variation at sector (3.55%) and individual (96.32%) levels remained similar and increased respectively. Additional inclusion of individual-level variables reduced the overall variance further, with none of the residual variation now explained at the borough level. Estimates remained unchanged after the addition of borough-level variables. Estimated associations between sector-level factors and mental health were therefore based on a simpler, two-level model that included only sector (level 2) and individual (level 1) variables.

### Table 2 Pearson correlation coefficients of the scores in the four derived factors of the Built Environment Assessment Tool

<table>
<thead>
<tr>
<th>General quality</th>
<th>Facilities</th>
<th>Green areas</th>
<th>Empty sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>General quality</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facilities</td>
<td>0.11</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Green areas</td>
<td>0.31</td>
<td>-0.03</td>
<td>1.00</td>
</tr>
<tr>
<td>Empty sites</td>
<td>0.10</td>
<td>0.07</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Effect of neighbourhood quality on common mental disorders
Table 5 shows the crude and adjusted associations between each of the sector-level factors (rescaled so that the possible range for each is 0–10) and CIS–R total score. After adjusting for other sector-level and individual-level variables, factor 1 (overall quality of the built environment) was inversely associated with total CIS–R score that is, there was strong evidence that individuals living in sectors with more desirable features such as better roads or more green areas had better mental health, after taking into account individual characteristics. There was also a significant association with factor 4 in the adjusted model only, but this was in the opposite direction; higher factor scores were associated with higher CIS–R scores.

We tested interactions between factor 1 and the following individual variables: gender, income and education. The only significant interaction was for gender and factor $1 (P=0.03), 95% CI$ $-0.58$ to $-0.05, P=0.02$ in which male respondents living in less desirable areas had significantly lower CIS–R scores than female respondents. No significant interaction was found for income ($0.004, 95% CI$ $-0.001$ to $0.008, P=0.10$) or education (secondary, $0.23, 95% CI$ $-0.26$ to $0.72$; university, $0.16, 95% CI$ $-0.35$ to $0.68$; overall $x^2=0.91, d.f.=2, P=0.63$).

Although our primary interest was to investigate the association of these factors with mental health, which was best represented by the continuous distribution in CIS–R total scores, we also explored associations with the most common ICD-10 disorders (World Health Organization, 1992), anxiety and depressive disorders, using logistic regression models. There were 154 (46.6%) cases of depression and 309 (9.1%) of anxiety. There was no evidence of any association with depression for factors 1, 2, or 4 (factor 1, OR=1.00,
Table 3  Characteristics of individuals, sectors and boroughs in the sample

<table>
<thead>
<tr>
<th>Variable</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuals (n = 3382)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, years; mean (s.d.)</td>
<td>36.9 (13.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income per capita in household; median (IQR)</td>
<td>100 000 (50 000–250 000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of supportive people; median (IQR)</td>
<td>3 (2.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Units of alcohol consumed daily; mean (s.d.)</td>
<td>1.8 (2.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>1358 (40)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>2024 (60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self-rated presence of disease, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>2782 (82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>600 (18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education level, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>559 (17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td>1607 (48)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>University</td>
<td>1216 (36)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married/cohabiting</td>
<td>1904 (56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Widowed</td>
<td>127 (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Separated</td>
<td>259 (8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>1092 (32)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Housing type, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very poor</td>
<td>125 (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>357 (11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>1474 (44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good</td>
<td>1188 (35)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luxurious</td>
<td>228 (7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sectors (n = 210)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scores; mean (s.d.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor 1: General quality (possible range 0–15)</td>
<td>8.97 (2.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor 2: Facilities (possible range 0–6)</td>
<td>2.43 (0.55)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor 3: Green areas (possible range 0–2)</td>
<td>0.97 (0.77)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Factor 4: Empty sites (possible range 0–2)</td>
<td>1.66 (0.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Episodes of violent crime reported to local police; median (IQR)</td>
<td>1.8 (0.7–5.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boroughs (n = 31)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education budget per capita; median (IQR)</td>
<td>146 404 (115 264–224 365)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health budget per capita; median (IQR)</td>
<td>17 037 (14 558–22 536)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of social organisations per population; mean (s.d.)</td>
<td>0.0018 (0.0013)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IQR, interquartile range.
1. Chilean pesos.

Table 4  Components of variance in total Revised Clinical Interview Schedule score (as a continuous variable) at the individual, sector and borough level: multilevel modelling

<table>
<thead>
<tr>
<th>Variance (s.e.)</th>
<th>Null model</th>
<th>Model 1 (null+sector variables)</th>
<th>Model 2 (model 1+individual variables)</th>
<th>Model 3 (model 2+borough variables)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 3 (borough)</td>
<td>3.59 (1.25)</td>
<td>0.08 (0.21)</td>
<td>0.00 (0.00)</td>
<td>0.00 (0.00)</td>
</tr>
<tr>
<td>Level 2 (sector)</td>
<td>2.42 (0.64)</td>
<td>2.15 (0.60)</td>
<td>0.51 (0.36)</td>
<td>0.50 (0.36)</td>
</tr>
<tr>
<td>Level 1 (individual)</td>
<td>58.28 (1.46)</td>
<td>58.33 (1.46)</td>
<td>51.29 (1.28)</td>
<td>51.29 (1.28)</td>
</tr>
</tbody>
</table>

1. Sector variables in model 1 were the four derived factors from factor analysis.
2. Individual variables were age, gender, presence of disease, income, education, marital status, housing type, number of supportive individuals, and alcohol use. Model 2 also included the sector-level variable episodes of violent crime reported to local police.
3. Borough variables were education budget per capita, health budget per capita, and number of social organisations.

DISCUSSION

This is the first large population-based study of common mental disorders and the built environment of small geographical sectors of a Latin American city using a directly assessed contextual measure and multilevel modelling in the analysis. We found strong evidence of an association between the quality of the built environment in these sectors and common mental disorders, before and after adjusting for individual variables. However, in line with previous reports, the contribution of these sectors to the total variance in common mental disorders was small and most of it was explained by individual factors. None the less, these results represent probably some of the most persuasive evidence found so far to establish an association between the quality of our surrounding built environment and the presence of common mental disorders.

The Built Environment Assessment Tool

We developed and tested a quick and reliable method to assess the built environment using a walk-through method. The great majority of the items reflected the built...
environment, but there were a few - such as the presence of stray dogs - that represented the observable residential environment rather than something built. We had previous experience with developing a similar instrument for a study in South Wales (Dunstan et al., 2005) and we studied carefully other similar instruments (Sanson et al., 1997; Cohen et al., 2000; Weich et al., 2002; Hembree et al., 2005). Our measure showed good interrater reliability and internal consistency; in the absence of a gold standard, however, it is difficult to assess its criterion validity. Overall, the psychometrics of this tool are comparable with those found for the two similar tools developed in the UK (Weich et al., 2001; Dunstan et al., 2005).

### Variance in mental disorders according to geographical aggregation

Studies in Western countries using multi-level models to estimate area-level variation in common mental disorders have found little or no variation at higher levels of aggregation, after accounting for individual differences (Weich, 2005). The contribution of smaller area effects to the total variance in common mental disorders usually fluctuates between 0.5% and 4% before adjusting for individuals' characteristics, and drops to levels between 0% and 1% after adjustment (Weich, 2005). Our findings are in keeping with the higher end of previous estimates; we found that 3.8% of the variance in common mental disorders was explained at the small sector level in the unadjusted models, reducing to nearly 1% in the adjusted models.

As eloquently argued by Weich (2005), there may be a number of reasons to explain this lack of positive findings. For instance, sectors that are large or heterogeneous tend to yield negative results. However, in a previous UK study in South Wales using small geographical units (postcode areas with approximately 150 people) we also found little variation at this level (further details available from the authors). It is possible that using geographical units identified on the basis of an arbitrary geographic classification, which might not reflect neighbourhood unity, may influence the results. In this respect, Rejneveld et al. (2000) found that the clustering of common mental disorders was higher at neighbourhood level (sectors with similar types of building delineated by natural boundaries) than at postcode level using arbitrary geographical boundaries. We tried to deal with both of these possibilities, so we used small geographical sectors of approximately 300 people that were sufficiently homogeneous in terms of their neighbourhood. Our study did not measure the outcome (CIS-R) aggregated at household level and so it is possible that some of the variance found at higher (sector) or lower (individual) levels might reflect variance present at household level.

In our previous study in Wales we found that 37% (95% CI 27-49) of variance existed at household level (further details available from the authors). Although the CIS-R variance at borough level appears greater than at sector level in Table 4, the confidence intervals of these estimates (5.6%, 95% CI 1.8-9.4%) at the borough level and 3.8%, 95% CI 1.8-5.7 at the sector level) show that one cannot reach this conclusion. More importantly, once adjustment for other variables are introduced this borough variance comes close to nil but the sector variance remains only slightly attenuated. Yet the variance we found at borough level in adjusted models is considerable in comparison with other studies. The most likely explanation is that Santiago, like other cities in Latin America, is quite compartmentalised in terms of the quality and socio-economic status of the
geographical areas, with little variation within but more variation between boroughs.

Previous studies have been criticised because they tend to rely entirely on brief psychiatric self-reported questionnaires to measure the outcome. Our study used a detailed structured psychiatric interview to overcome this limitation. So, as it stands, we have to conclude that there seems to be little variation in prevalence of common mental disorder explained at area level and much of this variance resides at individual level. However, even if small sectors contribute little to this overall variance, is it still possible that some features of these areas may be associated with common mental disorders?

Quality of residential environment and common mental disorder

There have been only a handful of mental health studies that have used truly contextual and independent measures of the built environment throughout the world. In the UK there have been only two such studies. Weich et al. (2002) found a significant association between the prevalence of depression and properties with predominantly deck access (OR = 1.28, 95% CI 1.03-1.58) and of recent construction (OR = 1.43, 95% CI 1.06-1.91). It is worth noting that this was a cross-sectional study in which no multilevel modelling was used in the analysis (Weich et al., 2002). Using a similar contextual assessment of the built environment as in the study we report here and multilevel modelling in the analysis we did not find any significant association between the total score of an index depicting the neighbourhood and built environment. Nevertheless, our study used brief questionnaires to measure mental disorder and studied smaller samples than in this study (Weich et al., 76 sectors, n = 1887; our previous study, 51 sectors, n = 1500). A larger number of sectors could help to improve the accuracy of the estimates and provide greater power to test smaller effects.

We found strong evidence of an association (P < 0.05) between two factors of our index of quality of the built environment (BEAT) and common mental disorders, after adjusting for individual differences. These factors represented almost two-thirds of the total variance in the quality of the built environment and thus one can confidently conclude that they are good indicators of the built environment in the city of Santiago. We used a similar method as in the South Wales study (REAT; Dunstan et al., 2005), but there were some differences that might help explain the discrepant results. The BEAT assessed mainly the more private built environment such as houses, gardens or housing density. The BEAT assessed extensively other aspects of the built environment such as roads, pavements and public facilities. The BEAT provided a total score reflecting the quality of the residential environment, whereas we used four factors with their corresponding individual scores. Although it may seem intuitive that a better built environment might help us feel better, the precise mechanism by which the built environment influences our mental health is still a matter of conjecture.

Why did only two factors show significant associations in our study? The first factor represented the largest proportion of the variance and it was the most comprehensive indicator of the quality of the neighbourhood and built environment. Although the relative contribution of this factor to change in CIS-R score is approximately ten times smaller than that associated with individual variables such as being female, it is a factor amenable to change and it is widely spread. Interestingly, the only significant interaction across levels showed that women were more affected (higher CIS-R scores) than men when living in less desirable areas. This would be in keeping with our hypotheses because the women – especially those who did not work outside the home – probably spent more time in the areas studied than men and were therefore more exposed.

We found, rather surprisingly, that factor 4 (empty sites) was associated in the opposite direction: fewer sites were associated with higher CIS-R scores. However, factor 4 was not a key indicator of the area environment, contributing only 8% of the variance in our factor analysis, and in the unadjusted model (see Table 5) this association was not significant at a 5% level. Our assumption was that fewer empty sites, especially if they were unoccupied, would be a good feature of the sector; however, it is possible that our assumptions were baseless and that empty sites in Santiago may not represent abandoned, derelict places where rubbish accumulates, as in other settings. We expected that factor 2, representing 20% of the total variance, would be significantly associated with CIS-R scores. However, this factor was a rare combination of essential and leisure facilities and noise and traffic in the area. Our assumption here was that an increased number of facilities would represent an asset for the locality, but it may be that more facilities bring more noise and traffic to the area and that this is more important. Nevertheless, overall it is reassuring that the strongest and clearest association is for the best and most comprehensive indicator of the quality of the built environment. When we explored associations of these factors with ICD-10 categorical disorders the results were puzzling. We found that there was no association between these disorders and factor 1, representing the overall quality of the neighbourhood. Even more surprisingly, individuals who were depressed were more likely to live in areas with more public green areas, an association that we did not find when using CIS-R total scores. More in keeping with the other results, individuals living in areas with fewer empty sites were less likely to have an anxiety disorder, an association that we found for CIS-R total scores but in the opposite direction. It is difficult to find a reasonable explanation for these disparate findings, especially for those related to depressive disorders. However, our interest was to focus on population changes in mental health (symptom scores) rather than concentrate on specific subgroups, mainly because the former approach would be more informative for public health decision-makers (Rose, 1993).

Santiago is fairly well compartmentalised according to socio-economic grouping. Wealthy people live in areas completely removed from the areas where poorer people live, something not always found in UK cities with a much more mixed socio-economic distribution within geographical sectors. This clear and distinct geographical distribution might have helped reduce ‘contamination’ and accentuated the differences between the sectors selected in our clustered sampling strategy. We selected the sectors in our sample to represent an adequate spread of neighbourhood deprivation, so we expected this would ensure an adequate spread of residential quality. We think that a drop of one point in the total CIS-R score attributable to living in the sectors with better built environment quality is a meaningful change, bearing in mind the large proportion of people who might potentially benefit from interventions.
to reduce this difference. When a common threshold of common mental disorder case-ness with the C15-R (≥ 12) is used, those living in areas with better built environ-ment are approximately 20% less likely to meet caseness criteria than those living in areas with poorer built environments. Le-venthal & Brooks-Gunn (2003) found that families who moved from a very poor neighbour-hood to a non-poor neighbour-hood showed better mental health than control families who did not move. A similar issue related to mobility is whether or not individuals with poorer mental health may selectively move to more deteriorated areas rather than poorer areas making individuals unhappier (causation vs. selection). Unfor-tunately the design of our study does not allow adequate testing of this theory, and the stability of residence was not recorded.

**Strengths and limitations**

Our study benefited from using a truly con-textual and independent assessment of the built environment rather than measures de-rived from aggregating individual data. The small size of our surveyed areas ensured reasonable homogeneity within sectors. We used multilevel modelling to account for the hierarchical structure of the data. The study was large but its unique setting means its results are not necessarily gener-alisable to other cities throughout the world. Our independent measures at the highest level concentrated on the physical aspects of the environment, mostly because we thought that these could be measured reliably. Of course, the quality of the built environment also reflects something of the psychosocial environment, but we did not include these aspects in this study. This study should be taken as an invitation to explore this field further.

The assessment of the geographical sec-tors was undertaken almost 4 years after we finished the survey of the individuals. Although it is possible that the conditions in those neighbourhoods could have changed in the interim period, we did not find evidence that sectors had experienced major structural changes during the interval according to a survey of local government authorities (Secretaría Regional de Ch ilen de Planificación y Coordinación, 2005). A few sectors with a larger proportion of so-cially disadvantaged individuals were excluded from the analysis. The main reason for sector exclusion was the small number of people in the sector or the lack of data.

Common mental disorders are more preva-lent among socially deprived individuals; thus our estimates may be an underrepresen-tation of the true association. Finally, this is a cross-sectional study and as such we cannot infer the direction of causality. Equally, this kind of design cannot account for factors related to selective migration or population instability.

In conclusion, measuring the impact of the quality of neighbourhoods on mental health and understanding the complex interrelationships between individuals’ characteristics and their local environment are challenges that should be confronted, so that appropriate and effective inter-ventions can be implemented to improve the mental health of the population.

**ACKNOWLEDGEMENTS**

We thank the School of Architecture, Universidad Central de Chile, Santiago, for their support and cooperation in carrying out this study, and to the European Union for funding the Santiago Mental Disorder Survey. This research was supported by a grant from the Wellcome Trust.

**REFERENCES**


Cross-national prevalence and correlates of adult attention-deficit hyperactivity disorder


Background Little is known about the epidemiology of adult attention-deficit hyperactivity disorder (ADHD).

Aims To estimate the prevalence and correlates of DSM-IV adult ADHD in the World Health Organization World Mental Health Survey Initiative.

Method An ADHD screen was administered to respondents aged 18–44 years in ten countries in the Americas, Europe and the Middle East (n=11 422). Masked clinical reappraisal interviews were administered to 154 US respondents to calibrate the screen. Multiple imputation was used to estimate prevalence and correlates based on the assumption of cross-national calibration comparability.

Results Estimates of ADHD prevalence averaged 3.4% (range 1.2–7.3%), with lower prevalence in lower-income countries (1.9%) compared with higher-income countries (4.2%). Adult ADHD often co-occurs with other DSM-IV disorders and is associated with considerable role disability. Few cases are treated for ADHD, but in many cases treatment is given for comorbid disorders.

Conclusions Adult ADHD should be considered more seriously in future epidemiological and clinical studies than is currently the case.

Declaration of interest None. Funding detailed in Acknowledgements.
The retrospective assessment of childhood Adult ADHD. The ACDS has been
used in clinical trials of adult ADHD (Spencer et al., 2001; Michelson et al., 2003).

Four experienced clinical interviewers (all PhD-qualified clinical psychologists) conducted the clinical reappraisal interviews. Each interviewer received 40 h of training from two board-certified psychiatrists, specialists in the treatment of adult ADHD, and successfully completed five practice interviews. All clinical interviews were tape-recorded and reviewed by a supervisor. Weekly calibrator meetings were used to prevent drift. A clinical diagnosis of adult ADHD required six symptoms of either inattention or hyperactivity-impulsivity during the 6 months before the interview (DSM-IV criterion A; American Psychiatric Association, 1994), at least two criterion A symptoms before age 7 years (criterion B), some impairment in at least two areas of living during the previous 6 months (criterion C) and clinically significant impairment in at least one of these areas (criterion D). No attempt was made to operationalise DSM-IV diagnostic hierarchy rules (criterion E).

The DIS questions used to assess ADHD in the main survey were treated as independent variables in the subsample of clinical reappraisal respondents who reported recent symptoms to predict

<table>
<thead>
<tr>
<th>Country</th>
<th>Survey</th>
<th>Sample characteristics</th>
<th>Field dates</th>
<th>Adult ADHD subsample size</th>
<th>Response rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Belgium</td>
<td>ESEMeD</td>
<td>Stratified multistage clustered probability sample of individuals residing in households from the national register of Belgium residents (NR)</td>
<td>2001/2</td>
<td>486</td>
<td>50.6</td>
</tr>
<tr>
<td>Colombia</td>
<td>NSMH</td>
<td>Stratified multistage clustered area probability sample of household residents in all urban areas of the country (approximately 73% of the total national population)</td>
<td>2003</td>
<td>1731</td>
<td>87.7</td>
</tr>
<tr>
<td>France</td>
<td>ESEMeD</td>
<td>Stratified multistage clustered sample of working telephone numbers merged with a reverse directory (for listed numbers); initial recruitment was by telephone, with supplemental in-person recruitment in households with listed numbers (NR)</td>
<td>2001/2</td>
<td>727</td>
<td>45.9</td>
</tr>
<tr>
<td>Germany</td>
<td>ESEMeD</td>
<td>Stratified multistage clustered probability sample of individuals from community resident registries (NR)</td>
<td>2002/3</td>
<td>621</td>
<td>57.8</td>
</tr>
<tr>
<td>Italy</td>
<td>ESEMeD</td>
<td>Stratified multistage clustered probability sample of individuals from municipality resident registries (NR)</td>
<td>2001/2</td>
<td>853</td>
<td>71.3</td>
</tr>
<tr>
<td>Lebanon</td>
<td>LEBANON</td>
<td>Stratified multistage clustered area probability sample of household residents (NR)</td>
<td>2002/3</td>
<td>595</td>
<td>70.0</td>
</tr>
<tr>
<td>Mexico</td>
<td>M‘NCS</td>
<td>Stratified multistage clustered area probability sample of household residents in all urban areas of the country (approximately 75% of the total national population)</td>
<td>2001/2</td>
<td>1736</td>
<td>76.6</td>
</tr>
<tr>
<td>Netherlands</td>
<td>ESEMeD</td>
<td>Stratified multistage clustered probability sample of individuals residing in households that are listed in municipal postal registries (NR)</td>
<td>2002/3</td>
<td>516</td>
<td>56.4</td>
</tr>
<tr>
<td>Spain</td>
<td>ESEMeD</td>
<td>Stratified multistage clustered area probability sample of household residents (NR)</td>
<td>2001/2</td>
<td>960</td>
<td>78.6</td>
</tr>
<tr>
<td>USA</td>
<td>NCS‘R</td>
<td>Stratified multistage clustered area probability sample of household residents (NR)</td>
<td>2002/3</td>
<td>3197</td>
<td>70.9</td>
</tr>
</tbody>
</table>

ADHD, attention-deficit hyperactivity disorder; ESEMeD, European Study of the Epidemiology of Mental Disorders; LEBANON, Lebanese Evaluation of the Burden of Ailments and Needs of the Nation; M‘NCS, Mexico National Comorbidity Survey; NCS‘R, National Comorbidity Survey Replication; NR, nationally representative; NSMH, Colombian National Study of Mental Health.

1. Most World Mental Health (WMH) surveys are based on stratified multistage clustered area probability household samples in which samples of areas equivalent to counties in the UK were selected in the first stage followed by one or more subsequent stages of geographic sampling (e.g., towns within counties, blocks within towns, households within blocks) to arrive at a sample of households. In each of which a listing of household members was created and one or two people were selected from this listing to be interviewed. No substitution was allowed when the originally sampled household resident could not be interviewed. These household samples were selected from census area data in all countries other than France (where telephone directories were used to select households) and The Netherlands (where postal registries were used to select households). Several WMH surveys (Belgium, Germany, Italy) used municipal resident registries to select respondents without listing households. Eight of the ten WMH surveys considered here are based on nationally representative household samples: the two others are based on nationally representative household samples in urban areas (Colombia, Mexico).

2. Attention-deficit hyperactivity disorder was assessed only among respondents in the age range 18–44 years in the Part II sample of each survey.

3. Calculated as the proportion of the number of households in which an interview was completed to the number of households originally sampled, excluding from the denominator households known not to be eligible either because of being vacant at the time of initial contact or because the residents were unable to speak the designated languages of the survey.

respondent visual aids were translated using standardised World Health Organization (WHO) translation and back-translation protocols (these materials are posted at http://www.hcp.med.harvard.edu/wmh). Consistent interviewer training and quality control procedures were used in all surveys. Procedures for informed consent, which was obtained in all countries before beginning interviews, were approved and monitored for compliance by the institutional review boards of the organisations coordinating the surveys in each country.

**Adult ADHD**

The retrospective assessment of childhood ADHD in the WMH surveys was based on the Diagnostic Interview Schedule for DSM-IV (DIS; Robins et al., 1995). Respondents classified retrospectively as having met full ADHD criteria in childhood were then asked a single question about whether they continued to have any current problems with attention or hyperactivity-impulsivity. A clinical reappraisal interview of these respondents was carried out in a probability subsample of 154 respondents in the WMH sample in the USA using the Adult ADHD Clinical Diagnostic Scale, version 1.2 (ACDS; Adler & Cohen, 2004; Adler & Spencer, 2004), a semi-structured interview which includes the ADHD Rating Scale (ADHD–RS; DuPaul et al., 1998) for childhood ADHD and an adaptation of the ADHD–RS to assess current adult ADHD. The ACDS has been used in clinical trials of adult ADHD (Spencer et al., 2001; Michelson et al., 2003).

CROSS-NATIONAL PREVALENCE AND CORRELATES OF ADULT ADHD
questioned in the US clinical reappraisal study applies equally well to the other WMH countries - an assumption that cannot be tested here in light of the fact that no clinical reappraisal study for adult ADHD was conducted in any of the other countries.

Socio-demographic correlates were estimated using multiple imputation logistic regression analysis. Co-occurrence was assessed by obtaining multiply imputed estimates of odds ratios between adult ADHD and other DSM–IV disorders in logistic regression equations that controlled for age in 5-year age groups. Functional disabilities were also estimated using multiple imputation logistic regression. Twelve-month treatment was estimated using multiple imputation cross-tabulations. In each phase of analysis we generated estimates both separately for each of the ten samples and also in a combined cross-sample analysis that included nine dummy control variables to indicate country. Interactions were then estimated between the country dummies and the substantive predictors to evaluate the significance of between-country differences. Such differences, although few in number, are noted in the following presentation of substantive results.

Part I cases were weighted to adjust for differential probabilities of selection within and between households and to match sample distributions to population distributions on socio-demographic and geographic data. The part II sample was additionally weighted for the undersampling of part I respondents without core disorders. Because the sample design used this weighting as well as geographic clustering, all parameters were estimated using the Taylor series linearisation method (Wolter, 1985), a design-based method implemented in the SUDAAN software system (Research Triangle Institute, North Carolina, USA). All significance tests used two-sided Wald $\chi^2$ tests based on design-corrected multiple imputation variance-covariance matrices.

**RESULTS**

**Prevalence**

The estimated prevalence of DSM–IV adult ADHD in the total sample based on multiple imputation, using a combination of directly interviewed cases from the clinical reappraisal sample in the USA and multiply imputed cases in the remainder
of the samples, was 3.4%, s.e. = 0.4 (Table 2). Prevalence estimates were significantly higher than this average in France (7.3%, s.e. = 1.8) and significantly lower in Colombia (1.9%, s.e. = 0.5), Lebanon (1.8%, s.e. = 0.7), Mexico (1.9%, s.e. = 0.4) and Spain (1.2%, s.e. = 0.6).

Socio-demographic correlates

Multiple imputation prevalence estimates of clinician-assessed adult ADHD were significantly greater in the total cross-national sample among men and among people educated to less than university level (Table 3), but these effects were modest in magnitude (1.5 < OR < 3.0). No significant between-country difference was found in the magnitude of the effects of gender and education, although it is noteworthy that there was little power to detect such effects (further details available from the authors).

Co-occurrence with other DSM-IV disorders

Adult ADHD was significantly associated with a wide range of other 12-month DSM-IV disorders (Table 4). The strength of these associations in terms of odds ratios was remarkably consistent across classes of disorder, with OR = 3.9 (95% CI 3.0-5.1) for mood disorders, OR = 4.0 (95% CI 3.0-5.2) for anxiety disorders and OR = 4.2 (95% CI 2.8-5.8) for substance use disorders. A dose-response relationship exists between ADHD and number of other disorders, with the highest odds ratio (OR = 7.2, 95% CI 5.1-10.2) associated with having three or more other disorders. Within-country patterns were similar to those in the combined sample, with a predominantly positive sign pattern (68 of the 70 odds ratios in the ten separate countries were greater than 1.0) and 56% of the within-country odds ratios significant at the P < 0.05 level. However, this pattern was notably weaker in France (further details available from the authors).
Table 4  Bivariate lifetime co-occurrence of multiply imputed adult attention-deficit hyperactivity disorder and other DSM IV disorders (n = 11,422)

<table>
<thead>
<tr>
<th>Classes of co-occurring disorders</th>
<th>ADHD/Co&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Co/ADHD&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mood</td>
<td>11.1 (1.2)</td>
<td>24.8 (2.6)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>9.9 (1.0)</td>
<td>38.1 (3.1)</td>
</tr>
<tr>
<td>Substance use</td>
<td>12.5 (2.3)</td>
<td>11.1 (2.0)</td>
</tr>
</tbody>
</table>

Number of co-occurring disorders

| Exactly one                        | 5.4 (0.7)           | 20.4 (2.1)          |
| Exactly two                        | 10.3 (1.5)          | 129 (1.6)           |
| Three or more                      | 20.3 (2.4)          | 162 (2.4)           |
| Any                                | 8.5 (0.8)           | 49.5 (3.6)          |

ADHD, attention-deficit hyperactivity disorder; Co, comorbid disorder.
1. Conditional prevalence estimates of adult ADHD in the subsamples of respondents with the comorbid disorders.
2. Conditional prevalence estimates of the comorbid disorders in the subsample of respondents with adult ADHD.
3. All odds ratios significant at P < 0.05, two-sided test.

Table 5  Temporal priorities in first onset of co-occurring adult attention-deficit hyperactivity disorder and other DSM IV disorders

<table>
<thead>
<tr>
<th>Co-occurring disorder</th>
<th>ADHD first % (s.e.)</th>
<th>Other disorder first % (s.e.)</th>
<th>Both in same year % (s.e.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mood disorder</td>
<td>85.6 (2.5)</td>
<td>95.2 (2.4)</td>
<td>49.1 (1.3)</td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td>49.6 (3.9)</td>
<td>41.2 (4.0)</td>
<td>91.2 (2.0)</td>
</tr>
<tr>
<td>Specific phobia</td>
<td>34.3 (5.3)</td>
<td>54.8 (5.1)</td>
<td>110.0 (2.8)</td>
</tr>
<tr>
<td>Any other anxiety disorder</td>
<td>68.5 (4.1)</td>
<td>19.7 (3.2)</td>
<td>118.2 (2.2)</td>
</tr>
<tr>
<td>Substance use disorder</td>
<td>99.0 (0.7)</td>
<td>0.5 (0.5)</td>
<td>0.4 (0.4)</td>
</tr>
</tbody>
</table>

ADHD, attention-deficit hyperactivity disorder.
1. Number of respondents with co-occurrence of adult ADHD and the type of disorder specified.

Table 4  Disability in 30-day functioning associated with adult attention-deficit hyperactivity disorder (n = 11,422)

<table>
<thead>
<tr>
<th>Disability</th>
<th>% (s.e.)</th>
<th>With controls for socio-demographic data</th>
<th>With controls for socio-demographic data and other DSM IV disorders&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-care</td>
<td>4.2 (1.0)</td>
<td>1.5 (0.8–2.8)</td>
<td>0.9 (0.5–1.6)</td>
</tr>
<tr>
<td>Mobility</td>
<td>16.9 (1.9)</td>
<td>2.2* (1.6–2.9)</td>
<td>1.5* (1.1–2.0)</td>
</tr>
<tr>
<td>Cognitive</td>
<td>20.5 (2.6)</td>
<td>3.9* (2.8–5.4)</td>
<td>2.2* (1.5–3.3)</td>
</tr>
<tr>
<td>Days out of role</td>
<td>31.4 (3.0)</td>
<td>2.6* (2.0–3.5)</td>
<td>1.8* (1.3–2.5)</td>
</tr>
<tr>
<td>Social interaction</td>
<td>10.7 (1.7)</td>
<td>3.1* (2.1–4.5)</td>
<td>1.5* (1.0–2.2)</td>
</tr>
</tbody>
</table>

1. Based on logistic regression equations controlling for country, age, education, employment, marital status and income.
2. Based on logistic regression equations controlling for country, age, education, employment, marital status, income, any mood disorder, any anxiety disorder and any substance use disorder.

**<i>P < 0.05</i>, two-sided test.

**<i>PP</i> = 0.05, two-sided test.

Disability

Adult ADHD was associated with significantly elevated odds ratios of disability in two of the three WHO–DAS dimensions of basic functioning – mobility (OR = 2.2, 95% CI 1.6–2.9) and cognition (OR = 3.9, 95% CI 2.6–5.4) - but not in the third dimension of self-care (OR = 1.5, 95% CI 0.8–2.8) (Table 6). Adult ADHD was also associated with elevated risk of high number of days out of role (OR = 2.6, 95% CI 2.0–3.5) and with disability in social functioning (OR = 3.1, 95% CI 2.1–4.5). These associations become somewhat weaker but remain statistically significant when controls are introduced for co-occurring anxiety, mood and substance use disorders. Within-country patterns are again similar to those in the combined sample, with 82% of within-country odds ratios greater than 1.0 and 46% significant at P < 0.05 level (further details available from the authors). The Netherlands is the only country where reported disability was consistently and significantly lower than the results in the combined sample. Only a handful of other within-country odds ratios differed significantly from the cross-national averages.

Twelve-month treatment

Patterns of treatment for emotional or substance use problems in the 12 months before interview among respondents with adult ADHD differed much more markedly across surveys than did any of the other statistics examined in this report (Table 7). The highest proportion of cases receiving treatment was in the USA, where nearly half (49.7%) of respondents reported some type of care, followed by roughly half as many (19.9–23.8%) receiving treatment in three of the European countries (Belgium, The Netherlands and Spain), roughly half this proportion (9.4–12.4%) in four other countries (Colombia, France, Germany and Mexico) and only 1.1% in Lebanon. The majority of people receiving treatment were seen in the specialty mental health sector in all countries other than France and Italy, where the majority were seen in the general medical sector. It is important to recognise that these patients were generally seen not for problems with attention, concentration, impulsivity or phobia first v. 34.3% ADHD first). These patterns are very robust across countries (further details available from the authors).
ADHD, attention-deficit hyperactivity disorder; CAM, complementary and alternative medicine.

**DISCUSSION**

Our findings have to be interpreted in the context of several important limitations. First, the diagnoses of adult ADHD in both the DIS and clinical reappraisal interviews were based on adult self-reports. Childhood ADHD is diagnosed on the basis of parent and teacher reports because children with ADHD often are unaware of their symptoms (Jensen et al., 1999). Use of informants, such as spouses or work supervisors, to assess adult ADHD is much more difficult (although ideal in clinical settings), making it necessary to base assessment largely on self-report (Wender et al., 2001). Although the one study that compared adult self-reports with informant reports of ADHD symptoms in a non-clinical sample found fairly strong associations between the two reports (Murphy & Schachar, 2000), our use of self-report without confirmation by informant reports still has to be seen as a limitation.

More importantly, our use of imputation to estimate adult ADHD introduced several other important limitations that need to be recognised in interpreting our results. For one, the model relied on retrospective assessments of childhood symptoms in conjunction with only a single question about recent adult persistence. Even though these responses were strongly related to independent clinical assessments of adult ADHD in the US sample, the coarse classification created by relying on only a single question about recency limited the texture with which we could study correlates of adult ADHD. This coarseness reduces the precision of estimates and, with it, attenuates measures of association. In addition, the imputation model was based on a clinical calibration conducted only in the USA. We have no way of confirming the analytical assumption that the positive and negative predictive values estimated to calibrate the imputations are the same in the other countries studied – an assumption that is fundamental to the imputation method. This is especially problematic given that, as noted in the introduction, little research on adult ADHD has been conducted outside the USA, making it unclear if the same markers apply in other countries. Given the centrality of this issue, it is important that the CIDI assessment of ADHD is expanded for use in future surveys in countries other than the USA. Another limitation of the imputation model – which would be relevant even if the model were equally accurate in all countries – is that it understates the strength of associations of adult ADHD with covariates that, owing to limitations of sample size, were not included as predictors in the model. This means that the evidence regarding socio-demographic correlates of adult ADHD reported here is likely to be conservative.

Finally, a question can be raised about the validity of the DSM-IV ADHD criteria when applied to adults, considering they were developed with children in mind. Clinical studies make it clear that symptoms of ADHD are more heterogeneous and subtle in adults than in children (De Quiros & Kinsbourne, 2001), leading some clinical researchers to suggest that assessment of adult ADHD might require an increase in the variety of symptoms assessed (Barkley, 1995), a reduction in the severity threshold (Ratey et al., 1992) or a reduction in the DSM-IV ‘six of nine’ symptom requirement (Kooi et al., 2005). To the extent that such considerations in the criteria would lead to a more valid assessment than in the current study, our prevalence estimate is conservative.

Within the context of these limitations, the results reported suggest that adult ADHD as currently defined in the DSM-IV is a commonly occurring and often seriously impairing disorder. The 3.4% estimated prevalence is likely to be conservative for the reasons described above. Although we would expect to find some variation in prevalence from one country to another, the amount of cross-national variation in the estimated prevalence is small compared with estimates for other disorders (Denytenaere et al., 2004). This low variation might be due to methodological factors such as a general lack of awareness about ADHD that makes it difficult for respondents to discriminate between questions, or that leads to normative cultural interpretations of certain symptoms (e.g. a high tolerance of hyperactivity in boys). Another possibility, though, is that adult ADHD is less strongly related than other disorders to environmental determinants that can vary across countries.
The findings that adult ADHD is significantly more prevalent among men than women and among people with low rather than high educational levels are consistent with much previous research (Schall & Schwab-Stone, 2000) and, as noted above in the discussion of limitations, are likely to be underestimates of the strength of these associations owing to the attenuation introduced by the coarseness of the imputations. The failure to find an elevated prevalence of ADHD among unemployed people, however, is inconsistent with these same studies. Nonetheless, we do find that WMH respondents estimated to have ADHD report significantly more disability in role functioning, as indicated by more days out of role and more disability in social role functioning, than comparable respondents without ADHD. These results regarding role disability are consistent with much previous research on disability in adult ADHD (Able et al., 2007). It is noteworthy that the WHO–DAS dimension associated with the highest impairment in the current study is the cognitive disability dimension. This finding is as one would expect, given the nature of the disorder. However, the WHO–DAS might underestimate ADHD disability because some WHO–DAS dimensions tap areas where ADHD is not highly disabling (e.g. people with ADHD are often very mobile and overwork) and because the WHO–DAS does not assess many dimensions where people with ADHD are thought to function less adequately (e.g. poor sleep and nutrition, high rates of accidents, high levels of smoking). Moreover, people with ADHD often have poor insight into their functioning, possibly leading to underestimation of WHO–DAS scores. It might also be that the social and interpersonal disabilities associated with adult ADHD require more detailed probing to detect than provided in the WHO–DAS. Based on these considerations, along with the more general problem noted above that imputation leads to attenuation of associations, the disabilities due to ADHD are likely to be underestimated. This makes it all the more striking that adult ADHD is consistently associated across countries with substantial elevations in disability that cannot be accounted for by co-occurring disorders.

The estimate that adult ADHD often co-occurs with other DSM–IV disorders is consistent with clinical evidence (Biederman, 2004). Methodological analysis shows that the evidence of co-occurrence holds up when careful diagnoses are made aimed at adjusting for overlap of symptoms, imprecision of diagnostic criteria, or other methodological confounds (Angold et al., 1999). The results regarding co-occurrence in our report, however, are likely to be much less precise – both because diagnoses of co-occurring disorders are based on a fully structured interview that, due to its limited ability to make differential diagnoses, will cause overestimation of co-occurrence, and because the diagnoses of adult ADHD are based on coarse imputations that, due to their individual-level imprecision, will lead to attenuation of correlations with other variables and consequent underestimation of systematic co-occurrence (i.e. underestimation of odds ratios).

As one might expect from the early onset of ADHD, comparison of reports of age at onset showed that the estimated co-occurrence in the WMH surveys is due to temporally primary ADHD being related to the subsequent onset of other disorders. The main exception here is co-occurring specific phobia, which is typically temporally primary to ADHD. This last observation raises the question whether early successful treatment of childhood ADHD would influence secondary adult disorders, an issue that is beyond the scope of the current report to investigate. A related question is whether adult treatment of ADHD would have any effect on severity or persistence of co-occurring temporally secondary disorders. Long-term research is needed to answer these questions. The results reported here highlight the importance of such long-term research by documenting that adult ADHD is a relatively common disorder in a number of countries, often co-occurs with largely temporally secondary conditions, and that it is associated with substantial impairment in adult role functioning.

ACKNOWLEDGEMENTS

The surveys discussed above were carried out in conjunction with the World Health Organization World Mental Health (WMH) Survey Initiative. We thank the WMH staff for assistance with instrumentation, fieldwork and data analysis. These activities were supported by the US National Institute of Mental Health (R01-MH070884), the John D. and Catherine T. MacArthur Foundation, the Pfizer Foundation, the US Public Health Service (R3-MH066849, RO1-MH068964 and R01-DA016558), the Fogarty International Center (FICRA RO1-TW006481), the Pan American Health Organization, Eli Lilly and Company, Ortho–McNeil Pharmaceutical, Inc, GlaxoSmithKline and Bristol–Myers Squibb. The current report was additionally supported by Eli Lilly and Company. A complete list of WMH publications can be found at http://www.hcp.med.harvard.edu/wmh/. The Chinese World Mental Health Survey Initiative is supported by the Pfizer Foundation. The Colombian National Study of Mental Health is supported by the Ministry of Social Protection, with supplemental support from the Saldarriaga Concha Foundation. The European Study of the Epidemiology of Mental Disorders (ESEMeD) project is funded by the European Commission (contracts QLG–1999–01042; SANCDO 2004123), the Piedmont Region (Italy), Fondo de Investigación Sanitaria, Instituto de Salud Carlos III, Spain (FIS 00/0028), Ministerio de Ciencia y Tecnología, Spain (SAF 2000–158 CE), Departament de Salut, Generalitat de Catalunya, Spain, RETICS RD06/0011 REM–TAP and other–local agencies and by an unrestricted educational grant from GlaxoSmithKline. The Lebanese Evaluation of the Burden of Ailments and Needs of the Nation (LEBANON) survey was partially supported by anonymous private donations to the Institute for Development, Research, Advocacy and Applied Care (IDRAAC), Lebanon (http://www.idraac.org), and by the Lebanese Ministry of Public Health, the World Health Organization (WHO: Lebanon), Act 4 Lebanon, The National Arab–American Medical Association and unrestricted grants from Janssen Cilag, Eli Lilly and Company, GlaxoSmithKline, Roche and Novartis. The Mexico National Comorbidity Survey is supported by the National Institute of Psychiatry Ramon de la Fuente (INPRMFIDES 4280) and by the National Council on Science and Technology (CONAGT–G30544–H), with supplemental support from the Pan American Health Organization. The US National Comorbidity Survey Replication is supported by the National Institute of Mental Health (NIMH U51-MH160220) with supplemental support from the National Institute of Drug Abuse, the Substance Abuse and Mental Health Services Administration, the Robert Wood Johnson Foundation (RWJF grant 044780) and the John W. Alden Trust. Professor Kessler, as Director of the WMH Data Analysis Coordination Centre, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Professor Kessler was responsible for instrument design and supervision of data analysis. The co-authors were responsible for preparing the first draft of the manuscript, data collection and making critical revisions to the manuscript.

REFERENCES


Serotonergic function in children with attention-deficit hyperactivity disorder

Relationship to later antisocial personality disorder

JANINE D. FLORY, JEFFREY H. NEWCORN, CARLIN MILLER, SETH HARTY and JEFFREY M. HALPERIN

Background Impulsive aggression in adulthood is associated with disturbances in serotonergic function. In contrast, research examining this association in childhood has yielded inconsistent results.

Aims The current study examined the prospective relationship between serotonergic function measured in childhood and the later emergence of antisocial personality disorder.

Method Hormonal response to fenfluramine, an index of serotonergic function, was assessed in 58 children with attention-deficit hyperactivity disorder between 1990 and 1997 when they were aged 7–11 years. Approximately 9 years later these individuals were evaluated for antisocial personality disorder.

Results Lower serotonergic responsivity assessed in childhood predicted the development of antisocial personality disorder (t (56) = 2.25, P = 0.028).

Conclusions These results provide a critical link between the child and adult literature on the covariation of impulsive aggression and serotonergic function and suggest a potential explanation for inconsistencies in the childhood literature.

Declaration of interest None. Funding detailed in Acknowledgements.

The association between deficiencies in central nervous system (CNS) serotonergic functioning and adult impulse aggression, particularly among males, is well established, whether indexed as aggressive behaviour leading to incarceration, lifetime history of aggressive acts as measured by semi-structured interview, or self-ratings of dispositional aggression (see M anuck et al, 2005a for a review). Results from a parallel body of research with non-human primates support this association (e.g. M ehman et al, 1994; H igley et al, 1996). Sociopathy is also associated with dysregulated serotonergic function (e.g. O’K eane et al, 1992), although people with antisocial personality disorder do not always exhibit aggressive behaviour (i.e. antisocial personality disorder and impulsive aggression are not completely overlapping constructs).

In contrast to this large and consistent body of work, the evidence linking dysregulated serotonergic function and aggression among children is equivocal (K ruesi et al, 1990; S toff et al, 1992; C astellanos et al, 1994; H alperin et al, 1994, 1997; P ine et al, 1997). One reason might be the exclusive focus on cross-sectional studies of childhood and early adolescence when aggression is temporarily unstable (M offitt, 1993); longitudinal research that examines the neurobiological correlates of temporally persistent aggression might yield more consistent findings. In support of this view, one prospective study links lower 5-hydroxy-indole acetic acid (5-HIAA) levels in the cerebrospinal fluid (CSF) of boys with disruptive behaviour disorders to aggression scores/arrests measured 2 years later (K ruesi et al, 1992). The current study capitalised on the unique opportunity to observe the prospective relationship between serotonergic function assessed in childhood and the emergence of antisocial personality disorder in late adolescence/young adulthood. We predicted that a lower prolactin response to fenfluramine in childhood would be associated with the development of antisocial personality disorder assessed 9 years later.

METHOD

Baseline evaluation

Between July 1990 and May 1997, 7- to 11-year-old children were referred by clinicians to a research programme examining the relationship between serotonergic function and aggression in children with disruptive behaviour disorders. Children were screened using the IOWA Conners Teacher Rating Scale (Lonley & M illich, 1982). Subsequently, parents completed the Child Behavior Checklist (CBCL; Achenbach, 1991) and were interviewed with the Diagnostic Interview Schedule for Children (Shaffer et al, 1996).

The current report is based on the total number of participants with attention-deficit hyperactivity disorder (ADHD) who had undergone the childhood evaluation including a fenfluramine challenge and who were later interviewed for identification of Axis II psychopathology between August 2001 and February 2005. The sample included 52 males and 6 females and represents 53% of the total number of children who were administered the fenfluramine challenge and were eligible for follow-up in February 2005 (n = 110). Participation in the follow-up evaluation was not associated with gender, socio-economic status, age at initial evaluation, IQ, or parent and teacher ratings of psychopathology (P < 0.18). At baseline, all 58 children (mean age 9.24, s.d. = 1.19 years) met diagnostic criteria for ADHD and 47 were diagnosed with oppositional-defiant disorder; 21 met criteria for conduct disorder.

Fenfluramine challenge

Following the clinical evaluation and prior to the day of the challenge protocol, participants followed a low monoamine diet for 3 days and reported to the laboratory at 08.00 h after fasting overnight. An indwelling catheter was inserted into a forearm vein. Following an adaptation period, baseline blood samples were drawn at 09.45 h and 09.55 h. At 10.00 h, a 1 mg/kg dose of d,l-fenfluramine hydrochloride was administered orally. Blood samples were drawn 60, 120, 180, 240 and 300 min later for determination of plasma prolactin, fenfluramine and norfenfluramine concentrations. All samples were placed on ice prior to centrifugation (within 2 h). After
separation, samples were frozen at \(-80^\circ \text{C}\) until analysis. The lower limit of detection for the prolactin assay is \(<1.0 \text{ ng/ml}\). Intra- and interassay variability are less than 6.7% and 8.4% respectively. Blood samples for determination of plasma fenfluramine and norfenfluramine were drawn hourly. After separation, these samples were frozen at \(-20^\circ \text{C}\) prior to assay by gas chromatography with electrical detection. The lower limit of sensitivity for these assays is 2 ng/ml for fenfluramine and 3 ng/ml for norfenfluramine. Intra- and interassay variability for the two assays are less than 7%. Participants remained awake and fasting during the entire procedure, reclining in a bed and watching videotapes. All fenfluramine studies were completed prior to September 1997.

**Follow-up evaluation**

Participants who were evaluated for ADHD in childhood were contacted to participate in a study of the longitudinal course of ADHD. The mean (s.d.) age at follow-up was 18.44 (1.23) years; follow-up took place on average 9.4 years later (s.d.=1.9 years). Seventeen participants identified themselves as African American, 18 as non-Hispanic White, 18 as Black Hispanic and 5 as mixed race. The participants were generally of lower- to lower-middle socioeconomic status (SES) (mean SES score 44.59, s.d.=15.62; Nakao & Treas, 1994). At follow-up, all participants and a parent or adult relative (informant) were administered the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II; First et al., 1997).

Although diagnoses of personality disorder have not traditionally been assigned to people under 18, there is growing acceptance that personality disorder diagnoses can be made reliably in adolescence (e.g. Bernstein et al., 1993; Johnson et al., 1995) and DSM-IV-TR (American Psychiatric Association, 2000) allows for this if the symptoms are persistent and present for at least 1 year. In the case of antisocial personality disorder, a prior diagnosis of conduct disorder is also required and, by definition, it is the only personality disorder diagnosis that cannot be made in individuals under 18. In the current study, because the SCID-II interview was used for research purposes it was administered to all participants regardless of age; if they were found to meet criteria for conduct disorder based on the interview, the criteria for antisocial personality disorder were also evaluated. Interviewers were masked to baseline data, including whether children had been given a diagnosis of conduct disorder during the baseline study, and the interviews were conducted independently (i.e. the same interviewer did not interview the participant and informant). Proband and informant interviews were not conducted for all participants; one proband died prior to evaluation and three probands did not identify an informant. If either the participant or the informant endorsed criteria for antisocial personality disorder, the participant was considered to carry the diagnosis. Concordance (Pearson’s r) between ratings was 0.62.

Signed informed consent was obtained from participants over 18 years and from parents for participants under 18, in whom consent was obtained. The protocol was approved by the institutional review boards at Queens College, City University of New York and M unct Sinai School of Medicine.

**Data reduction and statistical analyses**

Baseline prolactin level was calculated as the mean of two blood samples collected prior to administration of fenfluramine and area under the curve (AUC PRL[fen]) was calculated using trapezoidal integration (Pruessner et al., 2003). Both baseline prolactin and AUC PRL[fen] were log-transformed prior to analyses to normalise the distributions. Owing to the moderate association between baseline prolactin and AUC (r=0.57), log AUC PRL[fen] was regressed on log baseline prolactin to obtain a baseline-independent measure of serotonergic responsivity. This served as the major outcome measure of serotonergic responsivity. Fenfluramine and norfenfluramine levels during the challenge were averaged across time but were unavailable for three participants. These averages were found to be correlated with log AUC PRL[fen] (r=0.45 and 0.39 respectively); log AUC PRL[fen] was thus adjusted for covariation with log baseline prolactin concentration and drug levels using linear regression. The hypothesis that lower serotonergic responsivity would predict the development of antisocial personality disorder was evaluated using two-tailed Student’s t-tests.

**RESULTS**

Seventeen participants (29%) met criteria for antisocial personality disorder according to self-report or informant report and results indicated that lower serotonergic responsivity in childhood predicted the development of antisocial personality disorder (r (56)=2.25, P=0.028). Unadjusted prolactin levels throughout the 5h challenge are shown in Fig. 1. The difference between groups was greater after also adjusting for drug concentrations (r (53)=2.71, P<0.001). Gender is associated with responsivity to fenfluramine among adults (McBride et al., 1990) and children (Koda et al., 1996), with females showing greater hormonal responsivity. Although there were no gender differences in serotonergic responsivity in the current sample (r (56)=0.26, P=0.79), no females in the current sample met criteria for antisocial personality disorder. Thus, it might be argued that gender accounts for the group differences reported here. We therefore conducted post hoc analyses excluding girls, but the results were unchanged after adjusting for log baseline prolactin (r (50)=2.27, P=0.028) and after adjusting for log baseline prolactin and drug concentrations (r (48)=2.93, P=0.005).

The two diagnostic groups were then compared for other characteristics measured in childhood and were not found to differ in age (P=0.76), full-scale IQ (P=0.42), parent ratings on the CBCL of attentional difficulties (P=0.75) or aggression (P=0.39). There was a tendency for parents to endorse higher levels of delinquent behaviour in childhood among those who later developed antisocial personality disorder (r (49)=−1.88, P=0.066) and teachers rated these children as more aggressive (r (53)=−2.23, P=0.03) on the IOWA Conners Aggression Scale. The two groups did not differ in the number of years between the fenfluramine challenge and evaluation of antisocial personality disorder (P=0.10) or in SES score measured at the second evaluation (P=0.25). Interviews to ascertain family history were conducted.

![Fig 1](image.png)  
*Fig 1: Plasma prolactin concentrations following administration of a 1mg/kg dose of dl-fenfluramine hydrochloride.* - - - - , Antisocial personality disorder; -, no antisocial personality disorder.
at the initial evaluation for 47 participants (81%). Results indicated that participants who met criteria for antisocial personality disorder had a higher percentage of relatives with the disorder, ($r = 0.45$, $P = 0.001$).

Criteria for antisocial personality disorder were evaluated if participants met criteria for conduct disorder on the SCID-II, even if they were under age 18. However, post hoc analyses included only participants who were 18 years or over at the time of the interview assessment ($n = 33$). Despite loss in statistical power, results indicated that lower serotonergic responsivity was associated with the development of antisocial personality disorder, ($r = 0.31$, $P = 0.03$).

Finally, an exploratory forward selection stepwise logistic regression analysis using maximum likelihood estimates was conducted to examine the relative influence of childhood conduct disorder, childhood aggression (IOWA Connors Aggression Scale) and childhood serotonergic function (baseline-adjusted log AUC PRL[ten]) in predicting a diagnosis of antisocial personality disorder. Results indicated that a model that included serotonergic functioning and childhood aggression was significant ($\chi^2(2) = 13.22$, $P = 0.001$, $-2LL = 53.79$). Serotonergic functioning entered the model first ($B = -3.05$, 95% CI $-5.43$ to $-0.68$, $z = 6.4$, $P = 0.012$), followed by childhood aggression ($B = 0.21$, 95% CI $0.04$ to $0.38$, $z = 6.2$, $P = 0.013$). Odds ratios are not reported because serotonergic functioning is represented by a log transformation of AUC, which is not a meaningful unit of measurement.

**DISCUSSION**

**Main findings**

These results are the first to demonstrate that dysregulated serotonergic function measured during childhood (i.e. lower prolactin responsivity to fenfluramine) predicts the emergence of antisocial personality disorder in early adulthood. Although consistent with a considerable body of adult human and non-human primate research showing that dysregulated serotonergic function is associated with impulsive aggression, these are remarkable findings given that serotonergic responsivity was assessed 9 years prior to the assessment of antisocial personality disorder and was not at the time correlated with the severity of childhood aggression (Schulz et al., 2001).

Moreover, these results provide a critical link between the child and adult literature on the covariation of impulsive aggression and serotonergic function because this is the only study that spans the developmental period between childhood and early adulthood.

**Other studies**

A potential explanation for inconsistencies in the childhood literature is that aggressive and delinquent behaviour across childhood and adolescence is temporarily unstable and less tightly linked to individual variation in serotonergic function than in adulthood. Developmental theories of antisocial personality disorder propose alternative trajectories from childhood to adulthood, including continuous or life-course-persistent and transitory or adolescence-limited antisocial behaviour (DiLalla & Gottesman, 1989; Moffitt, 1993). These theories propose that delinquent behaviour in children that persists into adulthood is a more heritable or biologically mediated form of behaviour that likely includes aggression and violence. This view is supported by a study showing that age at onset moderates the association between genetic and environmental influences on antisocial behaviour (Sulske et al., 1997) and a meta-analysis of twin studies of aggression (Miles & Carey, 1997), which concluded that heritability estimates of aggression increase from childhood to adulthood, whereas the relative magnitude of environmental influences decreases. In the current study, we observed that boys who went on to develop antisocial personality disorder had more relatives with such a diagnosis, which is consistent with our previous report that aggressive children with higher familial aggregation of aggressive and antisocial behaviours show a lower prolactin response to fenfluramine (Halperin et al., 2003).

In contrast to life-course-persistent antisocial behaviour, many adolescents engage in more ‘normative’ forms of delinquent behaviour, including substance use. Because the base rate of this behaviour is so high, it is correlated with neither childhood nor adult behaviour. Of the 21 children diagnosed with conduct disorder in childhood in the current sample, only about half ($n = 11$) went on to develop antisocial personality disorder, corresponding to the stability data reported by Lahey et al. (2005) in a longitudinal study of conduct disorder first assessed in childhood. In addition, seven individuals who met criteria for antisocial personality disorder did not carry a diagnosis of conduct disorder at the time of the childhood evaluation, although six of the seven met criteria for oppositional defiant disorder, another precursor for antisocial personality disorder. In these, features of conduct disorder emerged after the childhood evaluation, and indeed, were apparent at the time of assessment of antisocial personality disorder. It should also be noted that children who did not meet full criteria for conduct disorder during childhood would not necessarily be considered ‘non-aggressive’. Because all children met criteria for a disruptive behaviour disorder upon study entry, they might have exhibited features of conduct disorder that did not reach the threshold level for diagnosis.

The results from the current study are not informative regarding the origins of the covariation between dysregulated serotonergic activity and antisocial personality disorder. Pedigree studies suggest that indices of serotonin (e.g. whole blood, CSF 5-HIAA levels) are heritable (Higley et al., 1993; Abney et al., 2001), as are measures of aggression and antisocial behaviour (Miles & Carey, 1997; Slutske et al., 1997). Environmental conditions suggest that social adversity (e.g. peer rearing in non-human primates (Shannon et al., 2005) and poverty and unemployment (Manuck et al., 2005)) are associated with dysregulated serotonergic function and these social factors may interact with functional variants of serotonin-regulating genes to confer greater risk for antisocial personality disorder (Caspi et al., 2002). Because the current sample included only children with disruptive behaviour disorders, in whom we would expect to see a full complement of genetic and environmental risk predictors for antisocial personality disorder, we cannot disentangle the relative impact of these factors on serotonergic function prior to age 7. Moreover, we do not posit that dysregulated serotonergic function is the sole feature that leads to antisocial personality disorder, but consider it one aspect of a complex interplay between biological and psychosocial variables.

**Limitations of the study**

Limitations of the current study should be acknowledged. The sample size was small and prospective replication of the association is warranted. However, the extended
length of follow-up and the fact that the association between serotonergic functioning and antisocial personality disorder could not be explained by common risk factors (e.g., gender, IQ, SES) suggests that the finding is robust. The sample was predominantly male and the findings may not generalise to girls, but we note that the results were unchanged upon exclusion of females from the analyses. In addition, although it is of great interest to identify neurobiological features that are associated with specific aspects of antisocial personality disorder (e.g., Yang et al., 2005), the number of people meeting the criteria for such a diagnosis was considered too small to conduct these analyses.

Limitations of the method used for measuring central serotonergic function should be acknowledged. Throughout the 1980s and 1990s hormonal responsivity to fenfluramine was a well-established and frequently-used measure for assessing serotonergic function in the hypothalamic-pituitary axis and the procedure is considered to reflect ‘net’ serotonergic transmission from the raphe nuclei to the hypothalamus, including both presynaptic and postsynaptic functioning (Coccaro et al., 1989; Yatham & Steiner, 1993), although data suggest that the findings may generalise to the prefrontal cortex (Soloff et al., 2000). This method of assessment has been largely supplanted by neuroimaging which affords greater regional specificity of central serotonergic pathways and receptor activity. In addition, it should be noted that fenfluramine was withdrawn from the US market in 1997 when safety concerns curtailed its continued use in neurobiological research, which prevented us from repeating biological assessment in our young adult sample.

Antisocial personality disorder is a heterogeneous entity that includes overlapping constructs: criminal behaviour, impulsive aggression and a lack of remorse for transgressions. The literature suggests that regulated serotonergic function will be associated prospectively with impulsive aggression rather than other aspects of antisocial personality disorder, but this hypothesis awaits verification.

The age range of the current sample spans the time of the developmental transition from adolescence into young adulthood and it will be critical to follow these individuals over the next 5–10 years to determine whether their current levels of behaviour persist. The longitudinal nature of this study will enable us to examine specific childhood neurobiological and psychosocial predictors of the persistence of aggression and related forms of antisocial behaviour.

ACKNOWLEDGEMENTS

This work was supported in part by grants ROIPIH64414 (to J.M.H) and K01DE09779 (to J.D.J.) from the National Institute of Mental Health.

REFERENCES


Pruehsner, J. C., Kirschbaum, C., Meinschmidt, G., et al. (2003) Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. Psychoneuroendocrinology, 28, 916–931.


Personality disorder and impaired functioning from adolescence to adulthood

ANDREW E. SKODOL, JEFFREY G. JOHNSEN, PATRICIA COHEN, JOEL R. SNEED and THOMAS N. CRAWFORD

Background  Little is currently known about functioning and impairment during adulthood associated with the course of personality disorders.

Aims  To investigate the association of personality disorder stability from adolescence through middle adulthood with measures of global functioning and impairment, using prospective epidemiological data.

Method  A community-based sample of 658 individuals was interviewed at mean ages 14, 16, 22 and 33 years.

Results  Individuals with persistent personality disorder had markedly poorer functioning and greater impairment at mean age 33 years than did those who had never been identified as having such disorder or who had a personality disorder that was in remission, after co-occurring Axis I disorders at age 33 years were taken into account. Remitted disorder was associated with mild long-term impairment. Adult-onset personality disorders, however, were also associated with significant impairment.

Conclusions  Persistent and adult-onset personality disorders are associated with functional impairment among adults in the community. These effects are independent of co-occurring Axis I disorders.

Declaration of interest  None. Funding detailed in Acknowledgements.

An enduring pattern of inner experience and behaviour associated with occupational and interpersonal dysfunction is central to the concept of personality disorder (World Health Organization, 1992; American Psychiatric Association, 2000). In rigorous follow-along studies of personality disorders, however, rates of symptomatic improvement in patient populations (Shea et al., 2002; Zanarini et al., 2003; Grilo et al., 2004) and in non-patient (Lenzenweger, 1999) or community (Johnson et al., 2000a) populations are inconsistent with the stability hypothesis. Because personality disorders have their origins in childhood or adolescence, deficits in the development of affect regulation, conscience, impulse control or identity consolidation can be expected to have an adverse impact on a person’s adaptation to the occupational and interpersonal demands of young adult life, which may persist even beyond symptomatic improvement (Géger & Crick, 2001; Cohen & Crawford, 2005). The purpose of this study was to determine differences in functional impairment in adulthood between community youth who showed improvement in personality disorder psychopathology during the transition to adulthood, those who did not, and those who first developed such a disorder as young adults.

METHOD

Participants and procedures  Participants in the Children in the Community (CIC) study were a 1975 residence-based sample of mothers of children aged 1–10 years, in two upstate New York counties, supplemented with an additional sample residing in poor urban neighbourhoods to compensate for those lost to follow-up. The 821 mothers and one randomly sampled child were interviewed three times in their homes by trained lay interviewers: in 1983 (mean offspring age 13.7 years, s.d. = 2.8); between 1985 and 1986 (mean offspring age 16.3 years, s.d. = 2.8); and between 1991 and 1993 (mean offspring age 22.1 years, s.d. = 2.7) (Fig. 1). Comprehensive assessments of personality disorder were completed with 760, 749 and 719 of the offspring respectively at each of these three assessments. The families were representative of families in the north-eastern USA with regard to socio-economic status and most demographic variables, but they also reflected the sampled region, with high proportions of participants who were Catholic (54%), White (91%) and rural residents (25%) (Cohen & Cohen, 1996).

The findings reported here are based on data from 658 individuals who were interviewed a fourth time between 2001 and 2004 (mean age 33.1 years, s.d. = 2.9). After home interviews assessing a wide range of psychosocial variables had been completed, psychiatric interviews were administered over the telephone by professionals with a master’s or doctorate degree in social work or clinical psychology and at least 10 years of experience in the administration of semi-structured psychiatric research interviews. The 658 individuals in this sample did not differ from the remainder of the original sample with regard to the prevalence of behavioural or emotional problems at earlier assessments. The institutional review boards of the Columbia University College of Physicians and Surgeons and the New York State Psychiatric Institute approved the study procedures. Written informed consent or assent was obtained from all participants after the interview procedures had been fully explained.

Fig. 1  Children in the Community study assessment waves, at mean ages 14, 16, 22 and 33 years.
Assessments

Assessment of personality disorders

Personality disorders were first assessed in the CIC sample in 1983, when no instrument existed to measure these disorders in adolescents. Accordingly, the disorders were measured with relevant parent- and youth-reported items from the study’s longitudinal protocol that were selected to correspond with DSM–III (American Psychiatric Association, 1980) criteria for Axis II disorders. Additional items were added to the protocol from the Personality Diagnostic Questionnaire (PDQ; Hyler et al, 1988) and an early version of the Structured Clinical Interview for Personality Disorders (Spitzer & Williams, 1986), adapted to make them age-appropriate (for a detailed history of how symptom scales and diagnostic algorithms were developed, see Crawford et al, 2005). Following publication of DSM–IV (American Psychiatric Association, 1994), the personality disorder symptom scales and diagnostic algorithms were modified to maximise correspondence with DSM–IV diagnostic criteria and to produce consistent repeated measures of personality disorder assessed at mean ages 14, 16 and 22 years. From each data collection period 152 items were available to assess 88 (93.6%) of the 94 DSM–IV criteria for Axis II disorders. The concurrent validity of the CIC assessment procedure has been supported by findings showing that personality disorders are associated with impairment, distress and increased risk of Axis I disorders (Bernstein et al, 1993; Kase et al, 1999, 2001). The predictive validity of the assessment has been supported by findings indicating that adolescent personality disorders are associated with elevated risks of Axis I disorders, criminal or violent behaviour, and suicidal behaviour during early adulthood (Johnson et al, 1999, 2000).

The Structured Clinical Interview for DSM–IV Axis II Personality Disorders (SCID–II; First et al, 1995a) was first used in this sample to assess personality disorders at mean age 33 years. The SCID–II is a two-stage diagnostic procedure which includes a screening questionnaire, followed by a semi-structured interview to determine whether affirmative responses on the questionnaire indicate the presence of clinically significant symptoms. The SCID–II interview test-retest reliability has been found to be satisfactory: $k = 0.51$ for ‘any personality disorder’ in patients and $k = 0.48$ in non-patients (First et al, 1995b). For this study 40 interviews were tape-recorded (with the respondent’s permission) and then rated again by a second interviewer to assess interrater agreement. Interrater reliability was satisfactory for ‘any personality disorder’: $k = 0.62$ (Crawford et al, 2005).

Personality disorders at mean age 33 years also were measured with the pool of self-report items assessed in the CIC longitudinal protocol. However, because parent interviews were no longer conducted at this age, CIC scales and algorithms were augmented with other self-report items to replace the parent-reported data (Crawford et al, 2005). When CIC and SCID–II diagnoses were compared, concordance for ‘any personality disorder’ ($k = 0.45$) was modest, but approached the SCID–II interview’s $k$ value for test-retest reliability in non-patients. Concordance rates for any cluster A diagnosis ($k = 0.41$) and any cluster B diagnosis ($k = 0.60$) surpassed comparable findings in 12 out of 13 studies reviewed by M odestin et al (1998). Concordance for cluster C diagnoses ($k = 0.29$) was closer to the published average.

Assessment of Axis I disorders

Axis I disorders at mean age 33 years were assessed with the non-patient version of the Structured Clinical Interview for DSM–IV Axis I Disorders (SCID-I/NP; First et al, 1996).

Assessment of global functioning and impairment

Overall functioning at mean age 33 years was assessed with the Global Assessment of Functioning Scale (GAFS; American Psychiatric Association, 2000). The GAFS evaluates functioning during the past year on a scale from 1 to 100; scores higher than 70 indicate satisfactory mental health, good overall functioning and no more than minimal or transient distress or impairment. Scores between 61 and 70 signify mild impairment or distress, scores between 51 and 60 indicate moderate impairment and scores below 51 indicate severe impairment. In adult out-patients GAFS scores have high rates of interrater reliability (intraclass correlation 0.86) and are significantly related to responses on the Symptom Checklist – 90 – Revised general severity index (Hilsenroth et al, 2000). In our present study clinicians completed the GAFS after conducting the SCID–I and SCID–II interviews.

Psychosocial impairment was assessed with a six-item self-report index ($k = 0.86$) adapted from items used in the Medical Outcomes Study Short-Form General Health Survey (Stewart et al, 1988) and the Disorganizing Poverty Interview (Kogan et al, 1977). Items assess difficulties in carrying out responsibilities, completing tasks and getting along with others, disorganisation and lack of control, recurrent health or safety risks, and recurrent behaviour leading to embarrassment or shame. Items are rated on five-point Likert scales of frequency, ranging from 0 (‘never’) to 4 (‘always or almost always’), which produce total impairment scale scores ranging from 0 to 24.

Assessment of socio-economic status

An index of socio-economic status was computed as the standardised sum of standardised measures of years of maternal and paternal education, income and occupational status.

Data analysis

Analyses of covariance (ANCOVAs) were conducted to investigate associations between the diagnostic stability of ‘any personality disorder’ with clinician-reported GAFS scores and self-reported impairment scores at mean age 33 years. In addition to adjusting for effects of age, gender, and socio-economic status, these ANCOVAs controlled for the presence of an Axis I disorder at mean age 33 years, in order to assess the impact of personality disorders on functioning independently of Axis I psychopathology. Individuals were classified as having ‘persistent disorder’ if they had any personality disorder diagnosis at mean age 14, 16 or 22 years and any personality disorder diagnosis at mean age 33 years. Individuals who had any personality disorder diagnosis by mean age 22 years but not at mean age 33 years were classified as having personality disorder in remission. Individuals who had any personality disorder diagnosis at mean age 33 years, but not at prior assessment were classified as having adult-onset disorder. Thus defined, there were 64 participants with persistent personality disorder, 185 in remission, 38 with adult-onset disorder and 371 with no personality disorder at any assessment interval. Analyses were conducted at the level of ‘any personality disorder’ because of inadequate numbers of cases of specific disorders or disorders from each DSM–IV.
cluster, once cases were divided into persistent, remitted and adult-onset personality disorders.

To determine whether the change in diagnostic procedures for personality disorders between assessment 3 and assessment 4 (i.e. from CIC scales to SCID-II) had an effect on the findings, we replicated the above analyses using only the CIC scales at all time points to create personality disorder stability groups as described above.

RESULTS

Demographic characteristics of the sample, the prevalence of Axis I disorders and disorder clusters at mean ages 16, 22 and 33 years, the prevalence of Axis I disorders at mean age 33 years and mean GAFS and impairment scores at age 33 years are all presented in Table 1. The rates of co-occurring Axis I disorders by SCID-II personality disorder group were as follows: no personality disorder 23.1%; personality disorder in remission 30.3%; adult-onset personality disorder 57.9%; and persistent personality disorder 70.3%.

Personality disorder stability from adolescence to adulthood

Global functioning and impairment outcomes

A consistent pattern of findings was obtained with regard to the association of overall personality disorder stability with GAFS scores and total impairment scale scores (Table 2). The poorest functioning and greatest impairment were observed among individuals with persistent disorder (i.e. those identified as having a personality disorder by mean age 22 years and also at mean age 33 years); these individuals had significantly lower GAFS scores (mean 58.73) than those in the other groups, and their functioning was moderately to severely impaired. Their mean impairment scores were nearly twice as high as those of participants who were never identified as having a personality disorder.

Participants identified as having a personality disorder in remission at mean age 33 years had significantly lower GAFS and higher impairment scale scores than the individuals who were not identified as having a personality disorder at any assessment. However, the impairment experienced by those in remission was relatively mild and did not tend to be clinically significant (mean GAFS score 72.92).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, % (n)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>47 (309)</td>
</tr>
<tr>
<td>Female</td>
<td>53 (349)</td>
</tr>
<tr>
<td>Ethnicity, % (n)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>91.2 (600)</td>
</tr>
<tr>
<td>African American or other</td>
<td>8.8 (58)</td>
</tr>
<tr>
<td>Axis I disorder prevalence at mean age 33 years, % (n)</td>
<td>26.6 (175)</td>
</tr>
<tr>
<td>Personality disorder prevalence, % (n)</td>
<td></td>
</tr>
<tr>
<td>By mean age 16 years</td>
<td></td>
</tr>
<tr>
<td>Cluster A</td>
<td>32.7 (210)</td>
</tr>
<tr>
<td>Cluster B</td>
<td>16.7 (107)</td>
</tr>
<tr>
<td>Cluster C</td>
<td>17.9 (115)</td>
</tr>
<tr>
<td>Other PD</td>
<td>10.4 (67)</td>
</tr>
<tr>
<td>At mean age 22 years</td>
<td></td>
</tr>
<tr>
<td>Cluster A</td>
<td>3.3 (20)</td>
</tr>
<tr>
<td>Cluster B</td>
<td>4.6 (28)</td>
</tr>
<tr>
<td>Cluster C</td>
<td>5.2 (32)</td>
</tr>
<tr>
<td>Other PD</td>
<td>2.2 (14)</td>
</tr>
<tr>
<td>At mean age 33 years</td>
<td></td>
</tr>
<tr>
<td>Cluster A</td>
<td>15.5 (102)</td>
</tr>
<tr>
<td>Cluster B</td>
<td>6.5 (43)</td>
</tr>
<tr>
<td>Cluster C</td>
<td>10.6 (70)</td>
</tr>
<tr>
<td>Other PD</td>
<td>5.2 (34)</td>
</tr>
</tbody>
</table>

Table 1 Sample and variable characteristics

GAF, Global Assessment of Functioning Scale; PD, personality disorder.
1. Anxiety, disruptive and mood disorders were assessed at mean age 33 years.
2. Personality disorder was not considered present by mean age 35 years unless the DSM-IV diagnostic criteria were met at mean ages 34 and 36 years; or diagnostic criteria for a personality disorder were met at one of these assessments and were missed by no more than one criterion at the other assessment. The prevalence of personality disorder at mean age 34 years was 22.2% (167 cases out of a total 623) and the prevalence at mean age 36 years was 18.5% (122 cases out of a total 670).
3. Assessed at mean age 33 years.
4. Assessed at mean ages 34 and 36 years.

Participants identified as having a personality disorder at mean age 33 years but not at the prior assessments (i.e. adult-onset disorder) had an intermediate level of impairment, greater than that of individuals whose disorder was in remission but less than those with persistent disorder. The impairment in functioning experienced by this group was clinically significant, in the mild to moderate range (mean GAFS score 64.93).

Stability of disorder and functioning using CIC scales

The greatest discrepancy in the identification of the personality disorder groups came in the adult-onset category, in which 38 adult-onset cases were identified using SCID-II compared with only 12 using the CIC personality disorder scales. A consistent pattern of findings was obtained, however, using only the latter scales to create personality disorder stability groups (see Table 2).

DISCUSSION

This study is the first to investigate the effect of personality disorder stability on functioning over time in a community sample. Traditionally, because these disorders were assumed to be stable and enduring, consideration of the effects of improvement in psychopathology was not an
issue. Skodol et al. (2005) have shown in a large sample of patients that, overall, impair- 
ment in psychosocial functioning, espe- 
cially in interpersonal relationships, was 
more stable than the personality disorder 
itself. Nevertheless, for patients with bor- 
derline{d}erline personality disorder who showed 

improvement in psychopathologic symp-
toms, some improvement was seen in func-
tioning. Zanarini et al. (2005) have also 

demonstrated that patients with borderline 

personality disorder who experienced a 
symptomatic remission during a 6-year 

follow-up period functioned significantly 

better in social relationships and at work 

than similar patients with no remission. 

The prevalence of personality disorder 
in our sample ranged from 27.2% at mean 
age 14 years to 15.5% (SCID-II) by mean 
age 33 years. In a review of eight epide-
miological studies of personality disorder in 
adults, Torgeron (2005) found the preva-

lence for "any personality disorder" ranged 

from 3.9% to 22.7%, with a median 

prevalence of 11.6% and a pooled mean 

prevalence of 12.3%. The prevalence esti-

mates of personality disorders in our study 

that correspond to adulthood are well with-
in this range. Personality disorders among 

adults in the community have been shown 

to be associated with reduced quality of 

life, as reflected in subjective well-being, 

self-realisation, relationship to friends, 

social support, negative life events, re-

relationship to family of origin and neigh-

bourhood quality (Torgeron et al., 2001). 

Of the participants who were diagnosed 
in our study with a personality disorder by 

the age of 22 years, only 25.7% retained a 

personality disorder diagnosis by age 33 

years (on average 11 years later). In the 

studies of patients reviewed by Perry 

(1993), McDavid & Pilkonis (1996) and 

Grilo et al. (1998) about 50% of patients re-
tained their diagnoses over periods ranging 

from 6 months to 15 years. In these studies, 

mostly of borderline personality disorder, 

the lowest stability rate was found in ad-
elle<
de-

olescence, when personality is often con-

sidered to be in flux. In general, the stability 
of personality disorders has been found to 
have a strong negative correlation with 
the length of the follow-up period. Thus, 
in our study the substantial rate of 

remission probably reflects both the young 
age of the sample when the personality 

disorders were first diagnosed and the 

length of the follow-up. 

The findings of our study shed light on 
the association between stability of disorder 
during the transition from adolescence to 
adulthood and functioning and impairment 
in adulthood. First, our findings suggest 
that adults in the community with persist-
ent personality disorder (i.e. that has been 
present since adolescence or early adult-
hood) are likely to experience poor func-
tioning and marked (moderate to severe) 
impairment in adulthood. The mean GAFS 
score obtained in this study for people with 
persistent personality disorder (58.7) is 
comparable with the mean GAFS score in 
a large sample of people (predominantly 
out-patients) with one of four types of per-
sonality disorder (57.6) reported by Skodol 
et al. (2002). These difficulties in functioning 

are not likely to be attributable to age, gen-
der or socio-economic status during adoles-
cence. Furthermore, the effects were 

independent of Axis I disorders at mean 
age 33 years, thus underscoring the impor-
tance of recognising and treating Axis II dis-
orders regardless of whether or not they 
occur together with Axis I disorders. These 
results are also consistent with those of 
Skodol et al. (2002) in that impairment in 
various domains of functioning in patients 
with personality disorders could not be ex-
plained by comorbid Axis I disorders, and 
with those of Trull (2001), who found 
similarly that borderline features in a non-
patient sample accounted for significant 
variance in functioning beyond that 
accounted for by Axis I disorders. Second, 
our findings suggest that individuals in the 
community who experience the onset of a 
personality disorder during adulthood are 
also likely to experience mild to moderate 
impairment that is clinically significant, 
although not as severe in most cases as that 
in earlier-onset and persistent personality 
disorder. Our findings are also of interest 
because they suggest that people with personality disorder who experience remission of 
symptoms of the disorder during the transi-
tion to adulthood may experience relatively 
little residual impairment by middle adult-
hood. That improvement in symptoms 
eventually will have a beneficial effect on 
functioning provides a reason to be optimis-
tic that many adolescents and young 
adults who exhibit personality disorder 
psychopathology may be able to function

### Table 2: Impairment and functioning at mean age 33 years by prior and current clinical personality disorder status (n = 658)

<table>
<thead>
<tr>
<th>Global indices of functioning and impairment</th>
<th>No PD</th>
<th>PD in remission</th>
<th>Adult-onset PD</th>
<th>Persistent PD</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCID-II diagnosis at age 33 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( n )</td>
<td>371</td>
<td>185</td>
<td>38</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td>Adjusted mean scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAFS</td>
<td>76.12 (SE: 2.4)</td>
<td>72.92 (SE: 2.4)</td>
<td>64.93 (SE: 1.5)</td>
<td>58.73 (SE: 1.5)</td>
<td>78.05 (df: 3, 650)</td>
</tr>
<tr>
<td>Psychosocial Impairment Scale</td>
<td>2.82 (SE: 0.4)</td>
<td>3.41 (SE: 0.3)</td>
<td>4.16 (SE: 0.3)</td>
<td>5.33 (SE: 0.3)</td>
<td>11.94 (df: 3, 650)</td>
</tr>
<tr>
<td>CIC PD diagnosis at age 33 years</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( n )</td>
<td>396</td>
<td>184</td>
<td>12</td>
<td>60</td>
<td></td>
</tr>
<tr>
<td>Adjusted mean scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAFS</td>
<td>75.21 (SE: 2.4)</td>
<td>71.78 (SE: 2.4)</td>
<td>64.92 (SE: 1.5)</td>
<td>62.05 (SE: 1.5)</td>
<td>37.05 (df: 3, 654)</td>
</tr>
<tr>
<td>Psychosocial Impairment Scale</td>
<td>2.87 (SE: 0.4)</td>
<td>3.44 (SE: 0.4)</td>
<td>5.50 (SE: 0.4)</td>
<td>5.47 (SE: 0.4)</td>
<td>13.54 (df: 3, 654)</td>
</tr>
</tbody>
</table>

CIC, Children in the Community; GAFS, Global Assessment of Functioning Scale; PD, personality disorder; SCID-II, Structured Clinical Interview for DSM-III-R Personality Disorders.

1. Adjusted for age, gender, socio-economic status and presence of an Axis I disorder at mean age 33 years.
2. Significantly different from no PD group (\( p < 0.05 \)).
3. Significantly different from PD in remission group (\( p < 0.05 \)).
4. Significantly different from adult-onset PD group (\( p < 0.05 \)).
5. Significantly different from persistent PD group (\( p < 0.05 \)).

\( \cdot \) 0.05.}
nearly as well as people without a history of such disorder. Declines in symptom levels from adolescence through early adulthood (e.g. Johnson et al., 2000a) are consistent with the hypothesis that many people ‘outgrow’ personality disorders during the transition from adolescence to adulthood as a result of maturation and socialisation, which promote the development of a stable sense of self and improved interpersonal, coping and impulse-control skills. Because personality disorder can often be treated effectively (Perry et al., 1999) and treatments have been adapted for adolescents with some success (Johnson et al., 2006), our findings suggest that mental health professionals who work with adolescents and young adults might be well advised to conduct an assessment of symptoms of personality disorder in these patients. Since those with the highest symptom levels for their age groups remain most at risk of persisting personality disorder (Crawford et al., 2004) and impairment (Johnson et al., 1999, 2000b), appropriate intervention with these patients might assist more young people to make the transition to adulthood successfully, with fewer interpersonal, occupational and other difficulties.

A potential limitation of this study is that clinician-administered, semi-structured interviews for personality disorder were conducted at the final assessment only. In order to determine whether the findings were influenced by change in the assessment of these disorders from the CIC symptom scales to the SCID–II clinical interviews at mean age 33 years, the analyses were repeated using only the CIC scales at each time point. The basic pattern of findings regarding the relationship between persistence of personality disorder and impairment in functioning was replicated in this additional set of analyses, providing strong support for the observed associations between stability of disorder and impairment in functioning. Extensive assessments of various domains of psychosocial functioning were not possible, but the most widely used measure of global functioning (the GAFS) was employed. A detailed description of the course of personality disorder psychopathology over the follow-up interval was not feasible using this study’s design. Thus, the assumption that personality disorders presenting before age 22 years and at age 33 years are in fact ‘persistent’ – as opposed to intermittent or recurrent – may not be justified. Furthermore, stability estimates are limited by the reliability of the personality disorder measures. The rates of improvement observed in rigorous, follow-along clinical studies, however, exceed by a substantial margin those that would be predicted on the basis of measurement error alone (Grilo et al., 2004). Finally, it was not possible to determine the association of the persistence of specific personality disorders or disorder clusters with impairment in adulthood, owing to limited statistical power.

It will be important for future studies to investigate the determinants of personality disorder stability. Identification of psychosocial factors that might promote reductions in symptom levels during the transition to adulthood might lead to new insights about how young people acquire the stable identities and interpersonal, coping and impulse-control skills that are characteristic of optimal development and functioning. It will also be of interest to examine the developmental course of adult-onset personality disorders in greater detail. Although a few investigators have examined predictors of later-onset personality disorders, such as the presence of Axis I disorders during adolescence (Kasen et al., 1999, 2001), many questions about the development and sequelae of adult-onset personality disorders remain unanswered.

ACKNOWLEDGEMENTS

This research was supported by grant MH-36971 from the National Institute of Mental Health to PC.

REFERENCES

American Psychiatric Association (1980) Diagnostic and Statistical Manual of Mental Disorders (3rd edn) (DSM–III), APA.

American Psychiatric Association (1994) Diagnostic and Statistical Manual of Mental Disorders (4th ed), APA.


Lenzenweger, M. F. (1999) Stability and change in personality disorder features: findings from a


Cognitive remediation therapy in schizophrenia

Randomised controlled trial

TIL WYKES, CLARE REEDER, SABINE LANDAU, BRIAN EVERITT, MARTIN KNAPP, ANITA PATEL and RENEE ROMEO

Background  Cognitive difficulties are prevalent in people with a diagnosis of schizophrenia and are associated with poor long-term functioning.

Aims  To evaluate the effectiveness of cognitive remediation therapy on cognitive difficulties experienced by people with schizophrenia.

Method  Participants with a diagnosis of schizophrenia, a social behaviour problem and a cognitive difficulty (n=85) were randomised to 40 sessions of cognitive remediation or treatment as usual in a single-blind randomised controlled trial. Working memory, cognitive flexibility and planning were measured at weeks 0, 14 and 40.

Results  There were durable improvements in working memory (advantage I.33 points, 95% CI 0.43–2.16, standardised effect size 0.34) as well as an indication of improvement in cognitive flexibility. Memory improvement predicted improvement in social functioning. Costs were lower in the cognitive remediation group following therapy but rose at follow-up. However, cost-effectiveness analyses showed that improvements in memory were achieved at little additional cost.

Conclusions  Cognitive remediation therapy is associated with durable improvements in memory, which in turn are associated with social functioning improvements in people with severe mental illness.

Declaration of interest  None.

Both longitudinal and cross-sectional studies of patients with a diagnosis of schizophrenia suggest that cognitive performance is poor and remains poor over the course of the disorder, and that such deficits, particularly in memory, limit functioning outcomes and the rehabilitation of particular life skills such as work and social functioning (Green et al, 2000; Wykes & Reeder, 2005). In order to remove this rate limitation a new rehabilitation technology, cognitive remediation therapy, was developed with the aim of improving cognition and thereby increasing the likelihood of improved functioning outcomes. Cognitive remediation therapy is an umbrella term for a number of different interventions defined by their procedural characteristics such as use of a therapist, use of a computer and the method of training. There is some evidence of efficacy for face-to-face therapy from small studies; however, no large study has investigated the effects and cost-effectiveness of face-to-face therapy. In addition, studies have been limited to people who fulfilled narrow entry criteria in terms of their cognitive difficulties. It is, therefore, not yet possible to identify whether this form of cognitive therapy will have an impact on those with a spectrum of cognitive difficulties. The key effectiveness questions for cognitive remediation therapy concern its likely success when the recipients have a variety of cognitive difficulties as well as a diagnosis of schizophrenia, and whether any cognitive improvements have an impact on functioning.

METHOD

Study design

We carried out a single-blind, randomised controlled trial of a new therapy to improve cognition in people with schizophrenia by comparing a group receiving 40 sessions of therapy with a group who received only usual treatment. We tested whether cognitive skills improved in the intervention group and whether this improved cognitive skill led to improvements in symptoms, social functioning and self-esteem. After baseline assessment, participants were randomised to either treatment or control and were then assessed at 14 weeks (post-therapy) and 40 weeks (6 months after therapy discontinuation). The trial registration number is ISRCTN 44277627.

Inclusion and exclusion criteria

We recruited participants from local community mental health teams in the South London and Maudsley National Health Service Trust in a structured geographical rotation from February 1999 to December 2002. Patients were included if they had been in contact with the services for at least 1 year, were at least 17 years old, had a diagnosis of schizophrenia based on DSM-IV (American Psychiatric Association, 1994) and evidence of both social functioning, defined as a problem on the Social Behaviour Scale (SBS; Wykes & Sturt, 1986), and thinking difficulties. Thinking difficulties were defined as a poor memory score on the Rivermead scale (Wilson et al, 1999), and/or cognitive flexibility on the Wisconsin Card Sorting Test (WCST; Heaton et al, 1993) below the 16th centile, and/or a poor score on the Hayling Sentence Completion Test (Burgess & Shallice, 1996).

Therapy

Several programmes are available to test, but the one chosen was first developed in Australia (Delahunty & Morice, 1993) and incorporates the teaching of strategic information processing, which has been identified as the training more likely to produce larger cognitive benefits following analysis in the most comprehensive review (Krabbenb & Aleman, 2003). This is a promising programme because, unlike the others, it has been shown to have specific effects when tested in a randomised controlled trial against another psychosocial programme (Wykes et al, 1999, 2003).

Therapy consisted of 40 face-to-face sessions, each involving a number of paper and pencil tasks that provide practice in a variety of cognitive skills that are set out in a manual (Delahunty et al, 2002). Therapy was delivered to individuals on at least 3 days per week until 40 sessions were completed. The therapists were graduate psychologists who had followed a dedicated training programme involving theory,
The three main outcome measures were:
(a) teaching (or facilitating learning of) new efficient information processing strategies;
(b) individualising therapy;
(c) aiding the transfer of cognitive gains into the real world.

The programme consists of three modules: cognitive flexibility, working memory and planning (Delahunty et al, 2002; Reeder et al, 2004). In each module there is a series of tasks, graded from 'extremely easy' to 'easy', so that an errorless learning environment can be provided. In the cognitive flexibility module, patients are given practice in engagement, disengagement and re-engagement activities for a particular cognitive set or between two sets. The working memory module requires the person to maintain two sets of information simultaneously and to carry out transformations on a held information set. The planning module consists of tasks in which the participant has to plan a sequence of moves to acquire a goal. The emphasis in this module is to organise information and to create and use sub-goals. One major change from the therapy as administered in previous studies (e.g. Wykes et al, 1999) was the emphasis of therapists on the possible uses of the strategies being taught within the participants' own lives, for example in going shopping. This was achieved by encouraging the participants to reflect on how the skills learnt in therapy might be used to achieve real-life goals (see Wykes & Reeder, 2005 for further details).

Therapist fidelity was checked against the records completed at the end of each session, the task sheets produced during the sessions and by direct observation. Participants did receive therapy that complied with the manual, and the majority of the tasks were delivered for most participants. These high levels of fidelity were maintained and supported by weekly supervision.

**Outcome measures**
The three main outcome measures were:
(a) cognitive flexibility - categories achieved from the Wisconsin Card Sorting Test;
(b) planning - the profile score from the Behavioural Assessment of the Dysexecutive Syndrome (BADS; Wilson et al, 1996);
(c) working memory - total raw score on the Digit Span test of the Wechsler Adult Intelligence Scale III (WAIS-III; Wechsler, 1981).

In addition to the main outcomes we also collected data on symptoms from the Positive and Negative Syndrome Scale (PANSS; Kay et al, 1987), a self-esteem score from the Rosenberg Self-Esteem Scale (SES; Rosenberg, 1965) and level of social functioning from the Social Behaviour Scale (Wykes & Sturt, 1986). Health, social care and criminal justice system resource use were assessed using the Client Service Receipt Inventory (Beecham & Knapp, 1992) retrospectively from healthcare staff or records, and/or by participant self-report for the relevant assessment intervals. Unit costs (at 2000–2001 levels) based on national statistics were attached to all resource use to calculate total health and societal costs.

**Procedure**
All participants gave written informed consent prior to inclusion in the trial. After baseline assessment participants were randomly allocated by an independent statistician using a concealed randomisation method. Participants assigned to the cognitive remediation condition received therapy within 2 weeks of randomisation. Therapy continued for 40 sessions (approximately 12 weeks). In addition to the assessments on outcome measures, data were also collected on clinical history, demographic characteristics and premorbid IQ as assessed on the National Adult Reading Test (NART; Nelson & Willison, 1991).

**Protecting against bias**
Symptoms were rated by a psychiatrist unaware of group allocation, who was based in a different building to the other researchers and the independent site of randomisation. Participants were informed that they should not reveal their group allocation prior to each assessment and none did so for the symptom assessment. Cognitive data were collected by independent assessors who, although initially masked to group allocation, were not unaware of all allocations since some participants revealed their randomisation group at the post-treatment assessment point. However, as these data were collected either by computer or under clear guidance and instruction, the effect of the revealing of group allocation is unlikely to be significant. Social behaviour data were collected from keyworker or relative informants who were independent of the trial but not masked to group allocation.

**Sample size and power of the study**
Previous studies of this programme have suggested that there will be improvement in both groups with repeated testing over time. We have therefore used the outcome data reported by Wykes et al (1999) to define a clinically significant difference as 71% of the experimental group improving compared with 31% of the control group. This difference is considered to be a clinically significant difference in proportions considering the amount of therapy time that would need to be allocated. This is an odds ratio of 0.184. We estimated that a sample of 29 people per group would have 80% power at the 5% significance level to detect this difference. The sample size was increased to 42 to take into account a possible 30% withdrawal rate.

**Statistical analyses**
Participants were analysed in the treatment group to which they were randomised irrespective of whether they adhered to their treatment. All outcome measures were analysed using linear mixed modelling with models fitted using restricted maximum likelihood methods based on the assumption of normality for the error terms. Models included baseline values of the outcome measure, and symptoms considered possibly to affect cognitive outcome following therapy as explanatory variables. The experimental factors, randomisation group (therapy or control) and time (post-treatment or follow-up) were included in the model as fixed main effects and a group × time interaction. In addition, random effects for participants were included. A significant interaction term implies a differential intervention effect at the two post-therapy time points. Where the interaction effect was not significant, the corresponding model was refitted excluding this term, to assess the overall group effect. A main effect of randomisation group would then be interpreted as an effect of the intervention therapy consistent across both time points.

We also chose to investigate whether the effects of treatment meant that cognitive scores were then within the normal range. This differential improvement rate was tested by chi-squared tests for each
main outcome measure, by investigating the changes to normal scores of all those who had abnormal scores at baseline.

Costs

Differences in mean costs and 95% confidence intervals were obtained using non-parametric bootstrapping techniques to account for any non-normality in their distribution (1000 repetitions), using Stata version 8.0 for Windows. Costs (including the costs of therapy where applicable) were adjusted for baseline values of equivalent cost categories and baseline total PANSS score.

To allow a cost-effectiveness analysis based on a more meaningful interpretation of the primary outcome measure, we also compared the percentage of ‘improvers’ in each group based on WAIS–III Digit Span raw scores (improvers were defined as those gaining 2 points or more on this measure since baseline). This was based on a relatively large effect size of 0.7 and was chosen because recent studies suggest that improvements of this size may contribute to functional improvements (Bryson & Bell, 2003). For cost-effectiveness ratios based on ‘improvers’, percentages were also compared using non-parametric bootstrapping, and were adjusted for baseline WAIS–III Digit Span raw score and total PANSS score.

It was not necessary to calculate ratios in scenarios in which one group had both lower costs and better outcomes, as the decision regarding which treatment is preferred is intuitively clear. Where one group had both higher costs and better outcomes, the additional cost per additional 1% of improvers on the WAIS–III Digit Span raw score was calculated by dividing the mean difference in costs by the mean difference in percentage of improvers.

Mechanisms underlying social functioning change

Finally, cognitive change is predicted to have an impact on social functioning. In order to test this model a regression was carried out with follow-up social functioning outcome as the dependent variable, therapy group as a factor, and cognitive change over the treatment period and baseline levels of social functioning and symptoms as covariates. The model first tested a group-dependent cognition effect by means of an interaction between cognitive change and group. If the interaction effect was not significant then it was excluded and the model rerun to assess the overall effect of cognitive change on functioning.

RESULTS

Eighty-five participants were recruited to the trial, of whom 43 were randomised to cognitive remediation therapy and 42 to the control condition (Fig. 1). Nearly three-quarters of the sample were men (73%; n = 62) and the mean age was 36 years; 47% (n = 40) were living in independent accommodation or with their family. Most had no experience of a stable relationship, and 40% (n = 34) had never lived independently. About half (n = 44) had been in touch with the psychiatric services for at least 10 years. The participants were therefore severely impaired in overall functioning, although some people had made some achievements such as marrying or having independent living arrangements.

Not all participants agreed to complete all assessments for a variety of reasons, including delusional ideation as well as refusal. There was no difference between the groups in the rate of withdrawals (\( \chi^2 = 0.047, \) d.f. = 1, \( P = 1.0 \)) and none of the potential baseline variables (cognition outcomes, self-esteem, social behaviour or symptoms) predicted withdrawal from the study (probability levels all above 0.16).

Overall, the intervention group participants received a mean of 36.9 (0–40) sessions of therapy, with a mean of 3.8 per week for those who started therapy, and at least 30 sessions being received by 93% of the sample.

Table 1 shows the types of primary medications and the mean dosage in chlorpromazine equivalents for those whose primary medication was a typical antipsychotic agent. Two people in the therapy group and one person in the control group received both typical and atypical medication. Of those prescribed typical antipsychotics, 11 received them in the form of depot preparations (4 in the therapy group and 7 in the control group).

As would be expected after random treatment allocation, the main cognitive outcomes were similar in the two groups at baseline, as were social behaviour and self-esteem. However, despite randomisation, the level of symptoms appeared to be greater in the therapy group (Table 2), but this variable was already included as a covariate in all models considered. The baseline means for the other main and secondary outcomes are presented in Table 2. The NART scores were 92.7 (s.d. = 13.3) for the therapy group and 92.4 (s.d. = 12.7) for the control group.

Fig. 1 Study profile.
Outcomes of therapy

Table 3 shows the group comparison results for all the outcomes. All statistically significant changes show an advantage for cognitive remediation. Working memory shows an improvement across both post-treatment time points and cognitive flexibility an improvement at the follow-up time point. Both differences are small to moderate effects. For working memory change the number needed to treat (NNT) is 3.1 to produce a clinical change of at least 2 points on the Digit Span test. For cognitive flexibility the NNT is 6.7 to improve by at least two categories on the WCST at follow-up.

Using the same mixed models analyses, drug effects were investigated using drug type as an additional explanatory variable. Neither working memory nor WCST outcomes were related to the type of medication prescribed (defined as either typical or atypical or when depot preparations were considered). However, there was a significant drug x group interaction ($F_{(2,70)=4.4}$, $P=0.016$) for planning scores. Further investigation of these effects suggests that cognitive remediation therapy had an overall effect on planning for those who received either clozapine or typical medication that was absent for those who received other atypical medications.

Normal score attainment following therapy

The normal range for the main outcome tests was conservatively estimated from the test manuals. For working memory this was within 1 standard deviation of the mean normal score, for cognitive flexibility it was above the 5th percentile and for planning it was above the low average score.

For working memory there were 27 participants in the therapy group and 19 in the control group who had abnormal working memory scores at baseline. Following the intervention, there was an advantage to therapy which was significant at the post-therapy assessment but failed to reach significance at follow-up (post-treatment: 43 therapy $v.$ 11% control, Fisher’s exact test $P<0.037$; follow-up: 32 therapy $v.$ 7% control, Fisher’s exact test $P=0.10$). For cognitive flexibility there was no difference at either the post-treatment or the follow-up assessment (post-treatment: 15 therapy $v.$ 17% control; follow-up: 17 therapy $v.$ 21% control). For planning, although almost double the number of people in the therapy group had a normalised score, there was no statistically significant effect (post-treatment: 32 therapy $v.$ 17% control; follow-up: 36 therapy $v.$ 19% control).

Other outcomes

These analyses are designed to detect whether there is a direct effect of cognitive remediation therapy on functioning outcomes, irrespective of the level of cognitive improvement detected. The results of the analyses are shown in Table 3. For both symptoms and self-esteem the results were in the expected direction, with the therapy group improving compared with the control group at post-treatment. There was evidence of an interaction such that any differential improvement at the baseline assessment was conserved through the therapy phase, as shown in Table 3.

Table I: Antipsychotic medication provided at baseline

<table>
<thead>
<tr>
<th>Therapy group (n=43)</th>
<th>Control group (n=42)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atypical medication, n</td>
<td></td>
</tr>
<tr>
<td>Clozapine</td>
<td>16</td>
</tr>
<tr>
<td>Olanzapine</td>
<td>8</td>
</tr>
<tr>
<td>Risperidone</td>
<td>7</td>
</tr>
<tr>
<td>Amisulpride</td>
<td>1</td>
</tr>
<tr>
<td>Quetiapine</td>
<td>2</td>
</tr>
<tr>
<td>Typical medication</td>
<td></td>
</tr>
<tr>
<td>Mean dosage, mg CPZeq</td>
<td>368</td>
</tr>
<tr>
<td>CPZeq, chlorpromazine equivalent.</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Cognitive, secondary and functioning outcome scores

<table>
<thead>
<tr>
<th>Baseline assessment</th>
<th>Post-treatment assessment</th>
<th>Follow-up assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRT group</td>
<td>Controls</td>
<td>CRT group</td>
</tr>
<tr>
<td>n Mean (s.d.)</td>
<td>n Mean (s.d.)</td>
<td>n Mean (s.d.)</td>
</tr>
<tr>
<td>Memory (Digit Span)$^1$</td>
<td>43 14.2 (3.9)</td>
<td>42 15.1 (3.9)</td>
</tr>
<tr>
<td>Cognitive flexibility (WCST)$^2$</td>
<td>43 2.4 (1.5)</td>
<td>42 2.2 (1.3)</td>
</tr>
<tr>
<td>Planning (BADS)$^3$</td>
<td>43 11.7 (4.6)</td>
<td>42 12.7 (5.1)</td>
</tr>
<tr>
<td>Symptoms (PANSS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>43 62.9 (16.4)</td>
<td>42 56.9 (14.7)</td>
</tr>
<tr>
<td>Negative symptoms</td>
<td>43 18.6 (17.3)</td>
<td>42 16.6 (17.2)</td>
</tr>
<tr>
<td>Positive symptoms</td>
<td>43 14.2 (5.4)</td>
<td>42 12.5 (5.0)</td>
</tr>
<tr>
<td>Self-esteem (SES)</td>
<td>43 17.3 (4.4)</td>
<td>42 16.7 (4.2)</td>
</tr>
<tr>
<td>Social behaviour (SBS)</td>
<td>43 11.6 (8.5)</td>
<td>42 13.7 (11.2)</td>
</tr>
</tbody>
</table>

BADS, Behavioural Assessment of the Dysexecutive Syndrome; CRT, cognitive remediation therapy; PANSS, Positive and Negative Syndrome Scale; SES, Self-Esteem Scale; WCST, Wisconsin Card Sorting Test; SBS, Social Behaviour Scale.

1. Maximum score 30.
Primary outcomes

- Change in social memory over the treatment period on social interaction scores.

Economic outcomes

- Societal costs: £46 in healthcare and social care costs.

DISCUSSION

The participants in this study had a wider range of abilities than participants in other studies and most other clinical information showed a wide spread of scores. In the overall comparisons these wider ranges tended to produce differences in the direction of this sample having poorer performance (e.g. on the WCST and Digit Span). In fact, although the sample was recruited as having fulfilled a number of different cognitive criteria, a large proportion (30% 6%) did not complete any categories post-treatment assessment disappeared at follow-up. However, there was no evidence of a direct effect of therapy on social behaviour scores.

Economic outcomes

Total overall health and societal costs are shown in Table 4. There is an advantage (although with highly skewed confidence limits) for the treatment period, with a difference of UK £1086 in healthcare costs and £1284 in societal costs in favour of the therapy group, but costs are higher at follow-up. The intervention dominates at follow-up. The intervention shows a significant effect for the therapy group (F1,34=9.2, P=0.006) but not for the control group (F1,34=0.06, P=0.16).

Effects of cognitive improvement

Since group could be shown to affect working memory, the effect of working memory change over the treatment period on social functioning outcome to follow-up was investigated using regression. There was no evidence of a group-dependent effect of cognition (F=0.996, d.f.=1.65, P=0.32), but after excluding the interaction term there was a significant effect of cognitive change (F=4.78, d.f.=1.66, P=0.03), suggesting that improvements in cognition were associated with improvements in social behaviour. When the effects were investigated in each group independently there was a significant effect for the therapy group (F1,34=9.2, P=0.006) but not for the control group (F1,34=0.06, P=0.16).

<table>
<thead>
<tr>
<th>Outcome measures</th>
<th>Interaction</th>
<th>Group effect (excluding non-significant interaction)</th>
<th>Estimated advantage to CRT† (95% CI)</th>
<th>Standardised effect size* (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary outcomes</td>
<td>Working memory (Digit Span)</td>
<td>F1,20=0.21, P=0.84</td>
<td>F0,8=5.82, P=0.019</td>
<td>1.33 (0.43 to 2.16)</td>
</tr>
<tr>
<td></td>
<td>Cognitive flexibility (WCST)</td>
<td>F0,8=6.924, P=0.011</td>
<td>NA</td>
<td>0.17 (−0.64 to 0.98)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td></td>
<td>1.017 (−1.8)</td>
</tr>
<tr>
<td></td>
<td>Planning (BADS)</td>
<td>F0,8=0.315, P=0.576</td>
<td>F0,8=0.092</td>
<td>1.1 (−0.18 to 2.4)</td>
</tr>
<tr>
<td>Secondary outcomes</td>
<td>Self-esteem (SES)</td>
<td>F0,8=4.44, P=0.039</td>
<td>NA</td>
<td>1.05 (−0.3 to 2.42)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td></td>
<td>−0.57 (−1.99 to 0.85)</td>
</tr>
<tr>
<td></td>
<td>Symptoms (PANSS)</td>
<td>F0,8=3.55, P=0.06</td>
<td>NA</td>
<td>−4.68 (−10.8 to 1.44)</td>
</tr>
<tr>
<td></td>
<td>Follow-up</td>
<td></td>
<td></td>
<td>1.59 (−4.8 to 8.04)</td>
</tr>
<tr>
<td>Social functioning</td>
<td>F0,8=0.438, P=0.53</td>
<td>F0,8=0.008, P=0.929</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 4** Baseline-adjusted mean costs and mean cost differences, including the cost of CRT

<table>
<thead>
<tr>
<th>CRT</th>
<th>Usual care</th>
<th>Mean difference† (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>Cost, Mean (s.d.)</td>
<td>n</td>
</tr>
<tr>
<td>CRT</td>
<td>Cognitive remediation therapy. 1. Adjusted for the baseline values of equivalent cost categories and baseline total Positive and Negative Syndrome Scale score.</td>
<td>41</td>
</tr>
</tbody>
</table>
of the WCST unlike participants in an earlier study (Wykes et al., 1999). The current group had also been in contact with psychiatric services for slightly longer and had poorer cognitive capacity than the previous study group (NART IQ 93 v. 104). Thus the sampling method achieved a wider variation in scores, with more people who had particularly poor cognitive performance on the outcome measures.

Despite the more chronic nature of impaired functioning in this sample, the results were similar to those of the previous pilot study. There was a durable improvement in working memory 6 months after the end of therapy, a significant improvement at follow-up in cognitive flexibility, and an advantage – but not a significant one – for planning (before medication was taken into account). Thus, for the primary cognitive outcomes there is evidence of overall effectiveness in a mixed group of participants, which in itself is an achievement as all the data point to a stable course for cognitive deficits. This fits into the growing evidence base showing that improvements in cognition are achievable even when the disorder has been evident for some time. In fact, nearly half of the people who scored at a very poor level on working memory performed within the normal range following therapy.

Following the end of therapy there was a continuing improvement in cognition, in cognitive flexibility. One possible explanation is that cognitive remediation therapy ‘jump starts’ engagement in the cognitive system through enhancing positive reward. This is achieved by the reinforcing nature of the tasks; the encouragement within therapy to engage these cognitive systems in everyday tasks; and the improved self-esteem and self-efficacy that further encourage engagement in new tasks, which provides continued practice.

The type of medication prescribed did not have any effect on the WCST or Digit Span outcomes. However, for the planning measure there was a significant group by medication interaction, suggesting that for some cognitive outcomes the type of medication could hinder or enhance the effects of cognitive remediation. The participants prescribed clozapine had a lower baseline planning score and therefore more chance for improvement. On the other hand, there was no difference in cognitive outcome between those taking typical and those taking atypical antipsychotics, suggesting that atypical medication does have a detrimental effect on the likelihood of change due to cognitive remediation.

In addition to cognitive outcomes, there were also some notable improvements in distal outcomes such as self-esteem post-treatment. These effects fit into the growing pattern (described by Twamley et al., 2003) of an effect on functioning of successful cognitive rehabilitation. There was an advantage in terms of both healthcare and societal costs for the therapy group at the post-treatment assessment, although this was not significant.

Methodological considerations
Data from other studies demonstrate high levels of stability of cognitive functioning without an intervention (Grég et al., 2004) and it might be proposed that these cognitive effects result only from attention and increased social contact. However, they are similar to those found in a previous trial in which there was an attention control group (Wykes et al., 1999, 2003), and although levels of social contact were higher in the therapy group there was no direct effect on social functioning outcomes. It therefore seems parsimonious to assume that the intervention therapy did produce the beneficial cognitive effects.

It is also vital to consider any possible challenges to the study validity, which for studies of psychological interventions in particular include rater bias. Although procedures were in place to reduce the chance of unmasking group allocation this did happen in some cases when cognitive data were collected. However, data quality checks (double scoring, data entry checking, etc.) were carried out each month to ensure that data were not obviously biased for any one rater or for any one participant. It was not possible to mask group allocation for the social functioning informants, but in this case there was no measurable effect of therapy, only interactions with cognitive improvements that were unknown to the key informants when they were providing the relevant information. It therefore seems likely that these acknowledged methodological difficulties did not compromise the study validity.

The sample size was small (although larger than any prior studies of this therapy) and so the power to investigate subtle effects was low. However, if this therapy is to be provided by health services it needs to show at least moderate effects and the current effect sizes are similar to those attributed to cognitive-behavioural therapy (Terrier & Wykes, 2004) suggesting that they may have some clinical relevance.

The effects of improving working memory
Poor memory has been highlighted in a number of studies as predicting poor overall outcome (Muser et al., 1991; Green et al., 2000). It was assumed that cognitive improvements would lead to functional change, and this is one of the reasons that cognitive remediation therapy was developed. However, few studies have measured the functional effects at a time when it might be possible for cognitive improvement to have had time to translate into functional changes. In this study there was support for a model in which a change in working memory had a beneficial effect on social behaviour 6 months after the end of therapy. This effect was only significant in the therapy group. The intervention was associated with lower costs at post-treatment assessment, and even when there were higher costs at follow-up these were small in relation to the beneficial outcome in working memory. The therapy itself has been estimated to cost £546.97 at 2000–2001 prices per patient (Wykes et al., 2003). The cost-effectiveness analysis has demonstrated that this translates into a small price to pay for memory improvements which are likely to produce further benefits on social behaviour.

The mechanisms through which cognitive change leads to functioning change have been somewhat elusive. Wykes & Reeder (2005) have suggested that for change to occur in routine behaviours, cognitive capacity of the sort measured by neuropsychological tests must change. This increases the efficiency with which routine cognitive schemas are implemented or held online. However, most functioning is not routine and must be flexible, so for novel behaviours it is necessary to change a further aspect of thinking – metacognition. Metacognitive processes and knowledge are required for the replacement of inefficient existing routine schemas and for the production of new, temporary high-level schemas. Current measures of cognition do not allow the separation of cognitive from metacognitive processes, and current measures of functioning also do not allow the differentiation of routine from novel actions. It is possible that the incomplete translation of working memory improvements into
social behaviour change is explained by the fact that it is not currently possible to differentiate the two forms of action. Working memory improvements are likely to translate into routine actions immediately, but not necessarily directly to non-routine actions unless metacognition has also improved. Future studies need to be able to differentiate behaviour change into routine efficiency and novel behaviour flexibility.

ACKNOWLEDGEMENTS

The study was supported by a grant from the Department of Health (RFG 757) who had no role in the study design, data analysis or interpretation, writing the report or decision to submit the paper for publication. The views expressed in this paper are those of the authors and not necessarily those of the funder.

REFERENCES


Target groups for the prevention of late-life anxiety

FILIP SMIT, HANNE COMIJS, ROBERT SCHOEVERS, PIM CUIJPERS,
DORLY DEEG and AARTJAN BEEKMAN

Background  Anxiety disorders in older people are highly prevalent, yet there is little evidence to guide targeted prevention strategies.

Aims  To identify subgroups at increased risk of developing anxiety in later life.

Method  Anxiety was measured with the Hospital Anxiety and Depression anxiety sub-scale in 1931 people aged 55–85 years followed over 3 years. Risk factors were identified that had a high combined attributable fraction, indicative of substantial health gains when the adverse effect of the risk factors can be contained.

Results  Factors significantly associated with increased risk of developing anxiety included sub-threshold anxiety, depression, two or more chronic illnesses, poor sense of mastery, poor self-rated health and low educational level.

Conclusions  The identified risk groups are small, thus providing prevention with a narrow focus, and health gains are likely to be more substantial than in groups not exposed to these risk factors. Nevertheless, more research is needed to produce evidence on target groups where prevention has optimal impacts.

Declaration of interest  None.

Anxiety disorders in later life are highly prevalent (Flint, 1994; Beekman et al, 1998; Jorm, 2000), compromise quality of life (De Beurs et al, 1999; M endlowicz & Stein, 2000), are associated with excess mortality (Van Hout et al, 2004) and generate substantial economic costs to society (Greenberg et al, 1999; Löthgren, 2004; Smit et al, 2006a). Prevention of anxiety might thus be a means of generating health gains in the population and reducing future costs. To maximise the impact of prevention strategies on patient outcomes and costs, evidence of target groups is needed (cf. Schoevers et al, 2006; Smit et al, 2006b). We report the results of an analysis of longitudinal epidemiological data to identify groups at increased risk of developing anxiety in later life who might benefit from targeted prevention strategies. This would help to set a rational agenda for preventive psychiatry.

METHOD

The analyses were based on data derived from the first two waves of the Longitudinal Aging Study Amsterdam (LASA). The sampling method and procedures of this study have been described elsewhere in detail (Beekman et al, 2002). At baseline a population-based sample was obtained comprising 3107 persons in the age group 55–85 years. Participants had given their informed consent and underwent face-to-face interviews at home. The random sample was stratified by age and gender. The older age strata and men were oversampled in anticipation of their higher attrition rates. After 3 years (mean = 1115 days, s.d. = 59) a total of 2164 (69.6%) participants were successfully re-interviewed and had complete data on their anxiety status. Loss to follow-up (n = 943) occurred mainly because the individuals were too ill or were no longer alive at the time of the first follow-up. Predictors of loss to follow-up were older age, male gender, lower educational level, functional limitations, chronic diseases and cognitive decline, but not anxiety status at baseline (Beekman et al, 2002). Corrective weights were used to account for the joint effect of intentional over sampling and attrition.

Measures

Anxiety  Anxiety was measured with the anxiety sub-scale of the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983). The HADS was constructed with the aim of avoiding overlap between symptoms of anxiety, depression and physical illness. Its anxiety sub-scale (HADS-A) consists of seven items, for example ‘Lately, worrying thoughts go through my mind’. Each answer is rated on a four-point scale, ranging from 0 (rarely or never) to 3 (mostly or always). The scale scores range from 0 to 21, with higher scores reflecting higher anxiety levels. The HADS-A has good psychometric properties (Myklethun et al, 2001). The scores were dichotomised at the cut-off score of ≥ 8 (Snaith, 2003). In this paper a HADS-A score equal to or greater than 8 is referred to as ‘anxiety’. Measurements were taken at baseline (t0) and at first follow-up (t1). Incident cases were identified when two criteria were met: absence of anxiety at t0 (HADS-A < 8) and presence of anxiety at t1 (HADS-A ≥ 8).

Risk indicators

It is appropriate to conduct indicated prevention, or early intervention, in people who have some symptoms of anxiety but who do not yet meet the diagnostic criteria of the full-blown disorder (Mrazek & Haggerty, 1994). Therefore, sub-threshold anxiety is a relevant risk indicator. Sub-threshold anxiety was defined as an HADS-A score above the population mean of 3 and below the cut-off of 8. Furthermore, it is appropriate to conduct selective prevention in people who are at a higher risk of anxiety because they are vulnerable and exposed to risk factors. Following the vulnerability-stress theory (Brown & Harris, 1978) and pertinent research (De Beurs et al, 2001; Schoevers et al, 2003, 2005), the following risk indicators were included.

Depressive symptoms. Depressive symptoms were ascertained with the Center for
Epidemiological Studies Depression scale (CES-D; Radloff, 1977). The CES-D consists of 20 items and its total score has a range between 0 and 60. Scores of 16 or over indicate clinically significant levels of depressive symptoms (Berkman et al., 1986). At this cut-off score sensitivity is 100% and specificity is 88% for DSM-IV Axis I depressive disorder (American Psychiatric Association, 1994) in the Dutch population older than 55 years (Beekman et al., 1997). In this paper CES-D scores of 16 or over are referred to as ‘depression’.

Chronic illness. Chronic illness refers to the most prevalent chronic physical disorders among older people, such as diabetes mellitus, chronic obstructive pulmonary disease, cardiovascular disease, arthritis and cancer (Kriegsman et al., 1996). The chronic illness variable was dichotomised as 0 (no illness or one illness) or 1 (two or more illnesses); because the majority of older people have at least one chronic illness, dichotomising at one illness would be unlikely to have much discriminatory or predictive power. It is worth noting that the physical disorders were reviewed in detail during the interview: symptoms were checked, and it was ascertained whether the participant was receiving medical attention for that particular physical disorder. In addition, the congruence between the self-reports and the medical files of the general practitioners was checked, and found satisfactory. Moreover, concordance between self-reports and general practitioners’ data did not depend on depression or anxiety status (Kriegsman et al., 1996).

Functional limitations. Functional limitations were measured with an adaptation of the Organisation for Economic Cooperation and Development (OECD) indicator for functional limitations (Van Sonsbeek, 1988); this variable was coded as 0 (none or one limitation) or 1 (two or more limitations).

Self-rated health. Answers to the question, ‘How do you rate your health?’ were coded as 1 (poor health, sometimes good/sometimes bad, fair) or 0 (good or excellent health).

Mastery. Low mastery was measured using the abbreviated (five-item) version of the (seven-item) Pearlin Mastery Scale (Pearlin & Schooler, 1978) and dichotomised at the median (1, score below the 50th percentile on the scale; 0, score above 50th percentile).

Other variables. The following socio-demographic variables were also included in the analyses: male gender (1, female; 0, male), old age (1, older than 75 years; 0, younger), low educational level (1, elementary or less; 0, more than elementary), living in an urban environment (1, living in Amsterdam; 0, living elsewhere) and small social network (1, fewer than 13 persons; 0, 13 or more persons).

It should be noted that all risk indicators were measured at t₀, thus well before the outcomes at t₁, and were dichotomised prior to the analysis, such that the index category (coded 1) was the assumed higher risk compared with the reference category (coded 0).

Analysis

All analyses took into account that the data were generated by a sampling design with intentional oversampling of the male and older age strata, and some amount of loss to follow-up. Therefore, the data were weighted such that the multivariate sample distribution over gender and age was exactly the same as in the general Dutch population in the age range of 55–85 years as reported by Statistics Netherlands (http://www.cbs.nl). In order to obtain correct 95% confidence intervals and probability values under weighting, all variance-related statistics were obtained with the help of the first-order Taylor series linearisation method as implemented in Stata version 9.0 for Windows. Weighted numbers are reported, rounded to the nearest integer, throughout the remainder of this paper. The subsequent analyses were carried out in several steps.

Analysis of incidence

Incidence was calculated in the cohort of the population at risk – that is, among those who were not categorised as HADS-A anxiety cases at baseline, and for whom the HADS-A anxiety status was available at follow-up after 3 years (n = 1931). The incidence rate was obtained with the help of a weighted Poisson model which was regressed on the HADS-A anxiety status at follow-up, while taking into account that not all participants had equal follow-up times.

Analysis of risks

The incidence rate ratio (IRR) helps to identify high-risk groups. For each risk indicator the IRR was obtained by regressing the outcome (1, incident case; 0, not an incident case) on the risk indicator in a weighted Poisson regression model, while adjusting for all other variables in the risk set. The IRR describes how much larger the incidence rate is in the exposed group relative to the incidence rate in the unexposed group, controlling for competing risks. Incidence rate ratio values larger than 1 signify an increased risk and values smaller than 1 indicate a lower risk in the exposed group.

For each of the risk indicators (or combinations thereof) exposure rates were calculated. The exposure rate gives the percentage of the population exposed to a risk indicator, or to a combination of risk indicators. Finally, the attributable fraction was calculated for risk indicators and combinations thereof. This indicates by how many percentage points the incidence of anxiety will be reduced when the adverse effect of the risk indicators is completely blocked (Miettinen, 1974; Rothman & Greenland, 1998). In other words, the attributable fraction puts an upper limit to the achievable health gain in the population when prevention is successful in containing the adverse effects of the risk indicators. A maximum likelihood estimate of attributable fraction was obtained with the aflogit-procedure in Stata for each of the risk profiles under a Poisson regression while adjusting for competing risks (Greenland & Drescher, 1993).

These statistics indicate the size of the group to be targeted (exposure rate), their risk (IRR) and the expected maximum number of preventable cases (attributable fraction). The last can also be used to quantify the economic benefits of avoiding the onset of new cases. Together, these indices of health gain and effort allow us to select high-risk groups for whom prevention is likely to be most cost-effective.

Identification of small, high-risk groups

Starting from the ‘long list’ of available risk indicators (see Table 1), a ‘short-list’ was compiled (see Table 2) using a conventional back-stepping procedure in a multivariate Poisson model. Only statistically significant risk indicators were retained in the model. There are two reasons to take this approach. First, the number of tests (in the
subsequent analysis) increases exponentially with the number of risk indicators, and extensive multiple testing would increase the likelihood of committing a type I error, i.e. incorrectly assuming that some associations are significant when in fact they are not. Second, extensive multiple testing would soon become very time-consuming and make the method less attractive for use.

The short-list of competitive risk indicators was then used as a starting point for generating risk profiles. Each risk profile contains at least one risk indicator and often a combination of risk indicators. For each risk profile the corresponding IRR, exposure rate and attributable fraction values were calculated. Therefore it is also possible to identify risk profiles that are associated with the best values for the IRR, exposure rate and attributable fraction overall.

For the selection of the ‘best’ risk profiles, we used the following criteria. First, we selected only risk profiles with an IRR of 5.00 or more - population segments with at least a five-fold risk of becoming anxiety cases. This was done for ethical reasons: we wanted to select only groups with seriously elevated risk levels. Second, we decided to target only population segments that formed 10% or less of the older population (i.e. where the exposure rate is 10% or less). This criterion was invoked in order to make future preventive interventions logistically and economically more feasible. When several risk profiles met these criteria, we opted for the risk profile associated with the highest attributable fraction value; that is, where we might expect the largest health gain. Here we need to point out that the criteria were arbitrary, and other thresholds could have been chosen; however, choosing other thresholds does not affect the principle of the methodology.

Systematic application of these criteria can be graphically depicted as tree-like structures (Lemon et al., 2003; see Figs 1 and 2). At the top of the tree we place the risk indicator which has the best starting value of IRR, exposure rate and attributable fraction. The risk indicator with the starting values is called the ‘parental’ node. ‘Child’ nodes can appear below the ‘parental’ node; in a ‘child’ node the ‘parental’ risk indicator is combined with the risk indicator of the ‘child’ node. At the level of the ‘child’ nodes the risk indicators are selected such that the IRR remains equal to or above 5.00 and the exposure rate drops below 10%. This process can be continued by adding more nodes to a branch. At the end of a branch one finds a ‘terminal’ node that satisfies the pre-set criteria (IRR ≥ 5.00 and exposure rate ≤ 10%). If there is a choice among several terminal nodes, then one selects the node associated with the highest attributable fraction value; that is, where the health gain at population level is more substantial.

These data-analytical strategies were pioneered by Smit et al. (2004) in the field of depressive disorder among people aged 18-65 years then improved and applied to late-life depression (Smit et al., 2006b) and cross-validated by using a different data-set and following a different analytical approach (Schoevers et al., 2006).

### RESULTS

#### Characteristics of the sample

The cohort at risk (n=1931) can be described as follows: 52.3% were female, 20.6% were older than 75 years, 37.0% had elementary education or less, 26.8% lived without a partner and 46.4% had a personal network of fewer than 13 people. In clinical terms the sample was characterised by presence of anxiety symptoms (32.0%), CES-D depression (6.9%), presence of two or more chronic illnesses (19.4%), two or more functional limitations (13.3%), poor self-rated health (31.4%) and a below-average sense of internal locus of control, i.e. low mastery (54.8%).

#### Incidence

In the cohort at risk (n=1931) the incidence rate was 1.82 new cases per 100 person-years (95% CI 1.51–2.19). Accordingly, if we were to follow 100 people at risk of developing anxiety over 1 year, we would be likely to observe 1.82 new cases. The incidence rate is higher in women (2.45, 95% CI 1.97–3.05) than in men (1.12, 95% CI 0.79–1.60).

#### Model with all risk indicators

Table 1 shows the exposure rate, incidence rate ratio, and the population attributable fraction for each of the risk indicators, after adjusting for the effects of all other risks in the model. In this multivariate model six risk indicators reached statistical significance for their respective IRRs. These were low education, sub-threshold anxiety, history of depression, presence of two or more chronic illnesses, low self-rated health and below-average levels of mastery. The attributable fraction of sub-threshold anxiety

<table>
<thead>
<tr>
<th>Risk indicator</th>
<th>Exposure rate, % (95%CI)</th>
<th>IRR (95%CI)</th>
<th>Attributable fraction, % (95%CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female gender</td>
<td>52.3 (50.0 to 54.6)</td>
<td>1.37 (1.07 to 1.65)*</td>
<td>18.4 (9.3 to 38.9)</td>
</tr>
<tr>
<td>Age &gt; 75 years</td>
<td>20.6 (18.9 to 22.2)</td>
<td>0.77 (0.53 to 1.13)</td>
<td>14.8 (5.7 to 36.3)</td>
</tr>
<tr>
<td>Elementary education only</td>
<td>37.0 (34.8 to 39.2)</td>
<td>1.63 (1.08 to 2.46)*</td>
<td>22.0 (4.8 to 37.0)</td>
</tr>
<tr>
<td>Urban environment</td>
<td>25.0 (23.1 to 27.0)</td>
<td>1.44 (0.97 to 2.13)</td>
<td>11.7 (4.5 to 23.2)</td>
</tr>
<tr>
<td>Network &lt; 13 people</td>
<td>46.4 (44.1 to 48.7)</td>
<td>1.27 (1.08 to 1.66)</td>
<td>9.5 (4.3 to 22.3)</td>
</tr>
<tr>
<td>Ever widowed</td>
<td>20.8 (18.0 to 22.5)</td>
<td>1.28 (1.09 to 1.42)</td>
<td>7.5 (4.0 to 17.7)</td>
</tr>
<tr>
<td>No current partner</td>
<td>26.8 (24.9 to 28.8)</td>
<td>0.86 (0.67 to 1.12)</td>
<td>6.3 (4.0 to 18.3)</td>
</tr>
<tr>
<td>Sub-threshold anxiety</td>
<td>32.0 (29.8 to 34.0)</td>
<td>4.11 (3.29 to 5.68)*</td>
<td>55.9 (39.4 to 67.0)*</td>
</tr>
<tr>
<td>Depression</td>
<td>6.9 (6.5 to 8.1)</td>
<td>1.72 (1.12 to 2.63)*</td>
<td>11.2 (0.0 to 29.5)*</td>
</tr>
<tr>
<td>Two or more chronic diseases</td>
<td>19.4 (17.7 to 21.2)</td>
<td>1.54 (1.04 to 2.30)*</td>
<td>13.7 (0.0 to 24.9)*</td>
</tr>
<tr>
<td>Two or more functional limitations</td>
<td>13.3 (11.9 to 14.8)</td>
<td>0.99 (0.60 to 1.65)</td>
<td>3.2 (3.0 to 5.4)*</td>
</tr>
<tr>
<td>Self-rated ill health</td>
<td>31.4 (29.3 to 33.5)</td>
<td>1.06 (1.08 to 2.52)*</td>
<td>23.4 (3.8 to 39.6)*</td>
</tr>
<tr>
<td>Low mastery</td>
<td>54.8 (52.5 to 57.1)</td>
<td>1.05 (1.06 to 2.58)*</td>
<td>30.3 (35.4 to 45.6)*</td>
</tr>
<tr>
<td>Total attributable fraction</td>
<td>87.9 (79.4 to 93.0)*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IRR, incidence rate ratio.  
1. Risk indicator at t.  
2. Obtained for all risk indicators with IRR >1.0.  
* P < 0.05.
is large, and indicates that 55.9% of new cases of anxiety can be prevented when all cases of sub-threshold anxiety can be identified and receive an adequate early intervention. It is worth noting that all the risk indicators account for 87.9% of future anxiety cases (‘total attributable fraction’ in Table 1). We will return to this point shortly.

Selecting a smaller set of risk indicators

In the next step we obtained a parsimonious multivariate model with fewer risk indicators (Table 2). This model is based on the smallest subset of statistically significant risk indicators (at P < 0.05). Five risk indicators were retained: sub-threshold anxiety, depression, self-reported poor health, low mastery and elementary education only. Using the five selected risk indicators, 82.8% of future cases of clinically relevant anxiety can be identified (‘total attributable fraction’ in Table 2). In the complete model with all risk indicators (Table 1) this percentage was only marginally higher. The implication is that the parsimonious model is nearly as good for predictive purposes as the one that contained all available risk indicators. It should be noted that we obtained nearly identical results for a parsimonious model in which the indicator ‘poor self-rated health’ was replaced by ‘presence of at least two chronic illnesses’, but then both variables are highly correlated (OR = 5.70; s.e. = 0.67; P < 0.001). For that reason we also included ‘presence of at least two chronic illnesses’ in the subsequent analyses.

Selecting ‘optimal’ risk profiles for indicated prevention

As is evident from Table 2, there is some benefit in selecting sub-threshold anxiety as a starting point for identifying the ‘best’ high-risk group for prevention. This group is certainly associated with a high risk; the drawback is that the corresponding group is large (32% of the population of older people) and it is difficult to see how prevention could be delivered to such a large population segment. Now a number of risk indicators can be added to the risk profile (Fig. 1). Adding depression offers a good solution: the IRR is still larger than 5, but the exposure rate has now dropped to 5.4%. Thus the combination of sub-threshold anxiety and depression can be seen as a risk profile that meets the pre-set criteria. Figure 1 also shows that adding ‘low mastery’ to ‘sub-threshold anxiety’ is a good step in building a risk profile, but the size of the corresponding target group is still too large, and a third risk indicator must be added. This results in four terminal nodes, all satisfying the pre-set criteria. Among these terminal nodes, it can be seen that joint exposure to ‘sub-threshold anxiety’, plus ‘low mastery’, plus ‘low self-rated health’ yields the best attributable fraction value, indicating a larger health gain at population level compared with the alternative risk profiles.

<table>
<thead>
<tr>
<th>Risk indicator</th>
<th>Exposure rate (%)</th>
<th>IRR (95%CI)</th>
<th>Attributable fraction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-threshold anxiety</td>
<td>32.0 (29.8-34.0)</td>
<td>4.10 (2.62-6.43)*</td>
<td>55.1 (0.41-0.68)*</td>
</tr>
<tr>
<td>Depression (CES-D score ≥ 16)</td>
<td>6.9 (5.8-8.1)</td>
<td>1.83 (1.24-2.73)*</td>
<td>12.1 (4.6-21.2)*</td>
</tr>
<tr>
<td>Self-rated ill health</td>
<td>31.4 (29.3-33.5)</td>
<td>1.93 (1.31-2.86)*</td>
<td>28.8 (11.4-42.7)*</td>
</tr>
<tr>
<td>Low mastery</td>
<td>54.8 (52.5-57.1)</td>
<td>1.70 (1.12-2.66)*</td>
<td>32.0 (5.4-51.1)*</td>
</tr>
<tr>
<td>Elementary education only</td>
<td>37.0 (34.8-39.2)</td>
<td>1.75 (1.21-2.53)*</td>
<td>24.2 (7.5-37.8)*</td>
</tr>
</tbody>
</table>

Parsimonious multivariate model of the risk indicators (Parsimonious multivariate model of the risk indicators (PP an anxiety can be identified (‘total attributable anxiety can be identified (‘total attributable only). Using the five selected risk indicators, only). Using the five selected risk indicators,

low mastery and elementary education low mastery and elementary education

indicators were retained: sub-threshold an-indicators were retained: sub-threshold an-

on the smallest subset of statistically signif-on the smallest subset of statistically signif-
nious multivariate model in which the indicator nious multivariate model in which the indicator

is nearly as good for predictive purposes as is nearly as good for predictive purposes as

of the corresponding target group is still of the corresponding target group is still

fraction' in Table 2). In the complete model fraction' in Table 2). In the complete model

cases of sub-threshold anxiety can be identi-cases of sub-threshold anxiety can be identi-
cases of anxiety can be prevented when allcases of anxiety can be prevented when all

Fig. 1 Selecting combinations of risk indicators where the incidence rate ratio is greater than 5 and the exposure rate is below 10% while maintaining the attributable fraction as high as possible, starting with the group of people with sub-threshold anxiety (i.e. indicated prevention).
impossible (or too complex) to identify sub-threshold cases for the purpose of prevention. Then one would like to conduct ‘selective prevention’ directed at people without symptoms but exposed to easily recognised risk indicators, for example risk indicators that are known to general practitioners, or can be retrieved from patient files. Ruling out ‘sub-threshold anxiety’ as a starting point, the next best candidate is ‘antecedent depression’ (Fig. 2). The corresponding population segment is not too large (exposure rate 6.9%), but the IRR falls below the pre-set criteria. The remaining risk indicators can then be added to the risk profile and the IRRs are increased to a level that meets the criteria. Most of the risk indicators in Fig. 2 are likely to be known by a general practitioner, whereas ‘mastery’ can be measured quickly with the help of a five-item scale and ‘self-rated health’ with only one question.

**DISCUSSION**

We wanted to identify population segments in whom prevention of late-life anxiety would stand the best chances of generating health gains at population level. This would help to guide research towards promising areas in preventive psychiatry. This is important because anxiety disorders are prevalent and diminish quality of life, but there is no empirically validated intervention for preventing onset of anxiety disorders in later life (Feldner et al., 2004).

**Main findings**

Our study shows that it is possible to use longitudinal epidemiological data to select risk indicators that warrant interest from the prevention perspective. These are risk indicators that are associated with a low exposure rate, representing small groups, high incidence rate ratios (IRR), indicating seriously elevated risk levels; and high population attributable fractions, indicating substantial health gains at population level. The methodology of identifying risk indicators for prevention is not new (Miettinen, 1974; Morgenstern & Bursic, 1982), but in the field of psychiatric epidemiology and prevention research it has rarely been applied. In this study, we applied it to late-life anxiety and came up with the following key findings.

First, the incidence of clinically relevant late-life anxiety is 1.82 new cases per 100 person-years, representing a substantial annual influx of new cases. Second, starting from a list of putative risk indicators, only a few were identified as interesting from the prevention perspective when the effects of the risk indicators were adjusted for competing risks. These are sub-threshold anxiety, depression, having a below-average sense of mastery, low self-rated health and having had only elementary education. It is worth noting that poor self-rated health and having two or more chronic illnesses are correlated variables that appear interchangeable. Third, the combined effect of being exposed to two, three or four selected risk indicators yields statistically significant and substantially interesting values on measures of potential health gain (IRR, attributable fraction) and effort (exposure rate). It is worth noting that the joint exposure to more risk indicators implies a smaller population segment. The intervention thus has a narrow focus, and the corresponding number of people who are the intended recipients of prevention becomes logistically manageable.

**Economic ramifications**

Once the costs of the disorder are known from a cost-of-illness study, then it is possible to combine the indices of effect and effort with the costs into an *ante hoc* cost-effectiveness analysis (Smit et al., 2004, 2006b). Here we will make the corresponding calculations for two hypothetical preventive scenarios: a ‘do nothing’ scenario, and a scenario in which people are targeted for prevention when they are depressed and have some anxiety symptoms.

In the ‘do nothing’ scenario (without any preventive intervention) one would see 18 200 new anxiety cases per 1 million people in a given year, because the incidence rate is 1.82 new anxiety cases per 100 person-years. A study carried out in the USA conservatively estimated that the direct medical per-patient costs of anxiety disorders were equivalent to £844 in UK currency. In a source population of 1 million people, the ‘do nothing’ scenario would thus entail a cost of £844 × 18 200 = £15 360 800 annually per 1 million source population. Now suppose that a preventive intervention is developed to contain the adverse effects of sub-threshold anxiety in people with depression. This intervention could be based, for example, on cognitive–behavioural therapy. To reduce intervention costs, it could be offered as self-help with minimal guidance. From Fig. 1 we now know that a completely successful intervention delivered to all people with depression and with sub-threshold anxiety (5.4% of the older population) would reduce the incidence of anxiety by 20.4%. In a hypothetical scenario in which 100% of the target group is reached and all receive a 100% effective intervention, then 3713 (20.4%) of the new cases would have been avoided. In a more realistic scenario of 60% coverage and a 30% success rate for the intervention (cf. Cuijpers et al., 2005), this would result in 3713 × 0.60 × 0.30 = 688 avoided onsets. Avoiding 688 onsets would thus save £844 × 688 = £580 700 per 1 million source population.

Clearly, the intervention would introduce costs of its own. We have calculated these as £285 per recipient of a preventive intervention of the type described above (Smit et al., 2006c). Again assuming a coverage rate of 60%, this would entail 3713 × 0.60 × 285 = £635 000. The averted costs (£580 700 per 1 million people) may not completely offset the costs of a preventive intervention (£635 000 per 1 million people).
people); nevertheless, the savings form a good starting point for cost-effective prevention of late-life anxiety. In short, we have a method at our disposal that could help to direct attention to high-risk groups in which preventive interventions are likely to become cost-effective. This is achieved at an early stage of the expensive and time-consuming cycle of development and evaluation of preventive interventions. Having said this, we need to add that ultimately the cost-effectiveness of a preventive intervention has to be established in a cost-effectiveness analysis alongside a randomised trial.

Strengths and limitations

Our findings have to be placed in the context of the strengths and limitations of this study. Its strengths are the use of population-based data; the prospective design, which enables the study of incidence and facilitates aetiological inference; and the measurement of exposures, which is not biased through post hoc rationalisation on the part of the participants because at $t_0$ they could not have any knowledge about their future health status at $t_1$. Furthermore, this study is among the first to show how a statistical technique can be applied to quantify potential health benefits and the effort required to generate these health gains. It thus supplies the sort of methodology which is of importance for setting a rational ‘research and development agenda’ for preventive psychiatry.

The limitations of this study consist in the not very detailed measurement of the exposures. We do not know for how long and how intensively the individuals were exposed. Moreover, the number of studied risk indicators is limited in that, for example, genetic and other biological risk indicators were not included. Another limitation is the measurement of anxiety with the HADS-A. This is not a diagnostic instrument. However, it has good psychometric properties (Miyklet al, 2001), and it may be valuable as a screening instrument, especially because anxiety disorders in older people are not well recognised.

Conceptually, it would be useful to distinguish between risk indicators that are amenable to change, such as anxiety and depressive symptoms, and those that are not. It should be noted that some risk indicators are not modifiable, such as chronic illness. However, their adverse psychological effects might be contained. Finally, there are risk indicators that are not modifiable and that have effects that cannot be brought under control through preventive interventions (such as gender); however, these risk indicators are valuable from the perspective of identifying groups at risk – which was the principal aim of this paper.

Currently there is no empirical evidence that prevention of anxiety can be successful in older people, but there are examples of effective prevention of anxiety in younger age groups (see Feldner et al, 2004) and in unipolar depression (Cuijpers et al, 2005). In this journal we have presented data on the effectiveness of preventing depression in adults (Willemsen et al, 2004) and on its cost-effectiveness (Smit et al, 2006c). We believe that developing and testing preventive interventions of anxiety disorders across the lifespan is an important and emerging research field, and this calls for a rational research agenda for the future, based on the data that we now have (cf. Smit et al, 2006b).

This study and related studies (Smit et al, 2004, 2006b; Schoevers et al, 2006) were conducted in an attempt to answer the question of whether it is possible to reduce the incidence of common, disabling and costly mental disorders in a cost-effective way. Our answers are only tentative and are best regarded as working hypotheses about directions where efforts to develop preventive interventions and to test these interventions in empirical cost-effectiveness studies are likely to stand the best chances of becoming fruitful. In a next step these hypotheses have to be tested in randomised prevention trials and cost-effectiveness studies. As yet, we are only beginning to see how prevention can be directed to high-risk groups such that the health gains are maximised, while the efforts and costs to generate these health gains are minimised.

Acknowledgements

This study is based on data collected in the context of the Longitudinal Aging Study Amsterdam, financed by The Netherlands Ministry of Public Health. The work done for the current paper was financed by The Netherlands Health Research Council (ZonMW) grant 13600008.

References


FILIP SMIT, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, and, Department of Clinical Psychology, Vrije Universiteit, Amsterdam; HANNE COMIJS, Department of Psychiatry, Free University Medical Centre, Amsterdam; ROBERT SCHOEVERS, Mentrum Mental Health Care, Amsterdam; PM CUIJPERS, Trimbos Institute, Utrecht, and, Department of Clinical Psychology, Free University, Amsterdam; DONLY DEEG, Department of Psychiatry, Free University Medical Centre, Amsterdam; AARTJAN BEEKMAN, Trimbos Institute, Utrecht, and, Department of Psychiatry, Free University Medical Centre, Amsterdam, The Netherlands

Correspondence: Dr Filip Smit, Trimbos Institute, Netherlands Institute of Mental Health and Addiction, PO Box 725, 3500 AS, Utrecht, The Netherlands, Tel: +31 30 2959254, fax: +31 30 2971111, email: FSmit@Trimbos.NL

(First received 9 February 2006, final revision 19 October 2006, accepted 7 November 2007)
All-cause mortality among recipients of electroconvulsive therapy

Register-based cohort study

TRINE MUNK-OLSEN, THOMAS MUNK LAURSEN, POUL VIDEBECH, PREBEN BO MORTENSEN and RABEN ROSENBERG

Background  Studies investigating mortality secondary to electroconvulsive therapy (ECT) are few.

Aims  To assess the risk of mortality from natural and unnatural causes among ECT recipients compared with other psychiatric in-patients over a 25-year period.

Method  Register-based cohort study of all in-patients admitted to a psychiatric hospital from 1976 to 2000. Cause-specific mortality was analysed using log-linear Poisson regression.

Results  There were 783 deceased in-patients who had received ECT compared with 5781 who had not. Patients who had received ECT had a lower overall mortality rate from natural causes (RR = 0.82, 95% CI 0.74–0.90) but a slightly higher suicide rate (RR = 1.20, 95% CI 0.99–1.47), especially within the first 7 days after the last ECT treatment (RR = 4.82, 95% CI 2.12–10.95).

Conclusions  Further investigation of the effect of ECT on physical health and the observed increased suicide rate immediately following treatment are needed, although the last finding is likely to result from selection bias.

Declaration of interest  None. Funding detailed in Acknowledgements.

METHOD

Data were obtained from different Danish registers and were linked using a unique personal identification number (the CPR number). All live-born children and new residents in Denmark are assigned a CPR number, and all information is registered using this number in all national registers (Pedersen et al, 2006).

The study population consisted of all patients admitted at least once to the Psychiatric Hospital, Aarhus, University Hospital of Aarhus during the study period April 1976 to October 2000. The sample did not include out-patients who only rarely receive ECT. Admissions at other hospitals were not included owing to a lack of information about the use of ECT. In total, 783 ECT recipients and 5781 other psychiatric in-patients died during the study period. Follow-up started from the first day of the first admission after 1 April 1976 and ended on 1 October 2000 or at the date of death, whichever came first.

Registers  Information about ECT at the Psychiatric Hospital in Aarhus was obtained from a local register created from written charts detailing every episode of ECT given at the hospital between March 1976 and August 2000. Details covered the treatment itself (seizure duration and treatment choice, i.e. bilateral or unilateral) and treatment dates. This register was previously used to describe predictors of first-time ECT recipients (Munk-Olsen et al, 2006), and documents that 95% of all treatments were unilateral, which is in line with the study by Stromgren (1973).

There are no private psychiatric in-patient facilities in Denmark, so the Danish Psychiatric Central Register contains information about all admissions to Danish psychiatric hospitals and psychiatric facilities since 1969 (Munk-Jorgensen & Mortensen, 1997). Diagnoses were entered as ICD–8 codes (World Health Organization, 1967) until the end of December 1993 and as ICD–10 codes (World Health Organization, 1992) from January 1994 onwards.

Information on causes and time of death was obtained from the National Register of Causes of Death (Juel & Hæweg-Larsen, 1999). If the time of death was the same day that the patient had been discharged, we categorised the patient as being admitted when analysing the variable 'days since discharge'.

Diagnoses  ICD–8 and ICD–10 diagnoses (primary diagnoses) were divided into six diagnostic groups: schizophrenia (ICD–8, 295 (minus 295.79); ICD–10, F20); schizoaffective disorders (ICD–8, 295.79, 269.89; ICD–10, F25); bipolar disorders (ICD–8, 296.19, 296.39; ICD–10, F30, 31); unipolar.
depressive disorders (ICD–8, 296 (minus 296.19, 296.39, 296.89), 298 (minus 298.39); ICD–10, F32–39); other non-affective psychosis (ICD–8, 297, 298.39, 301.83; ICD–10, F21–29 (minus F25); and 'other disorders' (remaining diagnoses).

Data analysis
Data were analysed using the log–linear Poisson regression with SAS GENMOD version 8.02 for Windows, with person-years as an offset variable (Laird & Ollier, 1981; Andersen et al, 1993). All variables except gender were treated as time-dependent. All relative risks (RR) were adjusted for gender, diagnosis, calendar period (5-year intervals) and age (10-year intervals). Wald’s test was used to calculate 95% confidence intervals (CIs); CIs not including 1.00 indicated a statistically significant difference (at a 5% level) from the reference group.

RESULTS

A total of 783 ECT patients died during the study period of around 25 years: 593 (76%) from natural causes and 190 (24%) from unnatural causes. The relative risk of mortality compared with the other psychiatric in-patients was 0.86 (95% CI 0.79–0.94). Among the 783 deceased ECT patients, 447 (57%) had been diagnosed with unipolar affective disorders, 178 (23%) with bipolar affective disorders, 55 (7%) with schizoaffective disorders, 46 (6%) with schizophrenia, 9 (1%) with other non-affective psychosis and 48 (6%) with 'other disorders'.

Mortality from natural causes
The mortality rate from natural causes was lower for ECT patients than for other psychiatric in-patients (RR=0.82, 95% CI 0.74–0.90, Table 1). Data were reanalysed adjusting for days since last ECT. The first 7 days after receiving the last ECT were characterised by an increased mortality risk (RR=2.11, 95% CI 0.94–4.74), but results were based on only 6 patients and CIs overlapped with 1.00 (indicating a non-significant finding at the 5% level).

Owing to the decrease in mortality from natural causes among ECT-treated patients (Table 1), we performed additional analyses including specific causes of death (Table 2) and found a decreased mortality rate for respiratory diseases (RR=0.73, 95% CI 0.55–0.95) and other natural causes (RR=0.67, 95% CI 0.56–0.82).

Mortality from unnatural causes
Overall, there was a slightly increased but statistically insignificant mortality rate from unnatural causes among ECT patients (RR=1.10, 95% CI 0.92–1.30, Table 2). A total of 41 ECT-treated patients had died from accidents but this did not translate into an increased risk compared with other in-patients (RR=0.80, 95% CI 0.56–1.14, Table 2).

A total of 149 ECT patients died by suicide during the study period, 112 (75%) of whom had unipolar or bipolar affective disorder. Patients treated with ECT had a marginally significant trend towards an increased risk of dying by suicide (RR=1.20, 95% CI 0.99–1.47) compared with other in-patients (Table 2).

Table 1

<table>
<thead>
<tr>
<th>Relative risk from natural causes among ECT patients and other psychiatric in-patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>ECT patients</td>
</tr>
<tr>
<td>Other psychiatric in-patients</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Cause of death</th>
<th>n</th>
<th>Relative risk (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unnatural causes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All unnatural causes</td>
<td>190</td>
<td>1.10 (0.92–1.30)</td>
</tr>
<tr>
<td>Suicide</td>
<td>149</td>
<td>1.20 (0.99–1.47)</td>
</tr>
<tr>
<td>Accidents</td>
<td>41</td>
<td>0.80 (0.56–1.14)</td>
</tr>
<tr>
<td>Natural causes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All natural causes</td>
<td>593</td>
<td>0.82 (0.74–0.90)</td>
</tr>
<tr>
<td>Cardiovascular diseases</td>
<td>163</td>
<td>0.85 (0.70–1.03)</td>
</tr>
<tr>
<td>Malignant neoplasms</td>
<td>124</td>
<td>1.01 (0.81–1.25)</td>
</tr>
<tr>
<td>Diseases of old age</td>
<td>90</td>
<td>0.98 (0.76–1.26)</td>
</tr>
<tr>
<td>Respiratory diseases</td>
<td>77</td>
<td>0.73 (0.55–0.95)</td>
</tr>
<tr>
<td>Other causes</td>
<td>139</td>
<td>0.67 (0.56–0.82)</td>
</tr>
</tbody>
</table>

ECT, electroconvulsive therapy.
1. Adjusted for gender, age (10-year intervals), calendar period (5-year intervals), diagnosis and previous admissions.
2. Adjusted for gender, age (10-year intervals), calendar period (5-year intervals), diagnosis, previous admissions, and days since last ECT treatment.
A 1.23, 1.01–1.52).

The increased suicide risk among ECT patients led us to examine the effect of the number of treatments. This varied from 3 to 17 among those who had died by suicide less than 7 days after discharge. Of the 6 ECT patients who died by suicide, 4 had received 10 or more treatments. The number of treatments received was not different from that of ECT patients who had died by suicide more than 1 week after discharge or ECT patients who had not died by suicide.

**DISCUSSION**

**Mortality from natural causes**

The overall mortality rate from natural causes was lower among ECT than other in-patients. Analysis of cause-specific mortality revealed that ECT patients had a lower mortality from respiratory diseases and the group of remaining disorders, and a marginally lower mortality from cardiovascular diseases than patients never treated with ECT. Other studies have found similar, reduced rates among ECT patients (Avery & Winokur, 1976; Prudic & Sackeim, 1999; Geddes, 2003), but these studies were based on fewer ECT patients and shorter study periods. The consistency of our findings with these studies strengthens our belief that the reduced mortality from natural causes is not an artefact.

The observed decreased risk of mortality from natural causes in ECT recipients could be a result of selection bias if patients with poor physical health are less likely to receive the treatment. However, in Denmark the only contraindications to ECT have been cerebral and other aneurysms. The Danish guidelines are in line with recommendations from the American Psychiatric Association that ‘no absolute medical contraindications to ECT exist’ (American Psychiatric Association, 2001). Furthermore, many medical conditions are contraindications for pharmacological treatments (Andersson et al., 2002), and we therefore do not believe that selection bias has influenced this result.

Patients treated with ECT are a selected group. First, ECT remains the first choice for treatment-resistant depression (Thase & Rush, 1995) and where a rapid, definitive response is required because of the severity of a psychiatric or medical condition (American Psychiatric Association, 2001). Patients treated with ECT have longer hospital stays than other patients (Munk-Olsen et al., 2006), which testifies to the severity of their illness. Second, ECT is used for treating patients with severe medical conditions for which anti-depressants cannot be used. However, these aspects do not explain the decrease in risk of mortality from natural causes but suggest that ECT patients should have a higher mortality owing to these selection mechanisms.

The decreased mortality from natural causes might result from the rapid effect of ECT on depression, which may be accompanied by a more quickly established, better overall well-being compared with treatment with pharmacological agents (Philibert, 1995; Prudic & Sackeim, 1999). The decrease in mortality among ECT patients might arise because these patients respond positively to this kind of therapy and are therefore more likely to receive ECT again upon recurrence of symptoms. The repeated use of this modality and its rapid efficacy could produce the overall reduction in mortality (Prudic & Sackeim, 1999). Another potential explanation might be that ECT patients are more likely to be admitted, which makes them less likely to become seriously ill without staff noticing and therefore more likely to receive quicker treatment and better care.

A single factor is unlikely to be the sole explanation for the decrease in mortality from natural causes among ECT-treated patients, and further studies focusing on different causes of death are therefore needed. Register data cannot be used to study whether ECT might have a positive effect on blood pressure, level of stress or other factors. The observed decrease in mortality from natural causes is unlikely to be a result of bias, but the effect of ECT on psychiatric patients’ physical health needs further exploration.

Mortality risk from natural causes was decreased for ECT patients compared with other psychiatric in-patients, but this does not mean that their mortality rates were decreased compared with people never admitted to a psychiatric hospital. Psychiatric patients are known to have increased mortality (from natural and unnatural causes) compared with the general population (Harris & Barraclough, 1998; Hoyer et al., 2000; Osby et al., 2001).

**Mortality from unnatural causes**

Overall, ECT patients had only a marginally significant increased suicide rate compared with other psychiatric in-patients. The
present study did not compare suicide or other mortality rates with those of the general population. However, our results are in line with other national studies of Danish patients (Hoyer et al., 2000, 2004) that have included some of the present regional sample. These results are in line with the general literature on suicide risk in people with depressive disorders (Harris & Barraclough, 1998).

Some studies have focused on ECT safety and have described mortality rates directly associated with ECT (Shiwach et al., 2001; Nuttall et al., 2004). Shiwach et al. (2001) reported that 30 of 8148 patients who received ECT during a 5-year period in Texas died within 14 days of ECT. Eight of these patients had died by suicide. The authors concluded that suicide might indicate failure of ECT. However, the severity of the psychiatric illness and the time required for the treatment to have a positive effect should be considered when assessing ECT as a preventative measure for suicide (Shiwach et al., 2001).

Our results should be interpreted with caution given the methodological limitations (i.e., the lack of information on variables associated with suicide) (Sharma, 2004). Our data came from registers and unfortunately no information was available on specific indications for receiving ECT, pharmacological treatments and dose, discontinuation of medication prior to a course of ECT, suicide attempts, patient well-being before and after the treatment and whether patients were on pass (temporary leave) during the admission.

A total of six ECT-treated patients died by suicide within 1 week of receiving treatment. These patients did not receive less treatment than the other ECT patients (mean 10.50 treatments per series v. 10.22 for all ECT patients, range 3–17). The elevated risk of suicide was not therefore a result of early treatment discontinuation. Psychiatric admission and discharge are associated with an increased risk of suicide (Hoyer et al., 2004), as found in this study. However, the elevated suicide rates in relation to time since last ECT treatment persisted even when admission status and time since discharge were taken into account. Hence our findings can not be ascribed to the well-established general association between suicide and psychiatric admission status.

Unipolar affective disorder is a strong predictor for receiving ECT (Munk-Olsen et al., 2006), and risk of suicide is particularly high among people with affective disorders (Mortensen et al., 2000). A total of 73 (49%) of the ECT patients who died by suicide during the study period had unipolar affective disorders. Since ECT is often administered to patients thought to be suicidal the risk might have been even higher if this group had not received ECT. Our results do not contradict the assertion that ECT might be under-utilised for suicidal patients (Nemeroff et al., 2001).

Others have reported a positive effect of ECT on short-term suicide rates, but it should be borne in mind that suicide risk has been a primary indication for the use of ECT since its introduction (Prudic & Sackeim, 1999), and therefore one would expect an elevated suicide rate among ECT patients. In the present study suicide risk was marginally increased among ECT recipients and was mainly, although not exclusively, confined to the first week after treatment. Therefore we could speculate that ECT returns the risk to the same (high) level of other patients with severe depression. However, another explanation is that almost a third of patients reported suicidal thoughts and acts prior to ECT in a study by Kelner et al. (2005), and if patients are assigned ECT because they are at risk of suicide this will introduce bias. This bias would be expected to be stronger close to the time at which the treatment was administered. We believe that this bias (confounding by indication) is a likely explanation for the moderately increased suicide risk among ECT patients in our study and the more marked increase shortly after treatment.

Suicidal intent in patients with depression is rapidly relieved by ECT (Kelner et al., 2005) and we cannot exclude the possibility that the risk during the first week after the last ECT session would have been higher if ECT had not been given to this particular subgroup of patients. This does not, however, exclude the possibility that elevated suicide rates immediately following ECT might be because ECT produces rapid improvement of depressive symptoms such as psychomotor retardation but does not eliminate suicidal impulses. This would indicate that patients should be assessed and monitored particularly carefully during this first period after relief of the depressive symptoms.

Implications

Compared with other psychiatric in-patients, ECT patients had an overall decreased mortality rate from natural causes but only a slightly increased suicide rate. A reduced mortality rate from natural causes among ECT patients has also been found in other studies, and there is a need for further investigation of the treatment and its potential effect on physical health. Although ECT patients are psychologically and physically severely ill, the decrease in mortality from natural causes implies that the treatment does not endanger but rather may have a positive effect on physical health.

The increased suicide rate among ECT patients shortly after treatment is probably a result of bias but our results also suggest that assessment and monitoring of suicide risk, and continuation treatment (e.g., antidepressants) continue to be important after the relief of depressive symptoms and ECT has been terminated.

ACKNOWLEDGEMENTS

We thank Dr. L. S. Stromgren who was the originator of the ECT-register and Dr. S. Bjerum for his help with data collection. This study was supported by the Stanley Medical Research Institute, The Danish Research Council (grant no. 22-02-0333) and the Health Insurance Foundation (Sygekassernes Helsefond, grant no. 11/212-95). Psychiatric epidemiological research at the National Centre for Register-Based Research is in part funded through a collaborative agreement with the Centre for Basic Psychiatric Research, Aarhus, Denmark.

REFERENCES


Dementia Screening Questionnaire for Individuals with Intellectual Disabilities

SHOUMITRO DEB, MONIKA HARE, LINDSAY PRIOR and SABYASACHI BHAUMIK

Background Many adults with Down's syndrome develop Alzheimer's dementia relatively early in their lives, but accurate clinical diagnosis remains difficult.

Aims To develop a user-friendly observer-rated dementia screening questionnaire with strong psychometric properties for adults with intellectual disabilities.

Method We used qualitative methods to gather information from carers of people with Down's syndrome about the symptoms of dementia. This provided the items for the Dementia Screening Questionnaire for Individuals with Intellectual Disabilities (DSQIID), which we then tested for its psychometric properties.

Results The DSQIID was administered to carers of 193 adults with Down's syndrome, 117 of whom were examined by clinicians who confirmed a diagnosis of dementia for 49 according to modified ICD-10 criteria. We established that a total score of 20 provides maximum sensitivity (0.92) and optimum specificity (0.97) for screening. The DSQIID has sound internal consistency (α = 0.91) for all its 53 items, and good test–retest and interrater reliability. We established a good construct validity by dividing the items into four factors.

Conclusions The DSQIID is a valid, reliable and user-friendly observer-rated questionnaire for screening for dementia among adults with Down's syndrome.

Declaration of interest None. Funding detailed in Acknowledgements.

Alzheimer's dementia is relatively common among adults with Down's syndrome and tends to manifest relatively early. As in the general population, increasing age and genetic predisposition act as risk factors (Ayward et al, 1997; Deb et al, 2000). Autopsy and neuroimaging studies (Deb et al, 1992) have shown an almost universal presence of Alzheimer's neuropathology among adults with Down's syndrome over the age of 45 years. Clinically, however, dementia is not universally manifested in this population (Mann, 1988; Prasher, 1995). One of the reasons for this discrepancy is the difficulty in diagnosing dementia among people with Down's syndrome in particular, and intellectual disabilities in general – especially during the early stage of dementia. Unfortunately, screening methods used for the detection of dementia among the general population are not suitable for people with intellectual disabilities because of floor effects. Moreover, we cannot standardise the cut-off thresholds for people with intellectual disabilities because those people vary considerably in their cognitive abilities. For the same reasons direct neuropsychological tests, including the Mini-Mental State Examination (MMSE; Folstein et al, 1975), are not useful for this population. Therefore, an observer-rated screening instrument which is primarily based on the reporting of behavioural changes following the onset of dementia is desirable (Deb & Braganza, 1999).

The two dementia screening instruments that are currently in wide use among people with intellectual disabilities, namely the Dementia Scale for Down Syndrome (DDSD; Gede, 1995) and the Dementia Questionnaire for Persons with Mental Retardation (DMR; Evenhuis, 1992, 1996), both have drawbacks. A questionnaire that is valid, reliable and easy to use could help to screen for dementia among people with Down's syndrome, which will help in timely diagnosis and treatment. We therefore developed a behavioural rating scale, incorporating carers' perspectives at the outset, for use by carers to screen for dementia in people with intellectual disabilities.

METHOD

Questionnaire development

We followed the steps described by Streiner & Norman (1999), which are widely accepted as gold standards for developing a new questionnaire. A qualitative interview method was used to inform the development of the questionnaire. Data gathered from interviews with carers of 24 adults with Down's syndrome and dementia were analysed qualitatively to derive 53 items for inclusion in the new questionnaire. The age of the 24 adults with Down's syndrome ranged from 48 to 72 years. Four people had mild, 16 moderate and 4 severe intellectual disabilities according to the ICD-10 criteria (World Health Organization, 1992). We named the questionnaire the Dementia Screening Questionnaire for Individuals with Intellectual Disabilities (DSQIID) because although the questionnaire was only tested among adults with Down's syndrome, we believe that it can be used equally effectively among all adults with intellectual disabilities. The project received approval from the Welsh Multi-centre Research Ethics Committee and we obtained written consent from each carer who completed the questionnaire.

The DSQIID is an observer-rated questionnaire, which is completed by carers of people with Down's syndrome who have known the individual for some time. The DSQIID is divided into three parts (see data supplement to online version of this paper). The first asks about the 'best' ability the person has or has had. The second contains 43 questions about behaviour or symptoms that are usually associated with dementia in adults with Down's syndrome. Each item is scored on a four-point scale: 'always has been the case'; 'always, but worse'; 'new symptoms'; and 'does not apply'. We adopted this scoring system to overcome the floor effect of the existing dementia screening scales, which only score current behaviour and not changes in behaviour (because in the general population the pre-existence of these skills is presumed). Items with a response of 'always been the case' or 'does not apply' are scored 0, those with 'always but worse' or 'new symptom' are scored 1.

Part 3 of the DSQIID contains 10 questions, all of which are comparative, for
example, 'speaks (signs) less' and 'seems generally more tired'. A response of 'yes' is scored 1 and a response of 'no' is scored 0. Scores from parts 2 and 3 are added to provide a total score. The 53 items of the DSQIID cover areas such as loss of memory, confusion, loss of skills, social withdrawal, behavioural changes, psychological symptoms, physical symptoms, sleep disturbance and speech abnormalities.

Questionnaire evaluation

Sample selection
Initially S.D. contacted colleagues in the UK requesting them to identify adults with Down's syndrome with and without dementia who might be suitable for inclusion in the study. M.H. publicised the study among her contacts who are primarily carers in Wales. S.B. approached those carers of adults with Down's syndrome on the Leicestershire register who had agreed to take part in research. The Leicestershire register holds information on over 3000 people with intellectual disabilities in the county. The adults with Down's syndrome who were included in the study had a range of intellectual disabilities.

Carers who expressed an interest were sent an information sheet, a written consent form, the DSQIID and a stamped addressed envelope. They were asked to return the completed DSQIID along with the completed consent form. We also asked the first carer to inform us of any other carer of the person with Down's syndrome who was willing to complete a DSQIID for that person – this was done to assess interrater reliability. Where appropriate, we immediately sent the same pack to the second carer, and thereby managed to gather data from 41 second carers of adults with Down's syndrome. We also sent the DSQIID again to the same carers immediately after we had received their completed first questionnaire. By this means we gathered test-retest data for 52 adults with Down's syndrome.

Inclusion criteria and matching
We did not match the groups with and without dementia for age and gender but subsequent analysis showed that those with dementia were significantly older than those without, which was expected. There was no significant difference in gender distribution between the groups. We did not match the two groups according to other possible confounders, such as hypothyroidism and depression, but on subsequent data analysis we did not find any significant intergroup differences in these variables (see Table 1).

Data analysis
We entered all data anonymously and analysed them using SPSS version 13 for Windows (Field, 2005).

RESULTS

Demographic data
We gathered data using the DSQIID on 193 adults with Down's syndrome from 28 centres in the UK. Local clinicians examined 117 of these adults and confirmed a diagnosis of dementia among 49 and the absence of dementia among 68 according to the modified ICD-10 criteria for the diagnosis of dementia among adults with intellectual disabilities (Aylward et al., 1997). Because some adults with Down's syndrome were recruited through carers and nursing staff, 76 were not examined by a clinician and therefore we do not have a dementia diagnosis for these participants.

We used receiver operating characteristic (ROC) analysis only on data from those who were examined by a clinician. We excluded 1 person with Down's syndrome from the ROC analysis because he had a mixed diagnosis of cerebrovascular events and dementia. We used data from all participants to analyse test-retest and interrater reliability.

The age of the whole cohort ranged from 23 to 77 years (mean 56 years, s.d. = 7.6); 51% were male. The age of the 49 adults with dementia ranged from 44 to 77 years (mean 56 years, s.d. = 7). The age range of 68 adults without dementia was 23–63 years (mean 44 years, s.d. = 10). Eighteen adults without dementia were over age 50. Independent-sample t-test showed that those with dementia were significantly older than those without (P < 0.001); 54% of those with dementia and 37% of those without were female. Although it was not possible to gather IQ scores from a cohort recruited from multiple centres, 35% had fluent speech, 37% could use short sentences, 15% speak a few words, 7% used sign language and 6% had no speech. Similarly, 13% lived totally independently, 6% lived independently but needed a lot of help, 35% were cared for by others and needed some help, and 46% were cared for by others and needed a lot of help for self-care. Therefore, it could be assumed that a proportion had severe and profound intellectual disabilities.

Comparative data for the adults with and without a diagnosis of dementia on the presence of depression, epilepsy, visual or hearing problems, and the use of anti-epileptics, antidepressants and thyroid hormones are presented in Table 1. A significantly higher proportion of adults with dementia had hearing (P = 0.014) and visual (P = 0.044) problems.

Psychometric properties
Feasibility
We sought comments from experts on the initial draft, and updated the questionnaire in the light of comments received. We piloted the draft questionnaire among six carers of adults with Down's syndrome and dementia to identify any practical difficulties before wider use in field-testing. Any ambiguity in the questions, difficulty in understanding wording and other practical issues related to the design of the DSQIID were rectified.

Content validity
We checked whether carers were consistently missing any particular item or providing the same answer. We also checked

<table>
<thead>
<tr>
<th>Variable</th>
<th>With dementia (n = 49)</th>
<th>Without dementia (n = 68)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thyroxine treatment</td>
<td>12 (24.5)</td>
<td>26 (40.6)</td>
</tr>
<tr>
<td>Depression</td>
<td>5 (11.4)</td>
<td>3 (5)</td>
</tr>
<tr>
<td>Antidepressant treatment</td>
<td>9 (18)</td>
<td>7 (11)</td>
</tr>
<tr>
<td>Epilepsy</td>
<td>13 (26.5)</td>
<td>6 (10.3)</td>
</tr>
<tr>
<td>Anti-epileptic treatment</td>
<td>17 (33.3)</td>
<td>9 (14.1)</td>
</tr>
<tr>
<td>Visual problems</td>
<td>23 (46.9)</td>
<td>18 (28.6)*</td>
</tr>
<tr>
<td>Hearing problems</td>
<td>19 (38.8)</td>
<td>12 (18)**</td>
</tr>
</tbody>
</table>

*P = 0.044; **P = 0.014
Table 2  Factor analysis of the 43 DSQIID items.

<table>
<thead>
<tr>
<th>Initial eigenvalues</th>
<th>Memory/confusion</th>
<th>Feelings of insecurity</th>
<th>Sleep problems</th>
<th>Behaviour problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Variance</td>
<td>41.17</td>
<td>3.0</td>
<td>2.12</td>
<td>2.16</td>
</tr>
<tr>
<td>01. Can't washy bathe without help</td>
<td>0.690</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03. Dresses inappropriately</td>
<td>0.579</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09. Can't find words</td>
<td>0.678</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Can't follow simple instructions</td>
<td>0.725</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Can't follow more than one instruction</td>
<td>0.733</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Stops in the middle of a task</td>
<td>0.596</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Can't read</td>
<td>0.433</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Can't write</td>
<td>0.538</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Confused at night</td>
<td>0.603</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Can't find way in familiar surroundings</td>
<td>0.564</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Loses track of time</td>
<td>0.640</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Not confident to walk over small cracks</td>
<td>0.473</td>
<td>0.406</td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. Can't recognise familiar persons</td>
<td>0.536</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Can't remember names of persons</td>
<td>0.688</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. Can't remember recent events</td>
<td>0.740</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32. Seems to go into own world</td>
<td>0.554</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37. Does not know what to do with objects</td>
<td>0.537</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43. Talks to self</td>
<td>0.541</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02. Can't dress without help</td>
<td>0.506</td>
<td>0.531</td>
<td></td>
<td></td>
</tr>
<tr>
<td>05. Needs help eating</td>
<td>0.990</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>06. Needs help using bathroom</td>
<td>0.577</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>07. Incontinence including accidents</td>
<td>0.412</td>
<td>0.420</td>
<td></td>
<td></td>
</tr>
<tr>
<td>08. Does not initiate conversation</td>
<td>0.537</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. Unsteady walk/loses balance</td>
<td>0.531</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. Can't walk unaided</td>
<td>0.472</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. Withdrawing from social activities</td>
<td>0.606</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. Withdrawing from persons</td>
<td>0.427</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31. Loss of interest in hobbies/activities</td>
<td>0.688</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38. Appears insecure</td>
<td>0.550</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39. Appears anxious or nervous</td>
<td>0.610</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40. Appears depressed</td>
<td>0.632</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04. Undresses inappropriately</td>
<td>0.492</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Changed sleep pattern</td>
<td>0.672</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Wakes at night</td>
<td>0.674</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Sleeps during the day</td>
<td>0.511</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Wanders at night</td>
<td>0.741</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Wanders</td>
<td>0.711</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42. Fits/ Epilepsy</td>
<td>0.508</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Obsessive or repetitive behaviour</td>
<td>0.737</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34. Hides or hoards objects</td>
<td>0.755</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35. Loses objects</td>
<td>0.544</td>
<td>0.581</td>
<td></td>
<td></td>
</tr>
<tr>
<td>36. Puts familiar things into wrong places</td>
<td>0.525</td>
<td>0.527</td>
<td></td>
<td></td>
</tr>
<tr>
<td>41. Shows aggression</td>
<td>0.715</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

DSQIID, Dementia Screening Questionnaire for Individuals with Intellectual Disabilities.

for possible floor or ceiling effects from the spread of overall scores from all carers. When preparing the questions we took into account interpretability, ambiguity, carers' reading level, avoidance of double-barrelled questions, jargon, value-laden words, positive and negative wording, and the length of items.

Construct validity
An initial principal component analysis using 'varimax rotation' created 13 factors (Field, 2005), which captured about 80% of the total variance. Subsequent scree plot analysis revealed that between four and five factors would be more appropriate but clinical grouping of items revealed that a four-factor structure was most appropriate for the DSQIID. Therefore, we carried out a forced four-factor analysis with four factors which included over 57% of the overall variance (Table 2). We excluded from the factor analysis the last 10 items of the DSQIID that were rated on a two-point scale as either 'yes' or 'no' as opposed to other items that were rated on a four-point scoring system (see data supplement to online version of this paper).

Factor 1 has most items involving symptoms of memory deficit and confusion; factor 2 includes primarily symptoms relating to frontal lobe dysfunction such as apathy and feelings of insecurity; factor 3 comprises primarily sleep and confusion-related items; and factor 4 symptoms associated with behavioural problems. Apart from some minor overlap, the factors appear to reflect different clinical symptoms. About 41% of the variance is owing to factor 1, whereas the remaining three factors contribute less than 17% of the variance. The first new variable contains the maximum amount of variation, whereas the remaining variables are orthogonal to the first and are independent of the first principal component. It is for this reason that the latter three factors contribute less to the variance as any common associations with first component items are ignored. This means that the items comprising factor 1, such as memory impairment and confusion, are only somewhat more important in screening for dementia in this population than the items in other factors.

Internal consistency
Cronbach's α for all 53 items in the DSQIID is 0.91.

Criterion-related validity
We assessed criterion-related validity by comparing the total score on the DSQIID with the clinicians' diagnosis of the presence or absence of dementia among 117
adults with Down’s syndrome. We used a ROC method to calculate the best fit between specificity and sensitivity. Out of 49 adults who had a clinical diagnosis of dementia, 4 scored less than 20 on the DSQIID. Out of 68 adults with Down’s syndrome who did not have a clinical diagnosis of dementia, 2 scored more than 20 on the DSQIID. Therefore, use of an overall score of 20 as a screening cut-off provided a specificity of 0.97, a sensitivity of 0.92, a positive likelihood ratio of 31 and a negative likelihood ratio of 0.08. Hence with a cut-off score of 20 a positive diagnosis of dementia is 31 times more likely in a person with dementia than in one without. Similarly, a negative diagnosis of dementia is 0.08 times more likely or 13 times less likely in a person with dementia than without. We therefore recommend 20 as the cut-off for the total score when using the DSQIID for screening for dementia among adults with Down’s syndrome. However, it is possible that there might be a different cut-off score for people with severe and profound intellectual disabilities whom we were unable to test separately. We therefore recommend the serial use of DSQIID over a period of time, particularly for people with severe and profound intellectual disabilities.

Reliability
The intraclass correlation for test–retest reliability (n=52) is 0.95, with a two-tailed level of significance of P<0.01 (>.80% power). The intraclass correlation for interrater reliability (n=41) is 0.9, and the two-tailed level of significance is P<0.01 (>80% power).

DISCUSSION
Development of DSQIID
Our approach in developing the DSQIID is somewhat unique in that, for the first time, we have adopted a ‘bottom up’ approach and incorporated the views of carers of adults with Down’s syndrome regarding the symptoms of dementia. The strategy has ensured that the DSQIID has good face and content validity. It also puts carers’ views to the forefront.

The DSQIID is easy to use, takes approximately 10–15 min to complete and can be completed either at home or in a clinic. The questions are simple and easy to understand, and the scoring system is simple and unambiguous. The screening cut-off is constant rather than variable, unlike the DSDS. The same cut-off score applies to adults with all levels of intellectual disabilities, unlike the DM R. We were not able to gather data on the level of intellectual disabilities among the whole study population but have included adults with all degrees of intellectual disabilities. Moreover, the original qualitative study used for the design of the DSQIID included adults with Down’s syndrome with mild, moderate and severe intellectual disabilities (Deb et al, 2007). We have found that adults with Down’s syndrome and severe intellectual disabilities showed a different manifestation of dementia in the early stage of the disease (primarily loss of skills) compared with those with mild-to-moderate intellectual disabilities (primarily memory deficit) (Deb et al, 2007). However, the end-stage symptoms of dementia are likely to be similar in both groups.

Other findings
Although we did not match those with and without dementia at the outset of the study, data analysis showed that, as expected, the dementia group was significantly older. The gender distribution of the two groups is similar. The two most common differential diagnoses of dementia are depression and hypothyroidism but there were no significant differences between the two groups on these variables. However, both hypothyroidism and depression may coexist with dementia and both are treatable conditions. There was a higher rate of hearing and visual problems among the group with dementia but the implication of these findings is not clear. Perhaps those with sensory deficits are more likely to develop dementia or this group might have been erroneously diagnosed because of their poor sensory skills.

We believe that for the DSQIID to be most effective the carers completing it should have known the person with intellectual disabilities for at least 6 months, and should have witnessed the change in behaviour since before the onset of dementia. Although we did not use this criterion for the field-testing, we believe that the carer should report only those behaviours that have existed for at least 6 months.

Limitations
It was not possible to test whether a different cut-off score for screening dementia should be applied for people with severe and profound intellectual disabilities. Moreover, sensitivity of the DSQIID can only be tested in a prospective study.

Inclusion of more adults with a clinical diagnosis of dementia might have improved the accuracy of the results.

Strengths
In order to avoid the floor effect, as can be seen with the M M SE, we have employed a scoring system by which only recent changes in behaviour are scored rather than all behaviours. This is a major strength of the DSQIID, which allows its use in a cross-sectional context. However, it is probably best to use the DSQIID at regular intervals over a period of time to identify the change in score. A further strength of the DSQIID relates to its robust psychometric properties, which existing scales often do not possess. Previous studies have included only a small number of people with dementia when validating scales (Evenhuis, 1992; Gedye, 1995), whereas in this study the number of people with dementia is much higher, and is very close to the 50 suggested by Steiner & Norman (1999). Moreover, the scores for test–retest and interrater reliability and internal consistency indicate that the DSQIID is very robust compared with existing scales.

ACKNOWLEDGEMENTS
We thank all the carers who took part in the study and the people with Down’s syndrome for whom they cared. We thank Dr Sayeed Hague for statistical advice, staff at the Leicestershire Register for People with Learning Disabilities and clinicians who...
helped with recruitment and examined people with Down’s syndrome for the diagnosis of dementia. The study was funded by the Bailey Thomas Charitable Trust.

REFERENCES


Foetal brain development in offspring of women with psychosis

MARY C. CLARKE, MARY CANNON, MATTHEW HOGG, MAUREEN N. MARKS, SUE CONROY, SUSAN J. PAWLBY, ANNE GREENOUGH and KYPROS NICOLAIDES

Summary  Cerebral ventricular enlargement and reduced cortical volume are correlates of chronic schizophrenia. We investigated whether genetic risk for psychosis is related to differences in foetal brain development as measured by prenatal ultrasonography. Routine foetal cerebral measures at 19–23 weeks of gestation were compared between the offspring of 35 women with a history of psychosis and 105 control women matched for gestational age. Overall, no significant differences were found between the high-risk and control groups. There was a non-significant trend in the adjusted analysis towards increased lateral ventricular width in the offspring of mothers with psychosis.

Declaration of interest  None.

Funding detailed in Acknowledgements.

Cerebral ventricular enlargement and reduced cortical volume are now well-replicated correlates of chronic schizophrenia (Wright et al 2000) and can be detected at the time of first presentation for treatment (Cahn et al, 2002). It is therefore likely that brain changes can be found much earlier in the disease process – perhaps even in foetal life. Pinpointing the time during development when abnormalities are first evident will have implications for investigating the causal pathways to later psychotic illness and identifying developmental genes that might be operating (Cannon & Clarke, 2005).

The aim of this study was to determine whether genetic risk for psychosis was related to observable differences in foetal brain development as measured by ultrasound. Our hypothesis was that the high-risk offspring would show increased lateral ventricular volume and decreased cerebral volume measurements compared with controls.

METHOD

Study design
This was a case-control study using foetal scan records from archives. Foetal scan data were retrieved for women with a prior diagnosis of psychotic disorder who had attended King’s College Hospital London for antenatal care between 1998 and 2002. From a search of referral records to the hospital perinatal psychiatry service over the same period, we identified 80 women with a history of psychotic disorder, including schizophrenia, schizoaffective disorder, bipolar disorder and post-partum psychosis. We were able to retrieve foetal scan data from the antenatal scan database for 35 of these women. The next three women, matched for gestational age, who were scanned after each index woman were taken as the control group (n=105).

Information on maternal age, height, weight, ethnicity and parity was obtained from the database for both groups.

This study was approved by the ethics committees of King’s College Hospital and the Institute of Psychiatry, London.

Foetal scanning
Routine second-trimester scanning for foetal anomalies was performed between 19 and 23 weeks of gestation at King’s College Hospital. The foetal scan report gives detailed measurements of the foetus and details of any abnormalities detected. Sonography measurements were made according to standardised procedures (Snijders & Nicolaides, 1994).

We extracted the following measurements of foetal cerebral growth from the scan report: biparietal diameter, head circumference, cisterna magna size, transverse cerebellar diameter and lateral ventricular width (the width of the posterior horn of the lateral ventricle measured at its widest point). Measures of overall foetal growth included femur length and abdominal circumference.

Statistical analysis
T-test and χ² test were used to compare foetal growth measures and maternal characteristics between the two groups. Conditional logistic regression analyses were performed to estimate the odds for being in the high-risk group for each unit change in foetal cerebral structures. Regression analyses were carried out using STATA version 8 for Windows.

RESULTS

Foetal growth
The two groups were well matched for mean gestational age at the time of the scan: 154 days (s.d.=9.5) for women with a history of psychotic illness v. 155 (s.d.=9.2) for controls. There were no significant differences in mean foetal femur length (37.9 mm, s.d.=4.8 v. 37.3 mm, s.d.=4.0; t=0.706; d.f.=135, P=0.48) or foetal abdominal circumference (173.2 mm, s.d.=0.5 v. 171 mm, s.d.=16.1; t=0.691, d.f.=135, P=0.49), indicating that there was no difference in overall foetal growth between the groups.

Maternal characteristics
Mothers with a history of psychiatric illness tended to be older (mean age 30.8 years, s.d.=5.8 v. 30.1, s.d.=6.1); heavier (mean weight 72.3 kg, s.d.=17.8 v. 67.5, s.d.=12.7); shorter (mean height 157.9 cm, s.d.=26.4 v. 165.3 cm, s.d.=9.6); and had greater parity (mean 1.2 children, s.d.=1.5 v. 0.75, s.d.=1.2). Only the difference in height was statistically significant (t=2.123, d.f.=103, P<0.05). There was no significant difference in ethnicity between the two groups (χ²=2.686, d.f.=2, P=0.26), but ethnic group was recorded for only three-quarters of the sample.

Table 1 presents odds ratios unadjusted and adjusted for maternal characteristics and measures of overall foetal growth. There were no significant differences between the groups for any of the foetal cerebral measurements in the unadjusted analysis. The adjusted analysis revealed a trend (P=0.06) for lateral ventricular size to be associated with genetic risk for psychosis. A unit increase in lateral ventricle width led to a 2.2-fold increase in the adjusted odds of being the offspring of a mother with a history of psychosis. Adjustment for parity differences between the groups contributed most strongly to this finding. On further examination, we found
a significant negative correlation between ventricle width and parity among women with a history of psychosis, such that each unit increase in parity led to a corresponding decrease in ventricle width. Ventricle width was greatest when mothers had had no previous pregnancies (7.3 mm). The comparison figure for the control group was 6.7 mm.

**DISCUSSION**

We found no significant overall differences in foetal cerebral measures between the group at high genetic risk for a psychotic disorder and the control group. However, the trend towards increased lateral ventricular width in the high-risk group is in keeping with our hypothesis and with studies indicating a relationship between increased foetal ventricular width and childhood neurodevelopmental disorders (Gilmore et al., 2001). The finding of increased ventricular width in first-born children of mothers with psychosis is also consistent with the increased risk of psychosis in first-born children (Kemppainen et al., 2001; Haukka et al., 2004). This, to our knowledge, is the first published report on foetal brain development in those at high risk for psychosis. A strength of our study is the high quality of the foetal ultrasound data. King’s College Hospital (which incorporates the Harris Birthright Centre) is a centre of excellence in foetal medicine in the UK with sonographers that are trained to a high level in standardised ultrasonography techniques (Snijders & Nicolaides, 1994).

There are a number of limitations of the study. First, the small number of cases reduced the power to find a significant association. Second, second-trimester scan data were missing for a large proportion of the high-risk group and since it is likely that the women who did not attend for their scan were more severely ill, this would have resulted in underestimating the differences between the groups. Third, we lacked information on maternal socio-economic status, smoking and pregnancy complications, which are likely to differ between the groups. Third, we lacked sufficient diagnostic information to allow us to carry out a subgroup analysis.

We recommend that future studies should be prospective and include detailed maternal diagnostic and demographic information.

**ACKNOWLEDGEMENTS**

M. Cannon is supported by NARSAD (2002 Grable Investigator Award), The Wellcome Trust and The Health Research Board (Ireland).

**REFERENCES**


**Table 1** The predictive value of foetal cerebral measures in identifying genetic risk for a psychotic disorder

<table>
<thead>
<tr>
<th>Intracranial measure (mm)</th>
<th>Mean (s.d.)</th>
<th>Unadjusted OR (95%CI)</th>
<th>P</th>
<th>Adjusted OR (95%CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cases</strong> (n = 35)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Biparietal diameter</td>
<td>54.43 (5.08)</td>
<td>0.94 (0.81^1.1)</td>
<td>0.4</td>
<td>0.86 (0.83^1.1)</td>
<td>0.7</td>
</tr>
<tr>
<td>Head circumference</td>
<td>195.54 (17.19)</td>
<td>0.99 (0.95^1.0)</td>
<td>0.9</td>
<td>0.99 (0.96^1.0)</td>
<td>0.9</td>
</tr>
<tr>
<td>Cisterna magna</td>
<td>5.70 (0.99)</td>
<td>1.09 (0.74^1.6)</td>
<td>0.6</td>
<td>1.12 (0.73^1.64)</td>
<td>0.6</td>
</tr>
<tr>
<td>Transverseboreal diameter</td>
<td>22.81 (2.40)</td>
<td>1.08 (0.82^1.4)</td>
<td>0.5</td>
<td>0.99 (0.65^1.48)</td>
<td>0.9</td>
</tr>
<tr>
<td>Ventricle width</td>
<td>6.98 (0.75)</td>
<td>1.22 (0.77^1.9)</td>
<td>0.4</td>
<td>2.2 (0.9^5.1)</td>
<td>0.06</td>
</tr>
</tbody>
</table>

1. Adjusted for maternal age, weight, height, parity and foetal growth (femur length and abdominal circumference).
2. Not adjusted for ethnicity owing to missing data.
Zuclopenthixol in adults with intellectual disabilities and aggressive behaviours

Discontinuation study

FRANK HAESSLER, THOMAS GLASER, MANFRED BENEKE, AKOS F. PAP, RALF BODENSCHATZ and OLAF REIS
on behalf of the Zuclopenthixol Disruptive Behaviour Study Group

Summary We investigated the effects of zuclopenthixol on aggressive behaviour in patients with intellectual disabilities by randomly withdrawing it after a 6-week period of open treatment. Of the 49 patients responding to the treatment, 39 took part in a randomised withdrawal trial. The placebo subgroup (n=20) showed more aggressive behaviour as indicated by outcomes observed by external raters on the Modified Overt Aggression Scale than did the continuing subgroup (n=19). The results indicate that discontinuation of zuclopenthixol in this population leads to an increase in aggressive behaviour.

Declaration of interest T.G., M.B. and A.F.P. are employees of Bayer Vital GmbH.

People with intellectual disabilities are at higher risk of mental health problems compared with people from the general population. Particularly, people with intellectual disabilities show serious behavioural disturbances, such as disruptive and aggressive behaviour. Among institutionalised individuals with profound intellectual disabilities, the incidence of self-injurious or aggressive behaviours ranges between 30 and 60% (Baumeister et al, 1998; Mikhail & King, 2001). Recent controlled studies of antipsychotic drugs focusing on risperidone reveal valuable effects on aggression and self-injurious behaviour in individuals with intellectual disabilities (Aman et al, 2002). However, risperidone produces adverse effects and is more expensive than conventional antipsychotic drugs which are rarely studied (Baumeister et al, 1998).

To our knowledge, the study reported here is the first multicentre, double-blind placebo-controlled trial of zuclopenthixol over the past 10 years involving adult patients with intellectual disabilities displaying severe aggressive behaviour.

METHOD

A randomised, double-blind placebo-controlled withdrawal study for parallel groups was conducted in six German centres. Forty-nine people aged 18-50 years, with mild to moderate intellectual disabilities (IQ 30-70), received open treatment with zuclopenthixol for 6 weeks because of exacerbations of aggressive behaviour. Zuclopenthixol was administered at a dosage of 2-20 mg per day. The dosage was adjusted once or twice daily as judged necessary by the clinician. Eligible participants were mostly individuals in institutional settings who had complex behavioural problems as rated on the Disability Assessment Schedule (Holmes et al, 1982). All participants scored below 39 on this instrument. After complete description of the study to the participants and their legal representatives, voluntary written informed assent or consent was obtained from the participants or their legal guardians (or both) for participation in the investigation.

After open treatment, those in the responders group (n=39) were randomised to continue or discontinue treatment for up to 12 weeks. Participants who discontinued treatment received placebo medication. Individual dosages were kept as stable as possible during the randomisation period. Concomitant use of other antipsychotics was not permitted throughout the study. Use of consistent doses of anticonvulsants as well as lithium, medication for extrapyramidal symptoms and benzodiazepines as an anti-epileptic escape medication was permitted. All concomitant medications were recorded. For all patients the Modified Overt Aggression Scale (MOAS; Yudofsky et al, 1986) was administered every 2 weeks. Several secondary measures, medical history and safety measures, including possible withdrawal symptoms, extrapyramidal signs, vital signs and weight, were recorded. Routine laboratory tests of prolactin and serum levels of zuclopenthixol were conducted.

The primary efficacy measures were binary variables derived from weighted sums of the MOAS aggression sub-scores. The weighting of these scores gives a higher impact on severe (physical) forms of aggression (Kay et al, 1988). Patients with a deterioration of at least 3 points in MOAS sum scores at two subsequent visits when compared with their state at randomisation were designated as non-responders. All patients without deterioration were considered to be responders unless they withdrew from the study because of insufficient efficacy, concomitant treatment or adverse events.

Exclusion criteria were the presence of a diagnosed neurological disorder (without epilepsy), psychotic disorder, infantile cerebral palsy, hypersensitivity to zuclopenthixol and cardiac abnormalities. Female participants who were sexually active and did not use an effective form of birth control were also excluded.

RESULTS

Results are reported here for the intention-to-treat sample only. The proportion of participants rated as responders, based on the weighted sum of MOAS scores 12 weeks after randomisation, was statistically significantly larger in the zuclopenthixol group (37%, n=7) than in the placebo group (5%, n=1); difference 32% (95% CI 3-61), Fisher's exact test P=0.020.

Figure 1 shows the Kaplan-Meier estimates of responder rates for the placebo group and for the zuclopenthixol group.
log-rank test, *P* = 0.005. Per protocol analysis yielded similar results.

Psychotropic adjunctive medications given after randomisation (*n* = 7) were equally distributed between the groups and involved the prescription of one benzodiazepine drug in each group. The number of adverse events and possible symptoms of withdrawal, such as nausea, insomnia, and diarrhoea, were recorded and did not differ between the groups.

**DISCUSSION**

These results are in agreement with the studies of Singh & Owino (1992), who found zuclopenthixol to be more effective than placebo, and Malt *et al.* (1995), who found zuclopenthixol to be superior to haloperidol in reducing unwanted behaviours. However, it should be noted that we used a discontinuation design in this study, and it was the withdrawal of zuclopenthixol that caused an increase in aggressive behaviour. In our study the beneficial effects of zuclopenthixol were found at low dosages between 6 and 18 mg (mean 11.4 mg). These dosages were lower than those in other studies in adults with intellectual disabilities and associated behavioural problems (Singh & Owino, 1992; Malt *et al.*, 1995). It is possible that these lower dosages might be responsible for the relatively high relapse rates in the continuation subgroup.

The anti-aggressive effects of zuclopenthixol may be explained by its dopaminergic mechanism, especially its high affinity to dopamine D1 receptors (Singh & Owino, 1992). Its high selectivity, together with the low dosages, may also explain the low rate of adverse effects. The psychopharmacological mechanism of zuclopenthixol differs slightly from the dopaminergic-serotonergic impact of risperidone; nevertheless, it provides a cost-effective alternative to the use of this atypical antipsychotic drug. Zuclopenthixol may be indicated especially in institutional settings, where patients and staff have to cope with severe forms of self-injurious and aggressive behaviours.

**ACKNOWLEDGEMENTS**

The study medication and placebos were provided by Bayer Vital GmbH; D 162, 51368 Leverkusen, Germany.

**REFERENCES**


Correspondence
EDITED BY KIRIAKOS XENITIDIS and COLIN CAMPBELL

Contents ■ Smoke-free mental health units ■ X-chromosome abnormality and schizophrenia

Smoke-free mental health units
Jochelson (2006) highlights the very important challenges that mental health units in the UK are likely to face in becoming smoke-free environments. Although there is very little doubt about the benefits of protecting patients and staff from the direct and indirect effects of smoking, the crude application of regulations of the English Health Act 2006 to all psychiatric settings might not be entirely beneficial and some patients might need to be exempt. Individuals presenting with severe psychopathology, those lacking capacity to agree to nicotine replacement treatment and individuals admitted under the Mental Health Act 1983 who have reduced civil liberties and limited access to outdoor space raise considerable concerns. Under these circumstances a forced nicotine withdrawal is likely. This iatrogenic phenomenon is associated with significant risks such as severe exacerbation or misinterpretation of psychiatric symptoms (Greenman & McClellan, 1991; Dalak & Meador-Woodruff, 1996), and pharmacokinetic changes resulting in increased concentration of psychotropic medications (Hughes, 1993).

Jochelson minimises concern that under these circumstances there might be an increased risk of aggressive behaviour in psychiatric patients. The reality is that it is very difficult to be certain because the literature offers controversial findings. In older studies, which report negative results, the information is mostly retrospective and qualitative, and studies have adopted different outcome measures and failed to control for a number of fundamental variables such as access to the outside, which may vary according to staff availability and patient status (e.g. under the Mental Health Act 1983), hospital setting (in-patients, out-patients, intensive care units, etc.), psychiatric diagnosis, degree of psychopathology, level of dependence, comorbidity with other addictive behaviours, motivation, etc. (For review see El-Guebaly et al, 2002.) This has resulted in the limited generalisability of the findings. More recent studies have controlled for these variables and have reported increased irritability and agitation among psychiatric patients, with disengagement from services and premature discharge (e.g. Prochaska et al, 2004). It is also noteworthy, if the ban is intended to enhance the long-term health of psychiatric patients, that experience emerging from other countries where smoking bans in psychiatric hospitals have already been implemented suggests that resolution of smoking after discharge is the most likely outcome, with questionable long-term effects (El-Guebaly et al, 2002; Lawn & Pols, 2005; Prochaska et al, 2006).

Effective measures to increase the chance of positive health benefits could be based on evidence emerging from the treatment of nicotine addiction in hospitalised patients. An effective strategy includes diagnosis and treatment planning with nicotine replacement therapy or bupropion, on-unit dedicated smoking cessation counselling, reasonably extensive behavioural support, and post-discharge referral for treatment of nicotine dependence (West, 2002). Eliminating the burden of tobacco use in psychiatric hospitals is a public health priority but must be delivered in such a way that risks are minimised in otherwise vulnerable individuals and healthcare systems are developed that are capable of delivering effective treatments.


D. Aronne, Department of Addictive Behaviour, St George’s Hospital Medical School, London and Department of Psychiatry, Warneford Hospital, Oxford, UK. Email: danielle.aronne@psych.ox.ac.uk

L. J. Simmons, St Mary’s Hospital, Portsmouth Hospitals NHS Trust, Portsmouth, UK.

doi: 10.1192/bjp.190.5.449

Jochelson (2006) has described the issues that arise for mental health units in England and Wales as a result of the Health Act 2006 which will ban smoking in public places. The proposed regulations will require most mental health units to ensure that the wards and the communal areas are smoke free. However, Jochelson does not consider the challenge to the implementation of the regulations presented by patients detained under the Mental Health Act 1983. These patients are detained in hospital against their will and are very likely receiving treatment to which they have not consented. Not only will they be deprived of their liberty but, if they are smokers, may also be forced to stop smoking. To compel a patient to stop smoking is unlikely to be a lawful use of the powers of the Mental Health Act 1983. To enforce a ban on smoking could be found to be an unjustifiable interference with the patient’s human rights, if subjected to a legal challenge (Mental Health Act Commission, 2006a).

Patients may be allowed to smoke outside the building, but for some patients on some units this may not be possible because of the risk posed to themselves or others. The regulations will allow units that normally provide accommodation for more than 6 months to have a designated smoking room. However, figures from a national census of mental health hospitals in England and Wales in March 2006 suggest
X-chromosome abnormality and schizophrenia

Van Rijn et al (2006) concluded that their findings suggested a link between an X-chromosomal abnormality and liability to schizophrenia which might be useful in the search for the genetic aetiology. Moreover they stated that a crucial role for X-chromosome abnormalities in this context has been proposed by Lishman (1998). In 1966 Hambert described a group of 75 XXY men, of whom 17 had hallucinations, 21 paranoid ideas, 9 "melamagical ideas" and 5 "short periods of mania-like disorder". Penrose (1966) claimed that "the effects of sex chromosomal disorders are more noticeable in relation to alterations in character and stability than to intellectual loss. Olander's (1975), working in the same research group as Hambert and Penrose, reported 16 women with schizophrenia among 31 with triple-X syndrome. Olander's own psychiatric investigation of these women revealed many with paranoid symptoms who did not meet his strict criteria for schizophrenia. He described a paranoid syndrome in 8, hallucinations in 4, confusion in 3 and catatonic symptoms in 2.

Van Rijn et al discuss the relationship between an extra X chromosome and psychosis through the 'decreased cerebral lateralisation' hypothesis. Netley & Rovet (1982) reviewed data which point to diminished cerebral cell numbers owing to lower mitotic rates which also result in the lower dermal ridge counts. I think that a lower cerebral cell number could give rise to decreased cerebral lateralisation, but this needs further investigation.

In my opinion, Van Rijn et al present no new data but have rediscovered what was known for a long time.


M. Otter Bennekom, The Netherlands. Email: maarten.otter@fitternet.nl
doi: 10.1192/bjp.190.5.450a

Authors' reply: Otter claims that our finding of high levels of schizophrenia symptoms in XXY men is a rediscovery of what has been known for a long time. He supports his claim by referring to reports on triple-X syndrome that were not published in peer-reviewed journals from the University of Gothenburg. We acknowledge that previous studies have also reported psychopathology in XXY men. However, we also point out that these studies have been limited in that they described men with Klinefelter's syndrome in psychiatric care or recorded hospital admissions. Our findings in a non-selected sample of XXY men, using valid and reliable dimensional measures of psychopathology, corroborate and extend the data derived from these earlier studies.

With regard to the novelty of the findings, it is interesting to note that none of the major reviews on Klinefelter's syndrome (Smyth & Bremner, 1998; Lanfranco et al, 2004) report a vulnerability for schizophrenia psychopathology, indicating that this is not a generally accepted feature. In addition, the aim of our study was not to provide a comprehensive review of psychopathology in X-chromosomal disorders, but we find the presence of schizophrenia psychopathology in XXY females very interesting as it supports our suggestion of a link between the X chromosome and schizophrenia symptoms.

Finally, Otter argues that reduced cerebral lateralisation in Klinefelter's syndrome has been suggested by neurobiological studies but is yet to be proved. However, a recent functional neuroimaging study has also presented evidence for reduced lateralisation in brain perfusion in XXY men (Itti et al, 2003).

In conclusion, we feel that the evidence put forward by Otter merely underscores the importance of our findings, as both triple-X and Klinefelter's syndrome have been associated with schizophrenia symptoms. Including both syndromes in genetic studies would advance the understanding of a link between the X chromosome and schizophrenia pathology.


S. Van Rijn Department of Psychiatry, Rudolf Magnus Institute for Neuroscience, University Medical Center, Utrecht, and Psychological Laboratory, Helmholtz Institute, Utrecht University, The Netherlands. Email: S_vanRijn@iss.uu.nl

A. Aleman Psychological Laboratory, Helmholtz Institute, Utrecht University and BSN Neuroimaging Centre, University of Groningen, The Netherlands

H. Swaab Department of Psychiatry, Rudolf Magnus Institute for Neuroscience, University Medical Centre, Utrecht and Department of Clinical Child and Adolescent Studies, Leiden University, The Netherlands

R. S. Kohn Department of Psychiatry, Rudolf Magnus Institute for Neuroscience, University Medical Centre Utrecht, The Netherlands
doi: 10.1192/bjp.190.5.450a
One hundred years ago

Sammlung Kleiner Schriften zur Neurosenlehre (Collection of Short Studies on the Neuroses).
Pp. 234, 8vo.

Professor Freud’s elaborate and painstaking efforts to elucidate the mechanism of hysteria and of various allied neurotic conditions, notwithstanding the opposition with which they have sometimes been met, have aroused a growing interest, and he has been induced to bring together the various studies (three of them in French) which he has published on this subject during the past fourteen years. The volume forms an admirable introduction to Freud’s work. We are enabled to follow the course of his thought — which has constantly undergone fresh modifications in various directions — from the perusal of all who desire to become acquainted with the work of one of the most original investigators in the Etiology of the Neuroses.” Charcot would have said it played no part, for he believed that the whole of the etiology was covered by heredity; but Freud is more than ever convinced that this is not the case, and he endeavours to set down as clearly as possible where he considers that sexuality intervenes as a factor, and in what direction his views have been modified by experience. Hysteria he still regards as “the expression of a special relationship of the individual’s sexual function,” and he still believes that childish experiences have an influence over the later direction of the patient’s hysterical state, but he no longer speaks of early sexual experiences as “traumatic,” and he recognizes the part herein played by hysterical imagination; “infantile sexual trauma” gives place to “infantilism of sexuality.” Along this line he has been brought somewhat nearer to Charcot’s position and attaches primary importance to heredity and constitution, adding, however, that he thinks more of “sexual constitution” than of general neuropathic disposition. Masturbation he regards as a main cause of neurasthenia, and coitus interruptus as producing neurosis of anxiety.

This view certainly requires for its justification the emphasis on heredity, for minor sexual aberrations are far too common to be regarded as injurious to a constitution that is not aboriginally unsound.

While these studies are mainly concerned with neurasthenia and allied states, in one interesting passage (pp. 124 et seq.) Freud suggests that in some cases paranoia resembles hysteria and imperative ideas in that its symptoms may be determined by the suppression of painful memories of a sexual character dating from early life. A case is brought forward in which this could be clearly shown. This idea is suggestive, and it is probable that many readers, recalling cases of systematised delusion with which they were intimately acquainted, may bring to mind instances in which an ancient episode of sexual nature which the patient had, so far as possible, pushed out of consciousness, serves to form part of the basis of the later auditory hallucinations.

In an essay on psychotherapy the author discusses the origin and development of his method in its therapeutic aspects. The method appears to have been originally due to Breuer, who called it the “cathartic” method; Freud prefers to call it the “analytic” method. It is entirely distinct from hypnotism (which Freud has abandoned for over eight years), being indeed the exact opposite of hypnotism. By the hypnotic method it is sought to put something into the patient; by the cathartic or analytic method it is sought to take something out of him. Freud illustrates the difference by reference to Leonardo da Vinci’s technical distinction between the different ways of art, the via di porre, or the painter’s way, of putting in something that before was not there, and the via di levare, or the sculptor’s way, of removing something that is there.

Freud’s style is always clear, attractive, and sincere, and this book is well worth the perusal of all who desire to become acquainted with the work of one of the subtlest and most original investigators in a difficult field.

Havelock Ellis

REFERENCE

Journal of Mental Science, January 1907, 172–173.

Researched by Henry Rollin, Emeritus Consultant Psychiatrist, Horton Hospital, Epsom, Surrey
doi: 10.1192/bjp.190.5.451
Book reviews

EDITED BY SIDNEY CROWN, FEMI OYEBODE and ROSALIND RAMSAY

Social Inclusion of People With Mental Illness

Julian Leff and Richard Warner are two of the key voices in contemporary social psychiatry. This comprehensive (though relatively short) book is written accessibly, so will be a valuable introduction to anyone involved with severe mental illness, but has a wealth of detail for the specialist reader. The narrative is international (albeit with an understandable emphasis on the USA and UK), building a coherent narrative from research – much of which is, by necessity, on schizophrenia. Embedded in this are supporting and illustrative comments by service users, descriptions of fascinating practice innovations (from a consumer-run pharmacy in Boulder, USA to crisis homes in Aro, Nigeria) and invaluable sections drawing on the authors’ extensive experience. Part 1 comprehensively reviews chronic psychosis, the history of mental health services, and treatment interventions, with the recurring theme of stigma and discrimination. Part 2 focuses on employment, with a clear description of the two planks of modern vocational rehabilitation: supported employment and social firms.

The book is optimistic, arguing that innovative model projects can be generalised to wider use, an optimism that may clash with clinicians’ experiences of mental health systems with top-down development and resource pressures. But it is also realistic, for example, acknowledging that not all people with chronic psychosis can be employed in the mainstream (though the research suggests 50–60% can, rather than 10–15% typically found), so work alternatives such as reformed sheltered work have a place.

The book is limited on forensic services and the challenges of fostering the social inclusion of mentally disordered offenders. There is the omission of disability discrimination legislation (and the wider context of government social policy) and, indeed, how mental health might be seen as part of the wider disability movement. And I would have welcomed more guidance on substance misuse, physical health and exercise, and on inequities in accessing healthcare. But these are quibbles. This is an excellent book, written with great compassion and with an emphasis on the person, citizenship and solutions.

David O’Flynn
Lambeth Hospital, Landor Road, London SW9 9NT, UK.
Email: david.o’flynn@lam.nhs.uk
doi: 10.1192/bjp.bp.106.031682

Cognitive Therapy of Schizophrenia

This is an immensely practical therapy manual for use in real-world clinical practice. The authors have avoided writing a step-by-step ‘how to’ text that would risk oversimplifying this complex area. Instead, they help readers conceptualise the huge variety of presentations falling within this diagnostic category and emphasise the underlying principles and attitudes that are essential to delivering this intervention effectively.

The relapse prevention section, for example, is not the familiar listing of intervention protocols around the detection of early warning signs. Instead, it provides the reader with an understanding of the process of relapse in order to guide formulation; gives advice on how to raise the topic of relapse during periods of remission; and emphasises the need to maximise the client’s sense of control. All of these are essential considerations if techniques are to be applied effectively in practice.

The slower pace and more elusive structure of cognitive-behavioural therapy (CBT) for psychosis can often leave therapists feeling confused and deskilled. Rather than present an idealised ‘expert’ account of therapy that can add to these feelings, the authors stay true to their therapeutic principles; they normalise many of the anxieties therapists are likely to feel and address many of the obstacles they are likely to face.

The book makes an intriguing read as the two authors are psychiatrists and so present an interweaving of the psychiatric and psychological perspectives that are too frequently seen pitted against each other. The result is perhaps a less thorough
consideration of a psychological explanation of psychosis than other texts in this area, but a clear presentation of the biological, social and psychological explanations for different symptoms of schizophrenia. This interweaving becomes a little confusing in the general assessment and formulation chapters, as the predominantly psychiatric assessment outlined does not seem readily to translate into a CBT formulation. This is made much clearer in the symptom-specific chapters.

Overall, the book reflects the authors' considerable experience disseminating this therapeutic approach for wider application in mental health settings. It makes for an interesting read, and is particularly recommended to mental health professionals who are already familiar with CBT and work within a psychiatric service setting.

Rebecca Rollinson Elizabeth Fry Building, University of East Anglia, Norwich NR4 7TJ, UK. Email: r.rollinson@uea.ac.uk

doi: 10.1192/bjp.bp.106.023192

Cognition + Addiction

This excellent book gives the reader an authoritative update on current psychological thinking in the addictions. The central theme is that the dominant views of addictive behaviour, of which social learning theory is a prime example, fail to take account of automatic cognitions and, therefore, fail to account adequately for relapse situations. The problem with existing theories is that they make the assumption that humans are rational decision-makers; readers will know that this is often not the case. Moreover, it is in the nature of dependence on psychoactive drugs that rational decisions are overruled by the desire for instant gratification.

The essential proposition running throughout the book is that psychological theories of addiction need to take account of automatic cognitive processes, that is processes that are both uncontrollable and mainly unconscious. In one chapter this is nicely described in computing language as 'drug related stimuli grabbing cognitive processing'. Evidence is presented for substance users having an attentional bias towards substance using situations, which in turn selectively brings to the fore related memories, reinforcing the addictive behaviour.

There are ten chapters. The first describes current psychological theories of addiction. The second and third chapters describe the evidence for automatic cognitive processes, notably attentional biases and automatic memories. Chapters four and five are on the more familiar clinical territory of motivational interviewing and understanding 'loss of control' from the perspective of automatic processes. Chapters six to eight cover special interest topics, namely genetics, opiate-specific cognitions and neurochemical processes. The final chapters bring together the implications for the psychological research described into clinical practice.

If there is a weakness, it is that the clinician will be left uncertain of the implications for day-to-day practice. The authors of the final chapters make a good attempt at answering this but, in truth, the point of the book is as much about laying down a challenge for practitioners as providing answers. The book deserves reading from cover to cover - stimulating, informative and well written.

Duncan Raistrick Leeds Addiction Unit, 19 Springfield Mount, Leeds LS2 9NG, UK. Email: duncan.raistrick@leedsmh.nhs.uk
doi: 10.1192/bjp.bp.106.033944
From the Editor’s desk

PETER TYRER

THE ALARM OF THE UNEXPECTED

One of our senior colleagues in the Royal College of Psychiatrists was recollecting the time when as a very young girl she was somewhat apprehensively travelling by train from Hampshire to London. As the train chugged its way past Basingstoke, Winchfield and Farnborough, she was reassured and dropped off to sleep. She woke as the train was drawing in to another station. ‘Wo-king’, she read on the platform as the train juddered to a halt. She was filled with immediate dread. The train had been hijacked, diverted or in some magical way been transported to China, and she would never see London again. She has never quite recovered and still feels discomfort whenever the word Woking is mentioned. This alarm of the unexpected – or the Woking effect – is demonstrated in some of the challenging papers in this issue. It can serve as a valuable wake-up call. So, after reading the editorial by Burke et al (pp. 371–372) dust off those old ideas that dementia is totally irreversible, that herbal remedies (especially Wo-king ones) have no long-term significance (Skodol et al, pp. 379–384), or that personality problems in adolescence are commonplace, and also unreliably assessed, so have no long-term significance (Skodol et al, pp. 415–420). Few things in psychiatry can be more alarming to the uninformed than the prospect of ECT, yet the paper by Mund-Olsen et al (pp. 435–439) suggests that, once the immediate effects of treatment are over, its recipients live longer than those with disorders not treated with ECT. In my early psychiatric life, one of the most alarming experiences was the invasion of an international symposium in Paris in 1974 by a group of sciento-logists. Fortunately, Martin Roth, the subject of our illuminating appreciation (Kerr & Kay, pp. 375–378), was in the chair and was completely ready for the unexpected. In his magisterial way, he thanked everybody, including the contributors not on the programme, for their contributions, and drew the meeting to a quiet close.

A NEW AFRICAN PSYCHIATRIC JOURNAL

The first regional conference of the World Psychiatric Association to be held in central Africa took place at the end of March in Nairobi. In addition to a full programme describing the challenges and opportunities in the region it also witnessed the birth of a new psychiatric journal for Africa. This journal, yet to be formally named, will cover the whole of Africa and will also be the official journal of the African Association of Psychiatrists and Allied Professionals, as well as a sister journal to our own International Psychiatry. Why should this matter, and why should a continent containing fewer than 1000 psychiatrists need its own journal? The reasons are many. We need to reduce the obstinately persisting 10/90 divide in mental health research (Saxena et al, 2006) and boost the efforts of those wishing to establish an evidence base for interventions that can no longer continue to be extrapolated from the well-resourced to the poor. There is something slightly obscene in the sometimes patronising way in which rich countries attempt to advise on how to provide services in poor ones and yet at the same time encourage covertly the emigration of psychiatrists to reinforce their own well resourced services. The more we can improve the capacity and self-confidence of the poorer countries the better will be their output in terms of service provision and research (Tyrer, 2005). We already have had some outstanding papers published from Ethiopia, Nigeria and Uganda in recent times (Gureje et al, 2005, 2006; North et al, 2005; Bass et al, 2006; Mogga et al, 2006), but these are not wholly indigenous as apart from the papers by Gureje and his colleagues they have all involved authors from the UK or USA. Such collaboration is not to be decried, but if young African researchers do not have their own journal to which they can submit their papers, both good and not so good, they will continue to look for help from outside, and this, although welcome, runs the danger of an agenda externally set for Western-oriented research. So we hope the new journal will also reduce the emigration of manuscripts, as much as that of psychiatrists, from Africa and both we at the Journal and many others in the Royal College wish the new journal well.


