THE BRITISH JOURNAL OF PSYCHIATRY

Volume 191, August 2007, pp.97-188

Psychiatry in pictures

Psychiatry in pictures
ALLAN BEVERIDGE

Highlights of this issue

Highlights of this issue
KIMBERLIE DEAN

EDITORIALS

The devil is in the detail: partnerships between psychiatry and faith-based organisations
GERARD LEAVEY and MICHAEL KING

REVIEW ARTICLES

Ethnicity and the Mental Health Act 1983
SWARAN P. SINGH, NAN GREENWOOD, SARAH WHITE, and RACHEL CHURCHILL

PAPERS

Relationship between daily suicide counts and temperature in England and Wales
LISA A. PAGE, SHAKOOR HAJAT, and R. SARI KOVATS

Cortical white-matter microstructure in schizophrenia: Diffusion imaging study

Frontal release signs and cognition in people with schizophrenia, their siblings and healthy controls
THOMAS M. HYDE, TERRY E. GOLDBERG, MICHAEL F. EGAN, MARC C. Lener, and DANIEL R. WEINBERGER
Emotion recognition and genetic vulnerability to schizophrenia
BENOIT BEDIOU, FATIMA ASRI, JEROME BRUNELIN, PIERRE KROLAK-SALMON, THIERRY D’AMATO, MOHAMED SAOUĐ, and IMANE TAZI

Long-acting injectable risperidone vs olanzapine tablets for schizophrenia or schizoaffective disorder: Randomised, controlled, open-label study
NICHOLAS A. KEKS, MICHAEL INGHAM, AKBAR KHAN, and KEITH KARCHER

Deployment-related stress and trauma in Dutch soldiers returning from Iraq: Prospective study
IRIS M. ENGELHARD, MARCEL A. VAN DEN HOUT, JOS WEERTS, ARNOUD ARNTZ, JOOP J. C. M. HOX, and RICHARD J. MCNALLY

Social identification and post-traumatic stress symptoms in post-conflict Northern Ireland
ORLA T. MULDOON and CIARA DOWNES
Br J Psychiatry 2007 191: 146-149.

Prevalence of dementia in intellectual disability using different diagnostic criteria
A. STRYDOM, G. LIVINGSTON, M. KING, and A. HASSIOTIS

Associations between common mental disorders and the Mental Illness Needs Index in community settings: Multilevel analysis
DAVID L. FONE, FRANK DUNSTAN, ANN JOHN, and KEITH LLOYD

Treatment of depression in primary care: Socio-economic status, clinical need and receipt of treatment
SCOTT WEICH, IRWIN NAZARETH, LOUISE MORGAN, and MICHAEL KING

Characteristics and activities of acute psychiatric in-patient facilities: national survey in Italy
GIOVANNI DE GIROLAMO, ANGELO BARBATO, RENATA BRACCO, ANDREA GADDINI, ROSSELLA MIGLIO, PIERLUIGI MOROSINI, BRUNO NORCIO, ANGELO PICARDI, ELISABETTA ROSSI, PAOLA RUCCI, GIOVANNI SANTONE, GIUSEPPE DELL’ACQUA the PROGRES—Acute group

SHORT REPORTS

Manifest disease and motor cortex reactivity in twins discordant for schizophrenia
MARTIN SCHÜRMANN, JUHA JÄRVELÄINEN, SARI AVIKAINEN, TYRONE D. CANNON, JOUKO LÖNNQVIST, MATTI HUUTTUNEN, and RIITTA HARI

Correspondence

Genetic hypotheses for schizophrenia
T. Crow
Authors’ reply
Jon McClellan, Ezra Susser, and Mary-Claire King

Anticipation and the genetics of psychosis
M. Blinc-Pesek and M. Agius

Attention-deficit hyperactivity disorder and mood disorders in adults
A. J. Kuan and A. H. Young

Authors’ reply:
P. Asherson, A. Stringaris, W. Chen, and E. Taylor

One hundred years ago

One hundred years ago
Henry Rollin

Corrigenda

Corrigenda

Book reviews

Journalists Under Fire: The Psychological Hazards of Covering War
Martin Deahl

Forensic Psychiatry: Influences of Evil
Anthony Maden
Risk and Resilience: Adaptations in Changing Times
Ian Goodyer

From the Editor's desk

PETER TYRER
Br J Psychiatry 2007 191: 188.
As far back as 300 BC native healers in Sri Lanka have been involved in the treatment of physical and mental illness. The above images show wooden masks used by traditional healers (kattadiya) in Sri Lanka to alleviate certain forms of mental distress believed to be caused by demonic spirits, especially in rural indigenous communities. The carvings symbolise the fears, dilemmas and anxieties of the patient. The healer wears special clothes and a mask (wesmuwu). He carries a torch (pandam) and engages in a ritualistic dance (nadagam) to the hypnotic beat of drums (yakbera). This creates an atmosphere which unifies the demonic spirits and the healer. This enables the healer to understand the nature of the illness and, after paying tribute to the demon, he is able to treat the ailment.
Highlights of this issue

BY KIMBERLIE DEAN

BRAIN STRUCTURE AND FUNCTION IN SCHIZOPHRENIA

In a large, community-based sample, Andreone et al (pp. 113–119) were able to confirm the findings of previous, smaller diffusion-weighted imaging studies in schizophrenia. Evidence of cortical white-matter microstructure disruption was detected in the frontal and temporo-occipital lobes of individuals with schizophrenia compared with controls. The authors propose that abnormalities in myelination might account for their findings and call for future research to further explore white-matter integrity and genes for myelination, particularly in early-onset or high-risk groups. Focusing on frontal lobe functioning, Hyde et al (pp. 120–125) found that frontal release signs were more frequently identified in a group with schizophrenia in comparison with a group of their unaffected siblings or healthy controls. The relationship between frontal release signs and neuropsychological impairment was strongest for those with schizophrenia. In another study comparing those with schizophrenia both with their unaffected siblings and with unrelated controls, Bediou et al (pp. 126–130) found that facial recognition was impaired in the former two groups compared with controls, but gender recognition was preserved. The impairments in facial recognition for patients with schizophrenia persisted over time.

POST-CONFLICT OUTCOMES: IRAQ AND NORTHERN IRELAND

In a prospective study of Dutch troops deployed to Iraq in 2005, Engelhard et al (pp. 140–145) found that levels of distress remained stable following deployment, except in a small minority. They also found that post-traumatic stress disorder (PTSD) rates based on questionnaire responses were much higher than rates determined by diagnostic interview and that much of the PTSD identified was not directly related to deployment. Muldoon & Downes (pp. 146–149) examined the population occurrence of post-traumatic stress symptoms in post-conflict Northern Ireland and found evidence of probable PTSD in 10% of respondents. Those individuals with probable PTSD were less likely to rate national identity as important, and more likely to report direct experience of the 'troubles'. Those of lower socio-economic status were also more likely to report severe post-traumatic symptoms.

PRIMARY CARE, IN-PATIENT AND COMPULSORY MENTAL HEALTHCARE

Weich et al (pp. 164–169) found no evidence of an association between socio-economic status and treatment/adherence to treatment for depression in primary care in England and Wales. They did find that the lowest treatment rates were among the older groups. Acute in-patient psychiatric treatment is offered in both public and private facilities in Italy. De Girolamo et al (pp. 170–177) report that public beds account for 45.8% of in-patient provision overall, with substantial variation in relative provision by region. Their survey also highlighted problems with the physical environment in many in-patient facilities, a longer duration of stay in private facilities, lower staffing levels in private centres, and the relatively low rate of involuntary admission in Italy overall. In a systematic review of ethnicity and the Mental Health Act 1983, Singh et al (pp. 99–105) conclude that Black and minority ethnic status is independently associated with likelihood of psychiatric detention in the UK. Detention rates were found to be lower for first-episode patients, however.

ECOLOGICAL PERSPECTIVES ON SUICIDE AND COMMON MENTAL DISORDER

Over an 11-year period in England and Wales, Page et al (pp. 106–112) found that high daily temperatures (above a threshold of 18°C) were associated with an increased risk of suicide, particularly of suicide by violent means. The authors did not confirm previously reported seasonal peaks in suicide rates but there was evidence of a peak during the 1995, but not the 2003, heatwave. The Mental Illness Needs Index was found by Fone et al (pp. 158–163) to be strongly associated with rates of common mental disorder in a small-area ecological comparison. In further analysis, they found that the association persisted on an individual level even after account was taken of individual risk factors.

DEMENTIA AND INTELLECTUAL DISABILITY

In a cross-sectional survey, Strydom et al (pp. 150–157) were able to demonstrate that symptoms associated with all dementia subtypes are found in older adults with intellectual disability. Alzheimer’s disease was most common and was in fact three times more prevalent than expected. The authors also found that Lewy body and frontotemporal dementia were more common than vascular dementia in this group, and that DSM–IV criteria were more inclusive than ICD–10.

INJECTABLE RISPERIDONE V. ORAL OLANZAPINE

In a 12-month, randomised, controlled, open-label study involving individuals with schizophrenia or schizoaffective disorder, Keks et al (pp. 131–139) found that injectable long-acting risperidone was not inferior to olanzapine tablets in the short phase of the study (13 weeks). Over the 12-month period, both groups demonstrated improvements in symptom scores and few were seen to drop out of the study because of adverse events.
The devil is in the detail: partnerships between psychiatry and faith-based organisations

GERARD LEAVEY and MICHAEL KING

Summary  Clergy continue to have a central role in many communities and the utility of their involvement in the care of people with mental health problems is increasingly argued. However, there has been a failure to examine the form and parameters of partnerships between faith-based organisations and psychiatry.

Declaration of interest  None.

The view that we live in an increasingly secular world is widely held in Western societies. However, the nature of secularism remains complex and disputed, divided by public–private perspectives. Thus, for some it is the decline of religion in institutional and public life, whereas for others it is the dwindling of spiritual and religious consciousness among individuals. For the secularist, the trend towards religious disaffiliation in Western societies is desirable and inexorable. Critics of this position point to an increasingly private, à la carte spirituality, often divorced from regular religious observance (Roof, 2001). The secular debate notwithstanding, we live in an increasingly plural society where religion is an integral and immutable part of identity for many people, governing aspects of their beliefs and behaviour. Religion and religious identity in many Western countries has ceased to be a marginal concern and has moved centre stage, generating not just political debate but new policy and legislation. In the UK this has been borne out by the inclusion of a religious affiliation question in the 2001 census. The issues of religious freedom and equality have been established within the Human Rights Act and the European Convention on Human Rights.

The past decade has seen a growing demand for health professionals to take better account of patients’ religious beliefs and establish links with faith-based organisations as partners in health and welfare services (Mental Health Foundation, 1997; National Institute for Mental Health in England, 2003). The rationale for this partnership is based on faith communities’ declared commitment to entwined spiritual and social values, and their deep-rooted social connections (Home Office Faith Communities Unit, 2004). These emerge in the health literature as the potential public health value attached to notions of social capital (McKenzie et al, 2002). Religion-based communities are considered to be exemplars of social capital ideals of reciprocity, integration, socialisation, activism and voluntarism, which are thought to solidify the community and benefit the individual. This counter-anomic vision of religion suggests that the incorporation of faith-based organisations as adjuncts to statutory sector health and welfare is a sensible move. However, although psychiatry and religion share similar concerns, their relationship has seldom been harmonious, with perhaps just cause for suspicion on both sides (Bhugra, 1997). In this editorial we outline some of the key issues in clergy–psychiatry partnerships, pointing to the reasons why building partnerships with faith groups and clergy is useful and necessary. However, although there is a need for dialogue and mutual understanding, there is also a need for psychiatry to examine the nature and boundaries of proposed relationships.

EXPLANATORY MODELS, CLERGY AND HELP-SEEKING

In the USA, analysis of the National Comorbidity Survey conducted in the 1990s revealed continuing use of the clergy alone and alongside professional and alternative practitioners for mental health problems. However, given that many self-help groups and recovery movements are faith-based, the study might have underestimated the role of religion in the delivery of services to patients with mental illness (Wang et al, 2003). Nevertheless, it is clear that faith-based organisations and their clergy are contacted by people with mental health problems, often in preference to consulting psychiatric professionals. This preference is influenced by somewhat intersecting cultural and service-related factors. These can be characterised in positive and negative terms.

Broadly speaking, healing has been observed as a central function of most religions and some people look to religion as a means of understanding suffering and as a beneficial way of coping with it. However, from a more negative standpoint, some of the reluctance of congregation members to consult psychiatry, psychotherapy or counselling services may be explained by a posited ‘religiosity gap’ between the religious patient and mental health professionals. There is some evidence, often anecdotal, that psychiatrists are viewed suspiciously by religious adherents (Mitchell & Baker, 2000). Consequently, as a form of self-protection such patients may conceal their religious beliefs, fearful that they will be regarded as further indications of mental pathology (Leavey, 2004). Additionally, ethno-cultural beliefs of patients and their families determine help-seeking. Studies in the UK indicate that culturally mediated religious beliefs influence differential access to, and engagement and satisfaction with, services (McCabe & Priebe, 2004).

Members of many minority ethnic communities are more exposed to risk factors for mental health problems (unemployment, poor housing and discrimination) than their majority White counterparts. Moreover, it is important to bear in mind that in many minority ethnic communities, particularly the recently arrived, clergy have a pivotal role as gatekeepers for services, advisors and mediators between government and communities. The clergy are popularly conceived of as knowledgeable and trusted brokers at local and personal levels. With respect to mental illness, clergy in closed, less secularised communities may play a pivotal part when people first present with psychological difficulties, thereby strengthening or challenging religious health beliefs and, in effect, advocating spiritual or secular intervention (Littlewood & Dein, 1995). Among religious adherents there may be a demand...
for clergy to be absolutist and directive, particularly when depression and anxiety are framed as moral disorders.

**PROBLEMS OF PARTNERSHIPS**

Generally, when clergy–psychiatry partnerships are proposed there are a number of underlying assumptions. Dominant among these is that dialogue and mutual understanding will assuage the suspicions and estrangement between the two sectors. It is also often assumed that all clergy share a biomedical conceptualisation of illness and are willing to offer time and resources in the care of psychiatric patients. However, we need to be clear what is meant by the notion of partnership. As we discuss, these assumptions ignore difficult issues on both sides.

First, although many clergy already provide pastoral care for emotionally distressed people, they may be reluctant to move further away from spiritual guidance – their ‘core business’ – towards a more secular enterprise. Second, there are considerable difficulties for psychiatry in the disjunction between biomedical and spiritual concepts of severe mental illness, their origins and their resolution. When religious individuals and their clergy have coincident beliefs about the supernatural origins of illness it seems likely that this will have serious implications for pathways to appropriate care and compliance with treatment. Although it is important to stress the heterogeneity of beliefs about suffering and healing found among mainstream organisations, this concern has particular resonance in relation to certain evangelical and Pentecostal churches which maintain deeply held beliefs and practices surrounding demonic possession, healing and deliverance rituals. Recent high-profile cases in the UK, such as that of Victoria Climbie (Laming, 2003), point to the strength of such beliefs among African Pentecostal churches and the potential for tragedy. How should mental health professionals engage with clergy who believe that sin or demonic possession lies at the root of a person’s illness? It seems unlikely for clinical and legal reasons that services could or should collude with religious healing rituals. Similarly, contested normative values, such as those related to sexuality, are not easily reconciled and may require clinicians to question and challenge fundamental tenets of certain faith groups. If this is the case, aside from any consideration of resources for training and personnel, the psychiatric role may become a persuasive and didactic one rather than a process of mutual enlightenment.

Third, our globalised post-modern world is characterised by diversity and pluralism. This is also true of belief systems and faith groups. Although it is likely that the larger, more established faith groups are more easily approached, how inclusive should be the dialogue and collaboration? How should we evaluate the representativeness of minority religious or faith groups? Should health providers reach out to groups that might be considered too ‘fringe’ – the independent Pentecostal churches, the Moonies or the International Society for Krishna Consciousness, for example? To ignore them might leave health providers vulnerable to accusations of discrimination and the possibility of litigation. Moreover, paradoxically, it may be among the smaller, more esoteric and less visible groups that dialogue with mental health services is most profitable.

**CONCLUSIONS**

Faith-based organisations across the religious spectrum have a pivotal role in the lives of many people, particularly those of Black and minority ethnic communities. As Western governments are beginning to realise, there is no doubt that the skills and capacity of faith communities are underused by health and welfare providers. However, the potential contribution of such organisations and how it would fit with existing statutory provision have yet to be fully examined. The purpose of this editorial is not to decry the need for dialogue and collaboration with clergy, but rather to begin a realistic appraisal of the difficulties, both clinical and legal, and the resources required for such partnerships.

**REFERENCES**


Mental Health Foundation (1997) Knowing Our Own Minds. Mental Health Foundation.


Ethnicity and the Mental Health Act 1983

Systematic review

SWARAN P. SINGH, NAN GREENWOOD, SARAH WHITE and RACHEL CHURCHILL

Background  Black and minority ethnic (BME) patients are disproportionately detained under the Mental Health Act 1983. There has been no systematic exploration of differences within and between ethnic groups, nor of the explanations put forward for this excess.

Aims  To systematically review detention and ethnicity, with meta-analyses of detention rates for BME groups, and to explore the explanations offered for ethnic differences in detention rates.

Method  Literature search and meta-analysis. Explanations offered were categorised, supporting literature was accessed and the strength of the evidence evaluated.

Results  In all, 49 studies met inclusion criteria; of these, 19 were included in the meta-analyses. Compared with White patients, Black patients were 3.83 times, BME patients 3.35 times and Asian patients 2.06 times more likely to be detained. The most common explanations related to misdiagnosis and discrimination against BME patients, higher incidence of psychosis and differences in illness expression. Many explanations, including that of racism within mental health services, were not supported by clear evidence.

Conclusions  Although BME status predicts psychiatric detention in the UK, most explanations offered for the excess detention of BME patients are largely unsupported.

Declaration of interest  None.

Over the past 20 years several studies have reported that a disproportionate number of patients from Black and minority ethnic (BME) populations within the UK are compulsorily detained under both civil and forensic sections of the Mental Health Act 1983 (Churchill et al, 1999; Bhiu et al, 2003; Morgan et al, 2004). However, some studies have not found this overrepresentation, with some evidence that it may not apply to certain groups, such as people with a first episode of psychosis (Cole et al, 1995; Burnett et al, 1999). There is also evidence that detention rates may not be excessive for all ethnic minority patients. Rates for Asian patients, for example, lie between those for Black (Black Caribbean and Black African) and White patients (Audini & Lelliott, 2002). The presence of such inequalities in service provision is important to service users, service providers and policymakers. For service users and carers, traumatic experiences of detention and coercion can lead to long-term aversion to mental healthcare. From a clinical perspective, such negative experiences cause mistrust and resistance to intervention, with delayed help-seeking and the necessity for further coercion (Singh, 2001; Morgan et al, 2004).

Several hypotheses have been put forward to explain this excess. These can be broadly divided into patient-related and service-related explanations (Littlewood, 1986). Patient-related explanations include higher rates of psychosis (Bebbington et al, 1994), perceptions of Black and minority ethnic patients being at greater risk (Lewis et al, 1990) and poorer insight in this group (van Os et al, 1996). Greater stigma associated with mental illness within minority communities leading to delays in help-seeking and more severe symptoms at presentation have also been offered as explanations (Harrison et al, 1989). Service-related explanations have focused on inherent racism within psychiatry (Littlewood & Lipsedge, 1997) with associated ‘Eurocentric’ misdiagnosis (Fernando, 1988) and perceptions among Black patients of services being inaccessible and inappropriate (Cochrane & Sashidharan, 1996). There are two narrative reviews of such explanations (Littlewood, 1986; Spector, 2001), but a systematic and structured review determining the strength of evidence for the various explanations for this excess is lacking. We conducted a systematic review of all UK literature on ethnicity and detention to:

(a) examine the evidence for greater detention of Black and minority ethnic patients within psychiatric services in the UK;
(b) explore differences between ethnic minority groups;
(c) determine the full range of hypotheses put forward to account for any such excess;
(d) examine the evidence for these hypotheses within the literature.

METHOD

A literature search was undertaken of studies relating to the Mental Health Act in the UK, both civil and forensic sections, published between 1984 and April 2005. The following bibliographic databases were searched: Medline, EMBASE, PsycINFO, the Cumulative Index to Nursing and Allied Health Literature (CINAHL), the Applied Social Sciences Index and Abstracts (ASSIA), the Health Management Information Consortium (HMIC), Web of Science, the Cochrane database, the System for Information on Grey Literature (SIGLE) and the National Research Register. The CD-ROM for the British National Bibliography was searched for relevant books. The electronic database search terms were divided into four sets: Mental Health Act terms; mental illness and forensic psychiatry terms; compulsory detention; and ethnic group terms. A combination of search terms from these sets was applied. Where Medical Subject Headings (MeSH) terms were available in the databases, these were exploded and combined. The bibliographies of relevant works were checked for articles missed by the initial search. Key review papers and published bibliographies in the area were also scrutinised for relevant studies.

Inclusion criteria

Studies had to fulfil the following inclusion criteria: publication in English; reference
made to the use of compulsion to detain a person under the Mental Health Act 1983 in England and Wales; provision of original data relating to the Mental Health Act; and inclusion of two or more ethnic groups in the study.

The relevance of the literature was initially ascertained from the titles. N.G. and S.S. independently looked at the titles of the first 250 studies in the database searches and agreed on the relevance of all but one article. Discussion of this article led to an improved understanding of the criteria and N.G. then continued with the remaining articles. Where titles appeared relevant, abstracts or equivalent summary information were studied. Just over two hundred \( n=210 \) hard copies of studies appearing pertinent from the abstracts were obtained. Further analysis of the full articles revealed that many of these did not fit the inclusion criteria and they were then excluded. Selected articles were read and the inclusion criteria applied independently by both N.G. and S.S. before the final selection was made.

### Personal communication with experts

Once the articles for the review had been selected, 24 experts were sent the list of included studies and asked whether there were any further studies they could suggest. Five experts responded; however, their suggestions for additional studies had already been considered. One expert did not provide any studies, but expressed unhappiness that we had excluded case histories and therefore considered our review to be ‘invalid’. We did explain that this was a meta-analysis of data-based studies and by definition case studies could not be included.

### Quality ratings

Literature quality was assessed using an adaptation of a scale (see data supplement 1 to the online version of this article) previously used in a similar review (Bhui et al, 2003). The resulting quality scores ranged from 0 to 14 and were divided into low (0–5), medium (6–10) and high (11–14). N.G. and S.S. rated five articles together to ensure consistent application of the scale and then the rest were rated independently. There was agreement on all but five studies, but discussion revealed that these differences were due to differing interpretations of the scale. Once this was resolved, complete consensus was reached on appropriate ratings for each study.

### Data extraction

For meta-analysis, raw data were extracted independently by N.G. and S.S. Categories of explanatory evidence emerged as successive papers were studied; data regarding explanations were extracted independently by N.G. and S.S. and consensus was reached regarding categorisation of explanations. Explanations were recorded as presented in the original paper and no attempt was made to interpret the text to fit any a priori hypothesis. Only explanations relating specifically to ethnic differences in detention rates were included. For instance in papers discussing ethnic differences in admission rates in general rather than Mental Health Act detention rates specifically, explanations were not included in the results. Some explanations were difficult to categorise, such as poor adherence, which could potentially be assigned to more than one category; a judgement was made in these cases as to the most appropriate category. Study authors sometimes offered similar explanations but for different reasons, especially for complex phenomena such as delay in help-seeking among Black patients, which in turn might lead to more disturbed presentation with greater risk of detention. Some authors attributed this delay to lack of social support, whereas others attributed it to denial of illness. Such explanations therefore appear in more than one category. Perception of Black and minority ethnic patients as more violent or at higher risk was categorised separately from studies showing differences in clinical presentation between ethnic groups.

### Level of evidence

Each study providing an explanation was scrutinised for the level of evidence for the explanation. Evidence was further categorised as primary evidence, secondary evidence or no evidence. Primary evidence was defined as direct evidence for an explanation provided by a study using its own data. This was further categorised as evidence at the level of an ‘association’ if the data demonstrated correlation between variables where confounders were not controlled and causal interpretations could not be made. An example would be studies where Black and minority ethnic patients were more likely to be detained but also more likely to be diagnosed with psychosis and it was not certain whether ethnicity or psychotic illness was the primary reason for the excess of detentions (especially if tests of association such as chi-squared tests rather than regression had been employed). Secondary evidence was defined as citations of other papers to support an explanation. These secondary citations were perused and key findings summarised, including (where possible) the strength of evidence for relevant conclusions drawn. A few authors discussed explanations for detention rates among Asian patients and these are distinguished from other explanations.

### Analyses

Meta-analyses were performed where aggregate data of minority ethnic and White compulsorily admitted patients were provided. Pooled odds ratios were calculated for the overall comparisons using the fixed-effects model. The chi-squared test for heterogeneity was then performed to determine whether there was significant heterogeneity in the odds ratios between studies. For comparisons in which there was significant heterogeneity, four possible source variables for the heterogeneity were investigated. These were patient type (civil, forensic, mixed), episode (first episode, mixed episode), quality rating (high, medium, low) and year of publication. Pooled odds ratios and 95% confidence intervals are presented for studies within each grouping created by the categorical variables. Year of publication was categorised as studies from 1980s, from the period 1990 to 1994, from 1995 to 1999 and from 2000 onwards. Meta-regression was performed, plotting the log odds ratio for each study against year of publication, using appropriate weighting. All meta-analysis was carried out using Comprehensive Meta-Analysis version 2.2 for Windows.

One study (Goater et al, 1999) included three sets of data (at admission, year 1 and year 5), each of which reported different detention rates among Black and minority ethnic patients. Each set was treated as independent and included separately in the meta-analyses.

### Terminology

In this paper the term ‘Black and minority ethnic’ is used to refer to participants of any ethnic group other than White. The term ‘Black’ refers to people of Black African, Black Caribbean and ‘Black other’ groups. The term ‘Asian’ is used for people...
RESULTS

Forty-nine studies met the inclusion criteria and were included in the review but only 19 provided raw data to permit meta-analysis. Table DS2.1 in data supplement 2 to the online version of this paper gives details of the 49 studies listed alphabetically by the first author. Research was mainly concentrated in major cities (71% of studies were from London, with 32% from the Institute of Psychiatry, the Maudsley Hospital or King’s College). Most studies were cross-sectional and relied on routinely collected data. Some studies included both retrospective and prospective data; just over half used only retrospective data and a fifth were prospective studies. Sample size varied from 20 patients (Anderson & Parrot, 1995) to 31,702 admissions (Audini & Lelliott, 2002), and just over half (53%) included fewer than 120 patients. Few studies were hypothesis-driven and only 39% stated inclusion and exclusion criteria. No study included power calculations.

Figure 1 is a forest plot of the studies included in the meta-analyses, with odds ratios and 95% confidence intervals for each study on a horizontal plane and the pooled effect displayed with a diamond marker. Table 1 provides a summary of the meta-analyses of four main ethnic group comparisons: Black and minority ethnic (BME) compared with White; Black compared with White; Asian compared with White; and Asian compared with Black. Within these ethnic group comparisons and where there were sufficient data, subgroups such as patient types and illness episodes were also analysed.

Ethnicity

Overall pooled odds ratios for BME compared with White (3.35, 95% CI 3.05–3.73, P < 0.0001) and Black compared with White (3.83, 95% CI 3.42–4.29, P < 0.0001) were similar. The odds for Asian compared with White (2.06, 95% CI 1.60–2.65, P < 0.0001) and Black compared with Asian (2.25, 95% CI 1.72–2.94, P < 0.0001) were both close to 2. Put slightly differently, compared with White patients, Asian patients were approximately twice as likely and Black patients approximately four times as likely to be detained.

Civil and forensic detentions

The pooled odds ratios of detention type showed that the excesses of BME (4.03, 95% CI 3.37–4.81, P < 0.0001) and Black (4.48, 95% CI 3.71–5.41, P < 0.0001) patients compared with White patients for civil detentions are greater than for forensic detentions (BME: 2.29, 95% CI 1.50–3.50, P < 0.0001; Black: 2.45, 95% CI 1.57–3.82, P < 0.001). The odds ratios differ significantly between the patient type groups for the Black v. White (P = 0.031) and the BME v. White comparisons (P = 0.017). The Black v. Asian comparison was non-significant (P = 0.115) and although the Asian v. White comparison was statistically significant, this should be viewed with caution because only one forensic study was included.

Illness episode

There was also an effect for illness episode across different ethnic comparisons, with first-episode BME (2.15, 95% CI 1.55–2.98, P < 0.0001) and Black patients (2.42, 95% CI 1.74–3.38, P < 0.001) less likely to be detained than later mixed-episode BME (3.53, 95% CI 3.16–3.95, P < 0.0001) and Black patients (4.06, 95% CI 3.60–4.59, P < 0.0001).

Quality

Studies rated as high quality in both the BME v. White and Black v. White comparisons showed lower summarised odds than low- and medium-quality studies. This effect was statistically significant in the Black v. White comparison (P = 0.03), but...
Table 1  Results of the meta-analyses: pooled odds ratios

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Number of data-sets</th>
<th>Odds ratio (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>BME v. White</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>21</td>
<td>3.35 (3.05–3.73)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Patient type</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civil</td>
<td>15</td>
<td>4.03 (3.37–4.81)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Forensic</td>
<td>2</td>
<td>2.29 (1.50–3.50)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mixed</td>
<td>4</td>
<td>3.12 (2.72–3.59)</td>
<td>0.003</td>
</tr>
<tr>
<td>Illness episode</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First episode</td>
<td>3</td>
<td>2.15 (1.55–2.98)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mixed episode</td>
<td>18</td>
<td>3.53 (3.16–3.95)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Black v. White</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>21</td>
<td>3.83 (3.42–4.29)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Patient type</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civil</td>
<td>15</td>
<td>4.48 (3.71–5.41)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Forensic</td>
<td>2</td>
<td>2.45 (1.57–3.82)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mixed</td>
<td>4</td>
<td>3.65 (3.14–4.29)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Illness episode</td>
<td>21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First episode</td>
<td>3</td>
<td>2.42 (1.74–3.38)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mixed episode</td>
<td>18</td>
<td>4.06 (3.60–4.59)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Asian v. White</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>5</td>
<td>2.06 (1.60–2.65)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Patient type</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civil</td>
<td>4</td>
<td>3.42 (2.31–5.07)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Mixed</td>
<td>1</td>
<td>1.45 (1.04–2.00)</td>
<td>0.028</td>
</tr>
<tr>
<td>Illness episode</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First episode</td>
<td>1</td>
<td>0.39 (0.113–1.37)</td>
<td>0.142</td>
</tr>
<tr>
<td>Mixed episode</td>
<td>4</td>
<td>2.21 (1.71–2.86)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Black v. Asian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Overall</td>
<td>5</td>
<td>2.25 (1.72–2.94)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Patient type</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civil</td>
<td>4</td>
<td>1.76 (1.18–2.64)</td>
<td>0.0006</td>
</tr>
<tr>
<td>Mixed</td>
<td>1</td>
<td>2.72 (1.90–3.88)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Illness episode</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>First episode</td>
<td>1</td>
<td>3.16 (0.87–1.45)</td>
<td>0.0800</td>
</tr>
<tr>
<td>Mixed episode</td>
<td>4</td>
<td>2.21 (1.68–2.91)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

BME, Black and minority ethnic.

not in the BME v. White comparison (P=0.16).

Publication date

Overall the odds ratio decreased significantly with study publication date for both the BME v. White (P=0.001) and Black v. White comparisons (P=0.001). The Asian v. White comparison approached significance (P=0.06) whereas the Black v. Asian comparison was non-significant (P=0.55). There was a statistical correlation between higher quality and recency of publication (P<0.01).

Explanations for the excess

Five categories of explanations emerged from the 49 studies included in the review. These were categorised as ‘patient-related’, ‘illness-related’, ‘service-related’, ‘culture-related’ and ‘patient–service interaction related’. Each category of explanation and literature offered to support it are presented in separate tables (Tables DS2.2–2.6 in data supplement 2 to the online version of this paper). The right-hand columns in each table describe the level of evidence offered for each explanation. Papers presenting evidence against that particular explanation are grouped at the end of each table.

Patient-related explanations

Patient-related explanations (Table DS2.2) included theories that higher rates of detention occur because Black and minority ethnic patients have higher rates of psychoses, are perceived as being at greater risk of violence and disturbed behaviour, have higher rates of comorbid drug use and have greater delays in help-seeking. Much of the evidence for these explanations came from secondary citations, with little primary evidence, especially for explanations such as comorbid drug use and delayed help-seeking. A few studies reported primary evidence that the effect of ethnicity could be entirely explained by an interaction between diagnosis and challenging behaviour. Some studies found that even when such variables were controlled for, BME status remained a predictor of detention.

Illness-related explanations

Explanations in this category (online Table DS2.3) related to different illness expression in Black and minority ethnic patients, with more challenging behaviour or violence, association with offending behaviour, poorer adherence and greater denial of illness, all of which could account for higher rates of detention. Much of the evidence was of a secondary nature, with one study reporting no ethnic difference in clinical presentation of psychotic disorders.

Service-related explanations

Service-related explanations (online Table DS2.4) included the possibilities that excess detentions could be explained by underrecognition and misdiagnosis of mental illness in Black and minority ethnic patients, lower likelihood of referral to specialist services, greater contact with the police, and racial stereotyping and discrimination within both the mental health and the criminal justice system. There was some secondary evidence of underrecognition of psychiatric problems in such patients and the possible role of racial stereotyping.

Other explanations

The other two sets of explanations, culture-related (online Table DS2.5) and patient–service interaction (online Table DS2.6), included a mixed set of explanations ranging from cultural differences in explanatory models of illness, stigma of mental illness in Black and minority ethnic communities, alienation from and mistrust of services due to negative perceptions and experiences, and unwillingness to seek help.
Of all these explanatory categories, culture-related explanations had the fewest supporting citations. Negative perceptions of services, with mistrust and poor engagement, dominated the service–patient interface explanations, but there was lack of supportive primary evidence.

Overall, racial stereotyping, labelling and discrimination against Black and minority ethnic patients was the most often cited explanation and appeared in 15 papers (31%); this was followed by alienation, dissatisfaction, negative perceptions and mistrust of psychiatric services (in 26% papers), greater perception of violence (22%), higher rates of psychosis (22%), delay in help-seeking and poor social support (18%) and misdiagnosis, underrecognition of mental illness with lower referral rates to specialist services (16%). If the perception of Black patients as more violent or at greater risk is considered as part of the 'racial stereotyping' category, then this 'race-based' explanation was offered in 53% of the studies. There was no primary evidence provided by most studies to confirm any of these explanations, and some papers presented data that contradicted these explanations – for instance, some studies showed that the effect of ethnicity could be accounted for by an interaction between age, gender, diagnosis and challenging behaviour.

DISCUSSION

Excess rates of detention among certain Black and minority ethnic groups have been a major cause of concern for service users, health service providers and policy makers. Reducing ‘disproportionate rates of compulsory detention of BME users’ is a key aim of the government report Delivering Race Equality in Mental Health Care (Department of Health, 2005). Psychiatry and psychiatric services have been accused of being explicitly and implicitly racist both in service provision and diagnosis (Fernando, 1988; Littlewood & Lipsedge, 1997; Sashidharan & Francs, 1999; Sashidharan, 2001; Chakraborty & McKenzie, 2002). Excess detention of Black and minority ethnic patients is not only a clinically important issue, it is also politically charged and ethically contentious, requiring a cautious and balanced approach to research and interpretation of data.

This review confirms earlier findings of an excess of compulsory detentions among Black and minority ethnic patients (Churchill et al, 1999; Bhuie et al, 2003; Morgan et al, 2004). However, our findings go further in identifying variations in detention rates between different minority groups, and also reveal differences between first and later episode illness, and between civil and forensic patients, publication dates and research quality ratings. The finding that studies rated as high quality (a rating that included an assessment of degree of control of possible confounders) tended to report a reduced excess of detentions supports the hypothesis that at least some of the excess is accounted for by confounding variables. The reasons for differences between minority ethnic groups remain unexplored and warrant further scrutiny as to whether these are related to socio-economic, cultural or help-seeking differences between groups, or different experiences and perception of racism. Our finding that forensic detention rates for BME White and Black White comparisons were lower than the rates for civil detentions was unexpected, given previous results of the overrepresentation of BME patients in secure psychiatric care (Lelliott et al, 2001). However, meta-analysis results should be interpreted with caution as only two datasets were included for the forensic sections.

The increasing detention rate across time, with lower rates for first-episode illness, suggests that the relationship between Black and minority ethnic patients and mental health services deteriorates over time. Parkman et al (1997) found that although Black and minority ethnic patients had decreasing satisfaction with each hospital admission, whether the admissions were compulsory or not did not have an independent effect on patient satisfaction. The relationship between engagement, satisfaction and detention needs to be further explored in order to identify both general concerns and those specific to Black and minority ethnic groups, using longitudinal, mixed-methods studies exploring the process and experience of care and detention over time.

We found that racism and racial stereotyping of Black and minority ethnic patients were the most common explanations offered for excess detentions, but without primary supportive evidence to justify these assertions. The second most common explanation was that these patients are alienated, mistrust mental health services and are dissatisfied with services. This also had little supporting evidence from the papers itself. Overall, few studies were hypothesis-driven or methodologically based on a testable theoretical or conceptual model. Even where ethnic differences were found, there was a disjunction between reported findings and proposed explanations, with no attempt to link or explore complex multidimensional interactions between variables.

One possible reason why explanations such as racism have become accepted as the ‘cause’ of excess detention is that authors of early papers that reported excess detentions speculated on several possible explanations for this new finding. Instead of robustly testing these hypotheses, subsequent research has presented these speculations as ‘evidence from previous research’. Although this often happens in scientific research, in politically sensitive and emotionally charged areas such as detention and ethnicity it is critical to distinguish fact from opinion and hypothesis from evidence. Racial discrimination undoubtedly occurs in British society and leads to much personal suffering and possibly also to mental illnesses (Bhuie, 2002; Karlsen & Nazroo, 2002). Racism may indeed play a part in ethnic inequalities in mental health care, but this needs to be scientifically explored rather than accepted as the only cause of such differences (Singh & Burns, 2006).

Inclusion of publication dates in meta-analyses for the BME White and Black White comparisons shows a reduction in the excess of detention rate with later publication date. This can be interpreted in two ways. Either the excess rates for Black and minority ethnic patients have reduced over time, or with better control of confounders in later studies the effect of ethnicity is partly accounted for by confounding variables.

There is also an important issue of possible publication bias, in which research reporting significant differences between groups is more likely to be published, be cited by other authors and to produce multiple publications than research not finding such differences. The former studies are therefore more likely to be identified in systematic reviews, which potentially leads to bias (Sterne et al, 2001; Dubben & Beck-Bornholdt, 2005). It was noteworthy here that some studies not finding differences in detention rates did not attempt to explain this finding (Holloway et al, 1988; King et al, 1994; Harrison et al, 1999; Riordan et al, 2004), although this was in
contradiction to much of the available literature. This suggests that statistically non-significant differences are perceived as less worthy of comment. Presumably, reporting and commenting on an absence of difference in rates was even less likely among authors whose main focus was not ethnicity and the Mental Health Act. This would mean their findings might not have been reported and therefore not included in this review and meta-analyses.

Internationally there is nearly twenty-fold variation in detention rates across Europe, with rates rising in England, Austria and The Netherlands (Zinkler & Pribe, 2002; Salize & Dressing, 2004). In The Netherlands immigrants from Morocco, Surinam and the Dutch Antilles have among the highest rates of psychiatric detention, but this excess is accounted for by the presence of more severe symptoms, risk behaviours, lack of treatment motivation and poor functioning in these groups (Mulder et al, 2006). Although there is no major difference in the attitudes of mental health workers and society with regard to the compulsory detention of people with mental illness across several European countries (Lepping et al, 2004; Steinert et al, 2005), it has been suggested that in England the mass-media-generated public concern about the dangers posed by the mentally ill, along with the high level of personal responsibility that psychiatrists are expected to carry, may influence decision-making and increase the tendency for detention even in the first episode of mental illness operate before presentation to mental health services. Hence, any potential solutions must go beyond the health sector and involve statutory as well as voluntary and community agencies.

**ACKNOWLEDGEMENTS**

The authors are grateful to Hugh McGuire for help with the literature searches, Liz Lockhard for help checking the earlier database, Professor Tom Burns for very helpful comments on the study and to the Department of Health for funding the review.

**REFERENCES**


*Studies that were part of the meta-analysis.

SWARAN P. SINGH, Health Sciences Research Institute, Warwick Medical School, University of Warwick, Coventry; NANN GREENWOOD, SARAH WHITE, Department of Mental Health, St George's University of London; RACHEL CHURCHILL, Institute of Psychiatry, London, UK

Correspondence: Swaran P. Singh, Health Sciences Research Institute, Warwick Medical School, Coventry CV4 7AL, UK. Email: S.P.Singh@warwick.ac.uk

(First received 25 August 2006, final revision 30 January 2007, accepted 21 February 2007)
Relationship between daily suicide counts and temperature in England and Wales

LISA A. PAGE, SHAKOOR HAJAT and R. SARI KOVATS

Background Seasonal fluctuation in suicide has been observed in many populations. High temperature may contribute to this, but the effect of short-term fluctuations in temperature on suicide rates has not been studied.

Aims To assess the relationship between daily temperature and daily suicide counts in England and Wales between 1 January 1993 and 31 December 2003 and to establish whether heatwaves are associated with increased mortality from suicide.

Method Time-series regression analysis was used to explore and quantify the relationship between daily suicide counts and daily temperature. The impact of two heatwaves on suicide was estimated.

Results No spring or summer peak in suicide was found. Above 18°C, each 1°C increase in mean temperature was associated with a 3.8 and 5.0% rise in suicide and violent suicide respectively. Suicide increased by 46.9% during the 1995 heatwave, whereas no change was seen during the 2003 heat wave.

Conclusions There is increased risk of suicide during hot weather.

Declaration of interest None.

For over a century seasonal fluctuations in suicide have been observed (Morselli, 1881; Durkheim, 1951), with studies usually showing an increase in the incidence of suicide in spring and early summer (Chew & McCleary, 1995; Preti, 2000; Petridou et al., 2002). Suicide by violent means has shown stronger seasonal fluctuation than non-violent suicide (Maes et al., 1993; Hakko et al., 1998; Preti & Miotto, 1998; Rasanen et al., 2002; Rock et al., 2003). Various meteorological factors have been suggested to be aetiological important in this seasonality, including hours of sunshine (Souetre et al., 1987; Maes et al., 1994; Salib & Gray, 1997; Preti & Miotto, 1998; Petridou et al., 2002), low humidity (Salib & Gray, 1997) and temperature (Souetre et al., 1987; Linkowski et al., 1992; Maes et al., 1994; Preti & Miotto, 1998). Several European studies have examined the possible role of temperature in suicide by looking at the relationship between monthly or weekly suicide rates and temperature (Souetre et al., 1987; Linkowski et al., 1992; Maes et al., 1994; Jessen et al., 1998; Preti & Miotto, 1998); none of these studies used a daily measure of temperature, despite daily temperature being a methodologically preferable measure of exposure because the health effects of excessive heat are likely to be immediate (Basu & Samet, 2002).

More generally there is evidence of a U-shaped relationship between mean temperature and all-cause mortality (Curriero et al., 2002), with (for example) minimum all-cause mortality in London occurring at 19°C (Hajat et al., 2002). During extremes of high temperature (i.e. heatwaves) increases in daily mortality have been clearly documented (Kovats & Eli-Kristie, 2006), but this effect has not been specifically investigated for suicide before. Our aims were first to assess the relationship between daily suicide counts and daily temperature, and second to assess the impact of specific heatwave episodes on suicide.

METHOD

Data

The study used an ecological design. Individual exposure to ambient temperature was not known for those who died by suicide; therefore temperature was assigned at the group level for each day of the period of study. The main statistical approach was time-series regression, which is a technique used for the analysis of longitudinal data that exhibit trends and/or cyclical patterns. Since correlations between suicides and temperature were assessed on a short-term (day-to-day) basis, neither the prevalence of baseline risk factors nor a population denominator needed to be known in order to perform time-series analysis, although time-varying confounders had to be accounted for (Schwartz et al., 1996).

The UK Office for National Statistics (ONS) provided mortality data on all deaths considered due to suicide between 1 January 1993 and 31 December 2003 in England and Wales. All deaths that occurred before 31 January 2000 were coded using ICD–9 codes (World Health Organization, 1978); thereafter ICD–10 codes were used (World Health Organization, 1992). Deaths assigned to the following codes were included: ICD–9 codes E950.0–E959.0, E980.0–E989.0 (excluding E988.8); ICD–10 codes X60–X84, Y10–Y34 (excluding Y33.9). Deaths recorded as being due to 'undetermined intent' were included in the study. For each death information was available on date of death, gender and method. Suicides were classified as ‘violent’ or ‘non-violent’ depending on the method used and in line with previous studies – see, for example, Maes et al. (1994). Self-poisoning by solid, liquid or gas were classed as ‘non-violent’ suicide and all other methods were considered ‘violent’. A small number (n=11) of ICD–9 coded deaths were classified as being due to ‘late effects’ of either self-inflicted injury or assault; these were discarded from the analysis as date of death might not have been close in time to the date of injury.

The Met Office (http://www.metoffice.gov.uk) provided central England temperature data for the period under study; these include daily maximum and minimum temperatures collated from readings at four monitoring stations in central England (Parker et al., 1992). Daily mean temperature was calculated from the average of the maximum and minimum temperatures. Duration of daylight for each day of the...
year was obtained for London from the website http://www.sunrisesunset.com.

Analysis strategy
All analyses were performed using Stata version 8.2 for Windows. The data were collapsed into daily counts of suicide, violent suicide and non-violent suicide for each date in the series. Basic description of the data was undertaken prior to graphical examination of suicide trends by season and over the entire time period.

Generalised linear modelling
The relationship between death counts and mean temperature was explored using Poisson generalised linear modelling. The initial strategy was to include potential time-varying confounders in a model that excluded temperature. Confounders were identified as follows:
(a) year of death was included in the model to control for any overall trend in suicide between 1993 and 2003;
(b) month of death was included to control for the effect of seasonality (a monthly term was felt to best represent any slow time-varying drivers of suicide seasonality);
(c) indicator variables for day of the week, Christmas and other public holidays were also included in the model, as these have previously been associated with suicide (Lester, 1979; Johnson et al, 2005a); ‘Christmas’ included 25 and 26 December for all years plus any public holidays that followed these dates, and ‘other public holidays’ included New Year’s Day, Good Friday, Easter Sunday, Easter Monday and all Bank Holiday Mondays for all years;
(d) length of daylight, which is correlated with daily temperature, was included because it has been suggested to be a causative factor in the high rates of suicide observed in May, June and July (Maes et al, 1994; Salib & Gray, 1997; Petridou et al, 2002).

Graphs showing deviance residuals between the model residuals and daily lag were produced to allow for assessment (and correction if needed) of autocorrelation within the data. Autocorrelation refers to consecutive days having non-independent suicide counts. An overdispersion parameter was calculated from the model by dividing the variance by the mean; overdispersion is taken to be absent or minimal if the parameter approaches 1.

Mean temperature was subsequently added into the model to estimate the effect of temperature on suicide counts after adjustment for all known time-varying confounders. Natural cubic splines were first used to assess visually the functional form of the adjusted relationship, thereby identifying whether the relationship was likely to be linear or not across the full range of temperatures. Natural cubic splines set ‘knots’ at regular intervals along the exposure variable (temperature) allowing the adjusted relationship between temperature and death counts to be assessed for each inter-knot segment. Following visual assessment of the relationship, linear terms could then be used for quantification of the effect for each inter-knot segment. The main outcomes of interest were all suicides, violent suicides and non-violent suicides. Secondary outcomes of interest were male and female suicides. Men are more likely to use violent methods of suicide (Maes et al, 1993), so it was predicted that male suicides would demonstrate a different temperature effect to female suicides. All age groups were considered together in order to retain power.

Episode analysis
An ‘episode analysis’ was undertaken to assess the effect of two separate heatwaves on daily suicide counts. Our hypothesis was that if high ambient temperatures were associated with higher death counts, then sustained periods of unusually high ambient temperature (i.e. heatwaves) would result in higher counts than is usual for that time of year. All suicide cases (rather than suicide subgroups) were used for the episode analysis because of the relatively few number of deaths on individual days. The two heatwave periods in the data-set were 30 July to 3 August 1995 (Rooney et al, 1998) and 4 August to 13 August 2003 (Johnson et al, 2005b); these periods were defined as starting when the maximum central England temperature surpassed average values by 8 °C and ending when temperatures returned to average (Met Office data). Expected mortality from suicide for the same calendar periods was calculated by averaging the counts from the nearest four years in the data-set for the 1995 heatwave (1993, 1994, 1996 and 1997) and for the 2 years prior to the 2003 heatwave (2001 and 2002). In view of the declining number of suicides after 2001, it was judged that to use years that preceded 2001 in the episode analysis of the 2003 heatwave might risk underestimating any effect of the heatwave. A 7-day moving average count was used to mitigate against a ‘day of week’ effect. Percentage excess mortality was then calculated as follows, with calculation of confidence intervals obtained by assuming a Poisson distribution.

RESULTS
Between 1 January 1993 and 31 December 2003 there were 53,623 deaths by suicide in England and Wales. There was missing information on date of death for 163 (0.3%) suicides; these deaths were not randomly spread over the years, with 98% occurring in the years 1993 and 1994 and the remainder occurring in 2000 and 2001. The reasons for these missing data are not known and these cases were dropped from the analysis. The mean number of suicides per day was 13.3 (s.d.=3.9). Three-quarters of all suicides were by men and this proportion remained constant over the study period. The proportion of violent suicides increased during the study period, with a Mantel–Haenszel test for trend indicating that the odds ratio of dying from a violent (v. non-violent) suicide was 1.07 (95% CI 1.06–1.07; P<0.001) for each progressive year.

Temporal trends
Graphical display of the data showed that the number of suicides per year was relatively consistent between 1993 and 2000, albeit with a gradual rise in violent suicides. Between 2001 and 2003 there appeared to be a reduction in yearly numbers of both violent and non-violent suicide and this is in line with recent ONS findings (Office for National Statistics, 2005). The highest monthly number of suicides (both violent and non-violent suicides) took place in January. This finding remained after monthly counts were adjusted for number of days in the month. Examining the data using a time-series plot averaged across 1 year shows that daily suicide counts remained fairly consistent until the beginning of November, when counts started to diminish (Fig. 1). Similar plots were observed for violent and non-violent suicide (data not shown).

The highest daily suicide count was recorded on 1 January (mean=17.6 deaths;
95% CI 13.6–21.7) and there was strong evidence that this was higher than the mean daily count for the rest of January ($t$-test $= 2.69; P=0.008$). In order to assess whether January 1 suicides were driving the data, deaths on this date were excluded from the data-set; however, the adjusted count for January remained the highest, with 424.4 deaths. There was good evidence that fewer suicides occurred during the Christmas period when compared with the rest of the year: relative risk 0.86 (95% CI 0.77–0.97; Wald test $= -2.39, P=0.017$). The evidence was weaker that public holidays incurred a lower risk: RR days incurred a lower risk: RR $= 0.94$ (95% CI 0.89–1.00, $P=0.058$). The largest number of suicides took place on Mondays, with numbers declining as the week wore on. A similar pattern was observed for violent and non-violent suicides.

**Regression analysis**

The use of natural cubic splines allowed the adjusted relationship between suicide counts and temperature to be modelled across the range of temperatures. A natural cubic spline model offered a better fit to the data than a linear term (likelihood ratio test $= 46.21, P < 0.0001$), implying that the adjusted relationship between temperature and suicide was significantly non-linear across the whole range of temperature. Spline functions demonstrated subtly different profiles for all suicide, violent suicide and non-violent suicide; however, in all three groups there was evidence of increasing risk at higher temperatures (Fig. 2). For all suicides and violent suicides there was visual evidence of a high temperature threshold effect at approximately $18^\circ$C, i.e. a ‘hockey stick’ plot was seen, with the gradient of the line becoming steeper above this temperature. A similar temperature threshold could be seen for non-violent suicide, although the estimate was less precise (as indicated by wider confidence intervals) and the overall gradient was flatter. When deaths recorded as being due to ‘undetermined intent’ were excluded from the analysis, the estimate did not substantially change. There was no evidence of significant autocorrelation within the data (when deviance residuals were plotted, autocorrelations were $< 0.05$ within a 7-day lag), therefore no control for autocorrelation was made in subsequent analyses. Overdispersion was approximately 1.06 for all the suicide models, which was considered adequate without further correction. Modeling an interaction between daylight and mean temperature failed to give a better fit to the data (likelihood ratio test $= 1.68, P=0.640$), indicating that there was no statistical interaction between these two variables.

The aim of the final Poisson generalised linear model was to control for all known time-varying confounders and calculate an adjusted relative risk (and hence percentage change) for suicide in relation to temperature. The focus of the study was the relationship between high temperature and suicide; therefore the final models assessed the linear effect of mean temperatures above the visually derived threshold of $18^\circ$C (with temperatures below $18^\circ$C set to zero). The final models showed that there was strong evidence for a small but significant effect of increasing temperature on all suicides and violent suicide. There was much weaker evidence for an effect of temperature on non-violent suicide (Table 1).

The effect of high ambient temperature on male and female suicide counts was estimated in a secondary analysis. The natural cubic spline plots (Fig. 2) show that the relationship between temperature and suicide is similar for male suicide and all suicides. The plot for female suicide appears to demonstrate a shallower curve than that for male suicide, with a steady positive gradient from approximately $15^\circ$C. The percentage increase in suicide counts above $18^\circ$C was similar for both genders, although the estimate was less precise for women.

**Episode analysis**

There were mixed results from the analysis of the heatwave events. The 1995 heatwave was associated with a marked short-term increase in mortality from suicide, whereas the 2003 heatwave was associated with virtually no change in mortality (Table 2, Fig. 3). In a post hoc attempt to explore the possible reasons for this lack of effect during the 2003 heatwave, the month prior to the 2003 heatwave was examined graphically. This demonstrated that there might have been an increase in mortality during an earlier hot spell between 13 July and 17 July, when mean temperatures increased to over $19^\circ$C for four consecutive days (Fig. 3). This brief heatwave in July 2003 might have resulted in some immediate increase in mortality, as reported in Table 2, although the confidence intervals were wide and encompassed zero.

**DISCUSSION**

This study used time-series regression to assess deaths from suicide over an 11-year period in England and Wales. After appropriate adjustment for time-varying confounders, there was evidence of a small but robust effect of temperature on suicide and violent suicide counts. The effect of temperature on suicide counts was demonstrated to be non-linear across the temperature range under consideration. High temperature thresholds were identified relatively easily for suicide and violent suicide, although this was less clear for non-violent suicide. When quantifying the effect of high ambient temperatures (mean $> 18^\circ$C), there was evidence of an increase in the relative risk of suicide and violent suicide for each $1^\circ$C rise in temperature. Secondary analyses revealed no clear difference in the relative risk of high temperature on male or female suicide. Episode analysis of two
heatwaves found that the 1995 heatwave was associated with a clear excess of suicides, whereas the 2003 heatwave produced little obvious change to the number of suicides expected for the time of year.

A large data-set was used for this study, which covered all deaths registered as being due to suicide in England and Wales between 1993 and 2003. The mortality and meteorological data were gathered as part of routine surveillance work by ONS and the Met Office and were unlikely to have been influenced by observer bias. The high quality of data provided by the death registration process in the UK meant that it was feasible to include ‘undetermined intent’ deaths as suicides in the study, which maximised the power of the study to detect an effect. Time-series analysis is a powerful technique with which to explore longitudinal data such as these, with the major advantage that it is not necessary to know the population denominator or the distribution of known (or unknown) individual risk factors within the population to interpret the findings. The ecological design of the study means that it is impossible to ascribe heat-related mechanisms to deaths occurring at times of high ambient heat, because individual exposure to high temperature was not known. Nevertheless, during hot periods it is highly likely that those who died were, like the majority of the population, exposed to high temperatures.

The scarcity of air-conditioning in the UK means it is unlikely that many of those who died were able to keep themselves cool at times of great heat.

Lack of evidence for seasonality

The results of our study did not support the finding of a spring or summer peak in deaths from suicide; suicide occurred most frequently in January. Neither was a seasonal effect seen for violent suicide, which has been identified as the subtype of suicide most likely to be associated with season (Maes et al., 1993; Hakko et al., 1998; Preti & Mirotto, 1998; Rasanen et al., 2002; Rock et al., 2003). Although the majority of previous European studies have

**Fig. 2** Relationship between temperature and suicide (1993–2003): mean temperature v. relative risk of (a) suicide, (b) violent suicide, (c) non-violent suicide, (d) male suicide and (e) female suicide. Broken lines indicate upper and lower limits of confidence interval. Natural cubic spline model, adjusted for year, month, day of week, Christmas, public holidays and hours of daylight.
shown a spring or summer peak in suicide, studies from the UK in the recent past have not. In particular, two studies that used ONS data for England and Wales between 1982 and 1999 found no—or very little—evidence of a seasonal effect (Yip et al., 2000; Simkin et al., 2003).

Harmonic analysis, an alternative time-series technique, has been shown to increase the likelihood of finding a seasonal effect for suicide (Hakko et al., 2002). However, harmonic analyses were used by Yip et al. (2000) and Simkin et al. (2003), who both failed to show any significant seasonality within recent UK data. It therefore seems unlikely that the choice of analysis technique has influenced the findings in this case. One explanation is that the effect of seasonality has become less important in recent decades in some countries (Yip et al., 2000; Ajdacic-Gross et al., 2005). Chew & McCleary (1995) looked at seasonality and suicide in 28 countries and found that countries with high levels of industrialisation and low numbers involved in agricultural work showed the least seasonality.

**Effect of temperature on suicide counts**

Despite no evidence of a seasonal effect on suicide, it was possible to show that temperature has a short-term effect on suicide counts. By modelling the relationship of temperature to death counts using natural cubic splines, a high temperature threshold was determined above which the effect of temperature increased linearly. This technique has not been used to study deaths from suicide before, although when used in other contexts has shown similar high temperature effects for hospital admissions for renal and respiratory problems (Kovats et al., 2004) and total mortality (Hajat et al., 2002). The relationship between suicide and mean temperature demonstrates a similar threshold to that of total mortality and temperature, which has been estimated to occur at about 19°C (Hajat et al., 2002). This is the first time that death from suicide has been shown to be contributing to the known increase in all-cause mortality at higher temperatures.

The effect of high temperature on death counts was seen for all subtypes of suicide except non-violent suicide, for which the evidence was weaker. In real terms this finding is likely to be important, as the relative risk estimates for each degree of temperature above 18°C indicate an increase in suicide and violent suicide of 3.8 and 5.0% respectively. It is not infrequent for the mean temperature in England to be above 18°C, with such values being recorded on 222 days over the 11-year period of the study. It is unlikely that deaths from suicide are merely being brought forward in time (or ‘harvested’) during hot weather—real additional suicides probably occur when temperatures are high. Future research should focus on whether other important subgroups, such as the elderly, are disproportionately affected by suicide in hot weather.

The 1995 heatwave resulted in a clear excess of suicide during the period of hot weather: the increase of 41.5% is well in excess of the 10.8% increase in all-cause mortality reported for the same period in London (Kovats et al., 2004). The absence of effect during the 2003 heatwave is therefore surprising and an explanation is required. Heatwaves that occur in early summer have been found to result in greater all-cause mortality than those that occur later in the year, implying that people most vulnerable to heat-related death may die during early periods of high temperature and/or that some adaptation to high temperature can occur (Basu & Samet, 2002). It is possible that both of these mechanisms were important during the 2003 heatwave. The earlier period of hot weather between 13 July and 17 July 2003 might have resulted in some excess deaths in those most susceptible to heat-related suicide, while also allowing some physiological or behavioural adaptation among other vulnerable individuals. The later heatwave in August 2003 may therefore have resulted in fewer deaths than would otherwise have occurred. This implies that a sudden increase in temperature may result in greater mortality from suicide than a gradual and sustained increase.

**Study limitations**

Because of the ecological design of the study it was impossible to link individual (or community) characteristics to heat-related suicide. Additional limitations of the study were potential misclassifications of outcome and exposure, which could have led to some bias in the results. First, deaths from the whole of England and Wales were used in the study and yet the temperature measure (the exposure of interest) was taken from four monitoring stations in central England. Second, it was not known exactly when the suicidal act took place, as only date of death was available for analysis. Third, it was assumed that there was no systematic difference between deaths designated as ‘undetermined

### Table 1  Adjusted percentage increase in suicide at temperatures above 18°C (lag = 0)

<table>
<thead>
<tr>
<th>Suicide type</th>
<th>Mean increase</th>
<th>Wald test P</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>All suicide</td>
<td>13.3 (3.9)</td>
<td>4.27</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Violent suicide</td>
<td>8.0 (3.0)</td>
<td>4.38</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Non-violent suicide</td>
<td>5.4 (2.6)</td>
<td>1.45</td>
<td>0.147</td>
</tr>
<tr>
<td>Male suicide</td>
<td>10.4 (3.3)</td>
<td>3.79</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Female suicide</td>
<td>3.5 (1.9)</td>
<td>2.02</td>
<td>0.044</td>
</tr>
</tbody>
</table>

1. Mean across entire temperature range.
2. Adjusted for year, month, day of week, Christmas, public holidays and hours of daylight.

### Table 2  Effect of heat waves on suicide (using 7-day moving average)

<table>
<thead>
<tr>
<th>Date of heat wave</th>
<th>Observed suicides</th>
<th>Estimated excess/deficit of suicides</th>
<th>Change in suicide from expected</th>
</tr>
</thead>
<tbody>
<tr>
<td>29 July to 3 August 1995</td>
<td>103.6</td>
<td>+33.1</td>
<td>+46.9 (15.6 to 86.8)</td>
</tr>
<tr>
<td>4–13 August 2003</td>
<td>118.9</td>
<td>−2.1</td>
<td>−1.8 (−17.8 to 18.4)</td>
</tr>
<tr>
<td>13–17 July 2003</td>
<td>56.6</td>
<td>+7.6</td>
<td>+15.5 (−12.7 to 56.1)</td>
</tr>
</tbody>
</table>

2. Comparison years for 2003 were 2001 and 2002.
Suicide counts during two heatwaves: (a) July and August 1995; (b) July and August 2003. Graphs show daily suicide counts and temperature, using a 7-day moving average filter.

**Implications**

Suicide is an important, preventable cause of premature mortality. This study has shown an effect of high temperatures on suicide counts that will probably become more important as global warming continues (Patz et al., 2005). It is possible that the population of England and Wales will adjust to higher ambient temperatures, although the speed of global warming may be too great for adaptation to occur. Those with mental illness are highlighted as an at-risk group in England’s heatwave plan (Department of Health, 2005), although this is because of their increased susceptibility to heat stroke (Bark, 1998) rather than for suicide prevention. In any case, it remains to be seen whether public health measures (designed after the 2003 heatwave) to prevent heat-related death have been effective or not (Kovats & Ebi-Kristie, 2006).

Replication of these findings is required in other populations and geographical regions. Seasonal effects of suicide have been shown most consistently in northern European countries and it is not known whether there is an effect of temperature on suicide in other regions (e.g., equatorial European countries) where high temperatures are more common. If a consistent association between short-term high temperature and suicide is identified, further attention needs to be paid to the mechanisms that underlie this effect.

**ACKNOWLEDGEMENTS**

The authors thank Allan Baker at the Office for National Statistics for supplying the mortality data. L.A.P. is supported by the National Institute of Environmental Health Sciences (NEIHS), National Institutes of Health as a Ruth L. Kirschstein National Research Fellow (F32 ES013690). The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the NEIHS. S.H. is supported by a Welcome Trust Research Career Development Fellowship. R.S.K. is funded by the European Commission Directorate-General for Health and Consumer Protection for the EuroHEAT project (agreement no. 2004322).

**REFERENCES**


Cortical white-matter microstructure in schizophrenia

Diffusion imaging study


Background  Several, although not all, of the previous small diffusion-weighted imaging (DWI) studies have shown cortical white-matter disruption in schizophrenia.

Aims  To investigate cortical white-matter microstructure with DWI in a large community-based sample of people with schizophrenia.

Method  Sixty-eight people with schizophrenia and 64 healthy controls underwent a session of DWI to obtain the apparent diffusion coefficient (ADC) of white-matter water molecules. Regions of interest were placed in cortical lobes.

Results  Compared with controls, the schizophrenia group had significantly greater ADCs in frontal, temporal and occipital white matter (analysis of covariance, P < 0.05).

Conclusions  Our findings confirm the presence of cortical white-matter microstructure disruption in frontal and temporo-occipital lobes in the largest sample of people with schizophrenia thus far studied with this technique. Future brain imaging studies, together with genetic investigations, should further explore white-matter integrity and genes encoding myelin-related protein expression in people with first-episode schizophrenia and those at high risk of developing the disorder.

Declaration of interest  None.

Diffusion-weighted imaging (DWI) is a relatively new technique capable of examining molecular water mobility in brain tissue by providing the apparent diffusion coefficient (ADC) of water molecules (Taylor et al., 2004), particularly in white matter, a highly organised tissue where water diffusion is restricted. The ADC is the critical measure for a detailed investigation of white-matter integrity and inferences can be drawn from it on white-matter microstructure, organisation and cytoarchitecture, which cannot be visualised using conventional magnetic resonance imaging (Basser, 2002). When brain tissue is disrupted, such as in neurological disorders involving white matter (for example multiple sclerosis), the ADC is abnormally increased (Nusbaum et al., 2000; Rovaris et al., 2002). Recently DWI has been used to explore white matter in schizophrenia, since this tissue has been suggested to have a major role in the pathophysiology of this disorder (Keshavan, 1999; Keshavan et al., 2005). Indeed, white-matter changes may alter intra-hemispheric connectivity and functional brain lateralisation in people with schizophrenia (Falkai et al., 1995; Delisi et al., 1997; Crow, 1998; Brambilla et al., 2005), potentially sustaining cognitive deficits. Several DWI studies conducted in recent years have consistently shown cortical white-matter disruptions (Taylor et al., 2004), although not all investigations have done so (Steel et al., 2001; Foong et al., 2002; see Table DS1 to the online version of this paper). However, previous diffusion imaging reports were limited by small sample sizes.

We used DWI to investigate cortical white-matter microstructure in a large community-based sample of patients with schizophrenia recruited from the geographically defined catchment area of South Verona in Italy. Our hypothesis, based on previously published findings of disrupted white-matter integrity in schizophrenia, was that people with schizophrenia would have increased ADC values.

METHOD

Sample

Sixty-eight people with a DSM-IV diagnosis of schizophrenia (American Psychiatric Association, 1994) were studied (Table 1). They were recruited from the geographically defined catchment area of South Verona (100,000 inhabitants) and treated by the South Verona community-based mental health service and by other clinics reporting to the South Verona Psychiatric Care Register (Amaddeo et al., 1997; Tansella & Burri, 2003). Diagnoses of schizophrenia were obtained using the Item Group Checklist of the Schedule for Clinical Assessment in Neuropsychiatry (IGC-SCAN; World Health Organization, 1992) and confirmed with the clinical consensus of two staff psychiatrists. The IGC-SCAN assessments were completed by two trained research clinical psychologists (C.P., L.P.) with extensive experience in using the SCAN instrument. They completed at least ten IGC-SCAN ratings with a senior investigator trained in SCAN assessment, after having conducted several IGC-SCAN assessments. Successively, reliability was checked in a further ten assessments with the senior investigator, masked to the results. Similar diagnoses were obtained for at least eight out of ten IGC-SCAN assessments. Moreover, the psychopathological item groups completed by the two raters were compared in order to discuss any major symptom discrepancies. In addition, we regularly assured reliability of the IGC-SCAN diagnoses by holding consensus meetings with treating psychiatrists and a senior investigator. It is noteworthy that the Italian version of the SCAN was edited by our group (World Health Organization, 1996) and that our investigators attended specific training courses held by an official trainer in order to learn how to administer the IGC-SCAN. Subsequently, diagnoses of schizophrenia were corroborated with the clinical consensus of two staff psychiatrists, according to DSM-IV criteria. Patients with a comorbid psychiatric disorder, alcohol or substance misuse within the 6 months preceding the study, a history of traumatic head injury with loss of consciousness, or epilepsy or other neurological diseases were excluded. All but two patients were receiving antipsychotic medication at the time of imaging. Specifically, 22 patients were taking typical antipsychotic drugs (13 haloperidol, 3 chlorpromazine, 2 fluphenazine,
2 clonazepam, 1 thioridazine, 1 zuclopenthixol) and 44 on atypical antipsychotic medication (25 on olanzapine, 9 on clozapine, 8 on risperidone, 2 on quetiapine). Patients’ clinical information was retrieved from psychiatric interviews, the attending psychiatrist and medical charts. Clinical symptoms were characterised using the 24-item Brief Psychiatric Rating Scale (BPRS; Ventura et al., 2000), which was administered by two trained research clinical psychologists (C.P., L.P.). The reliability of the BPRS ratings was established and monitored using similar procedures to those used for the IGC–SCAN.

Sixty-four people were recruited to constitute a healthy control group (Table 1). They had no DSM–IV Axis I disorder, as determined by an interview modified from the Structured Clinical Interview – DSM–IV Axis I Disorders, non-patient version (Spitzer & Williams, 1988), no history of psychiatric disorder in a first-degree relative, no history of alcohol or substance misuse and no current major medical illness. Members of the control group were hospital or university staff volunteers or patients undergoing magnetic resonance imaging (MRI) for dizziness without evidence of central nervous system abnormalities on the scan, as reviewed by the neuroradiologist (R.C.); their dizziness was due to benign paroxysmal positional vertigo or to non-toxic labyrinthitis. Control group participants were scanned only after a full medical history and general neurological, otoscopic and physical examinations, and after they had completely recovered from their dizziness. None was taking any medication at the time of participation, including drugs for nausea or vertigo.

This research study was approved by the biomedical ethics committee of the Azienda Ospedaliera of Verona. All individuals provided signed informed consent, after having understood all issues involved in study participation.

### Table 1 Socio-demographic and clinical variables of the sample

<table>
<thead>
<tr>
<th></th>
<th>Control group (n = 64)</th>
<th>Schizophrenia group (n = 68)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years: mean (s.d.)</td>
<td>40.70 (11.16)</td>
<td>41.39 (11.68)</td>
</tr>
<tr>
<td>Males/females, n</td>
<td>34/30</td>
<td>39/29</td>
</tr>
<tr>
<td>Right-handed</td>
<td>60</td>
<td>64</td>
</tr>
<tr>
<td>Ethnicity, %</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Education, n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary or secondary school</td>
<td>22</td>
<td>51***</td>
</tr>
<tr>
<td>High school</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>First degree or professional school</td>
<td>27</td>
<td>2</td>
</tr>
</tbody>
</table>

### Clinical variables: mean (s.d.)

- **Age at onset, years**: 27.46 (9.48)
- **Length of illness, years**: 14.40 (11.12)
- **Number of hospitalisations**: 3.79 (6.09)
- **Lifetime antipsychotic treatment, years**: 12.83 (10.76)
- **BPRS score**
  - **Total**: 45.38 (16.96)
  - **Negative symptom score**: 9.08 (3.13)
  - **Positive symptom score**: 11.74 (6.68)

Brief Psychiatric Rating Scale.

***p = 0.001, **p < 0.01, *p < 0.05

---

### Image analyses

The apparent diffusion coefficients of water molecules for cortical white matter were detected by using software developed in-house written in MatLab version 7 (The Mathworks, Natick, Massachusetts, USA). The ADCs were obtained by placing, bilaterally, circular regions of interest in the frontal, temporal, parietal and occipital cortex on the non-diffusion-weighted (b=0) echoplanar images in reference to standard brain atlases (Jackson & Duncan, 1996; Patel & Friedman, 1997) and according to previous studies (Sun et al., 2003; Wolkin et al., 2003; Kumra et al., 2004; Kitamura et al., 2005; Fig. 1). The regions of interest were then automatically transferred to the corresponding maps to obtain the ADCs. The ADC maps were obtained from the diffusion images with $b=1000$, according to the equation $b_{ADC} = \ln[A(b)/A(0)]$, where $A(b)$ is the measured echo magnitude, $b$ is the measure of diffusion weighting and $A(0)$ is the echo magnitude without diffusion gradient applied (Basser, 2002). The resulting ADC was expressed in units of $10^{-3}$ mm²/s. A trained rater (N.A.), masked to group assignment and patient identity, measured all scans. The intraclass correlation coefficients, which were calculated by having two independent...
Circular regions of interest were placed in cortical white matter on the b = 0 echoplanar images, and then automatically transferred to apparent diffusion coefficient (ADC) maps (A, frontal lobes; B, temporal lobes; C, parietal lobes; D, occipital lobes).

Anatomical landmarks

Frontal cortex
Regions of interest were positioned in the axial slice at the level of the genu of corpus callosum (standardised at 43.5 mm²), then in the inferior slice (standardised at 43.5 mm²) and in the two superior slices (standardised at 84.4 mm²), posteriorly and medially to the frontal horns of the lateral ventricles.

Parietal cortex
Regions of interest (standardised at 84.4 mm²) were placed in the axial slice when the lateral ventricles first disappeared and in the superior slice, posteriorly to the postcentral sulcus.

Temporal cortex
Regions of interest (standardised at 43.5 mm²) were positioned in the axial slice at the level of the lateral fissure and in the inferior slice, posteriorly and laterally to the lateral fissure.

RESULTS

Occipital cortex
Regions of interest (standardised at 43.5 mm²) were placed in the two inferior axial slices where the occipital horns of the lateral ventricles become visible, posteriorly to the occipital horns.

Statistical analyses
All analyses were conducted using the Statistical Package for the Social Sciences version 11.0 for Windows and the two-tailed statistical significance level was set at P < 0.05. Analysis of covariance (ANCOVA) with age and gender as covariates was performed to compare white-matter ADCs between the schizophrenia group and the control group. Pearson’s correlation and partial correlation analyses controlled for age were used to examine possible association between age and clinical variables respectively, and ADC measures.

Compared with the control group, the participants with schizophrenia had significantly greater apparent diffusion coefficients for frontal, temporal and occipital white matter (Table 2), even when taking educational level into consideration (right and left frontal ADCs, P = 0.09, 0.12; right and left temporal ADCs, P = 0.006, 0.009; right and left occipital ADCs, P = 0.006, 0.002, respectively; ANCOVA, age, gender and educational level as covariates). Similar results were found when the schizophrenia group was compared separately with control participants recruited from hospital and university staff (n = 33) (left frontal ADCs, P = 0.14; temporal ADCs: P < 0.001, occipital ADCs, P < 0.003) and with control participants who had been treated for dizziness (n = 31) (right frontal ADCs, P = 0.07; temporal ADCs, P = 0.01; occipital ADCs, P ≤ 0.01) (ANCOVA; age and gender as covariates). Also, no significant difference for any ADC measure was found between the two control subgroups (ANCOVA; age and gender as covariates, P > 0.05).

The ADC measures were still greater in the schizophrenia group than in the combined control group when both groups were stratified by gender, both in men (left frontal ADCs, P = 0.04; temporal ADCs, P < 0.001, occipital ADCs, P < 0.002) and women (right temporal ADCs, P = 0.12; left
temporal ADCs, $P=0.03$; right occipital ADCs, $P=0.06$; left occipital ADCs, $P=0.01$) (Mann–Whitney U-test).

Age was significantly and directly correlated with left temporal ADC measures in the control group ($r=0.28, P=0.02$) but not in the schizophrenia group ($r=0.16, P=0.18$). No significant association was shown between age and other ADC values (Pearson’s correlation, $P>0.05$) or between clinical variables (age at onset, length of illness, number of hospitalisations, BPRS scores, antipsychotic lifetime treatment) and white matter ADCs (partial correlation controlled for age, $P>0.05$). Furthermore, no significant difference for any ADC value was observed between patients treated with typical antipsychotic drugs ($n=22$) and those treated with atypical antipsychotics ($n=44$) (Mann–Whitney U-test, $P>0.05$). Also, patients with severe illness (BPRS $>41$; $n=37$) did not differ significantly on any ADC measure compared with patients with mild-to-moderate illness (BPRS $\leq 41$; $n=31$) (Mann–Whitney U-test, $P>0.05$). A BPRS total score of 41 was chosen as the cut-off level for mild or moderate illness, indicated by Leucht et al (2005).

**DISCUSSION**

This study found widespread regional white-matter disruption in schizophrenia, as shown by higher ADCs in frontal, temporal and occipital lobes. To our knowledge, this is the largest study to show disrupted white-matter cytoarchitecture in schizophrenia (Kanaan et al, 2005). Consistently, impairments of cortical white-matter integrity have been found in people with schizophrenia by a number of prior small diffusion imaging studies (Kubicki et al, 2007, see online Table S1). Specifically, abnormally increased water diffusion coefficients or abnormally decreased fractional anisotropy have been found in at least ten prior investigations of frontal lobes (Buchsbaum et al, 1998; Ardekani et al, 2003; Minami et al, 2003; Kubicki et al, 2004; Wang et al, 2004; Kitamura et al, 2005; Kubicki et al, 2005a; Szeszko et al, 2005; Hao et al, 2006; Shin et al, 2006) and in tempo-occipital lobes (Lim et al, 1999; Agartz et al, 2001; Ardekani et al, 2003, 2005; Minami et al, 2003; Kubrick et al, 2005a; Szeszko et al, 2005; Hao et al, 2006; Shin et al, 2006). However, some studies report preserved integrity of white matter in schizophrenia (Steel et al, 2001; Foong et al, 2002; Kubicki et al, 2002). Both ADC and fractional anisotropy are considered as complementary indices of white-matter microstructure organisation, providing evidence of disruption when increased and decreased respectively (Taylor et al, 2004). In our study, we did not report fractional anisotropy because the diffusion tensor sequence was not collected. Specifically, the ADC image provides a relative presentation of the diffusion coefficient in each pixel within the image, where low and high intensity values indicate respectively low and high diffusion (Basser, 2002). Abnormalities in cortical white matter may lead to impaired connection, which may ultimately alter the speed, quantity and/or quality of intrahemispheric communication, relevant to cognitive disturbances reported in schizophrenia (Krabbendam et al, 2005). This may be a result of reduced axonal density or myelination. Indeed, oligodendrocytes, which have the potential to influence myelination and synaptic transmission, have been found to be functionally abnormal in schizophrenia (Hof et al, 2002; Davis et al, 2003; Bartozok & Altschuler, 2005). None of the less, several factors may contribute to explain increased water white-matter diffusion, such as less dense packing of fibres, disruption of internal axonal integrity (reduced intra-axonal microtubular density), reduced degree of myelination or variation in membrane permeability to water. However, since white-matter is mostly composed of myelinated axons, the density of axonal membranes and myelin seem to play the major part (Beaulieu & Allen, 1994; Giedd, 2004).

Several earlier diffusion imaging studies reported frontal, temporal and occipital white-matter alterations within regions of interest identified by visual inspection of the individual anatomy, as in our method (Steel et al, 2001; Hopman et al, 2002; Minami et al, 2003; Wolkin et al, 2003; Kumra et al, 2004; Kitamura et al, 2005). In particular, we examined the middle and inferior frontal white-matter regions, which have been shown to be functionally altered in schizophrenia (Shenton et al, 2001), potentially sustaining executive function deficits (MacDonald et al, 2005; Brambilla et al, 2007). Moreover, temporal regions of interest were positioned in the medial temporal white matter regions, which are involved in modulating language domain in humans and are likely to have a key role in language abnormalities in schizophrenia (Seidman et al, 2003; Antonova et al, 2004). Finally, the occipital regions of interest were placed in medial occipital areas, which have been shown to be altered in schizophrenia by other diffusion imaging studies (Lim et al, 1999; Agartz et al, 2001; Ardekani et al, 2003, 2005; Minami et al, 2003; Kumra et al, 2004; Kubicki et al, 2005a; Szeszko et al, 2005; Hao et al, 2006; Shin et al, 2006). Furthermore, abnormalities in early-stage visual processing in schizophrenia have recently been shown, possibly contributing to higher-level cognitive deficits (Butler et al, 2005; Schechter et al, 2005). Therefore, our findings suggest that frontal and temporo-occipital white-matter disruption may in part support cognitive and language deficits in schizophrenia.

Taken together, these brain imaging findings indicate that cortical white-matter microstructure is disrupted in schizophrenia. Moreover, these results may be supported...
by post-mortem studies showing a quantitative reduction in white matter cells (Ambariz et al., 1996; Uranova et al., 2004). In particular, reduced expression of myelin and oligodendrocyte-related genes and proteins has been shown in schizophrenia, suggesting oligodendrocyte dysfunction (Flynn et al., 2003; Hof et al., 2003; Tkachev et al., 2003; Chambers & Perrone-Bizzozero, 2004). Specifically, neuregulin 1 (NRG1), a candidate gene for schizophrenia (Stefansson et al., 2002; Tosato et al., 2005; Williams et al., 2005), has been shown to have a key role in oligodendrocyte development and proliferation (Marchionni et al., 1993; Vartanian et al., 1999; Liu et al., 2001). Therefore, altered expression of NRG1 or other myelination-related genes may potentially result in abnormal oligodendrocyte function or myelination in schizophrenia (Hakak et al., 2001; O'Donovan et al., 2003). However, it remains to be elucidated whether cortical white-matter impairment mostly reflects brain maldevelopment or neurodegeneration. In particular, it would be of great interest to understand how and when the white-matter disruption in schizophrenia relates to the physiological processes of white-matter maturation (Bartzokis, 2002; Hafner, 2004; Harrison, 2004; Bresnahan et al., 2005). Indeed, recent reports suggest that intracortical myelination increases during adulthood, reaching its peak during the fifth decade of life, particularly in the frontal and temporal lobes (Bartzokis et al., 2003), in a constant state of well-regulated structural and functional change. Affected myelination in schizophrenia, which may itself be due to multiple genetic and environmental factors, may contribute to alter this temporally expanded view of brain white-matter development from adolescence until middle age. As proposed by Bartzokis, this would ultimately result in dysregulation of the temporal synchronous development of widely distributed neural networks in schizophrenia, being manifested in the heterogeneity of symptoms and cognitive impairments (Bartzokis, 2002). Interestingly, white-matter alterations (particularly of corpus callosum) and abnormal down-regulation of oligodendrocyte and myelination genes have been demonstrated in bipolar affective disorder as well as in schizophrenia (Brambilla et al., 2003, 2004; Tkachev et al., 2003). This sustains the notion that the two disorders may have similar white-matter pathophysiological pathways. Future brain imaging studies together with genetic investigations should further explore white-matter integrity and genes encoding myelin-related protein expression in people with first-episode schizophrenia and possibly bipolar affective disorder, and in the populations at high risk of developing these disorders.

Interestingly, we found a significant direct correlation between age and left temporal ADC values in the control group which was not present in the schizophrenia group. This is consistent with a recent investigation showing in controls, but not in patients, a significant negative effect of age on the integrity of the left superior longitudinal fasciculus, which connects the frontal and temporal cortex (Jones et al., 2006). Also, age-related decline of cerebral white-matter coherence in humans, which may represent subtle structural white-matter changes with normal ageing, has been demonstrated by diffusion imaging studies (Engelter et al., 2000; Pfefferbaum et al., 2000; O’Sullivan et al., 2001; Sullivan et al., 2001). Thus, as a speculative interpretation, it is possible that the effects of physiological ageing on white matter cannot be seen in schizophrenia owing to the presence, since early adolescence, of abnormal neurodevelopment and cytoarchitectural organisation of cortical white matter, particularly in the temporal region (Pantelis et al., 2005).

No significant association between ADC values and any clinical variable was found in our study, consistent with several prior reports exploring correlations between diffusion measures and clinical features in schizophrenia (Steel et al., 2001; Kumra et al., 2004, 2005; Jones et al., 2005; Kubicki et al., 2005a; Kitamura et al., 2005; Szczko et al., 2005). This suggests that cortical white-matter disruption in schizophrenia is not a secondary effect of chronicity, medication or psychopathology but is potentially related to the core pathophysiology of the disease. However, it should be mentioned that two small studies have found increased white-matter alterations in people with schizophrenia with more severe negative symptoms in the right insula (Shin et al., 2006) and the inferior frontal region (Wolkin et al., 2003). However, the latter group also showed a relationship between impulsivity/aggression and altered white-matter microstructure in the right inferior frontal region and insula in men with schizophrenia (Hoptman et al., 2002, 2004). Therefore, the correlation between white-matter cytoarchitecture and clinical symptoms in schizophrenia is still controversial and needs further investigation in large samples.

It should be noted that our schizophrenia sample mostly comprised treated patients with chronic illness, thus it is not clear whether white-matter disruption preceded the onset of the illness or appeared subsequently as a result of illness course or psychotropic treatment. However, length of illness or antipsychotic lifetime administration did not significantly affect ADC values, suggesting that cortical white-matter abnormalities may not be related to illness or medication. Also, we recruited a relatively larger number of participants than prior diffusion imaging studies, with a good match between those in the schizophrenia and control groups, providing adequate power. Part of our control group was selected from individuals undergoing MRI scanning for dizziness, which may represent a methodological limitation. However, these participants were fully recovered at the time of scanning and had no evidence of central nervous system abnormalities on the scan. Finally, no particular white-matter tracts could be detected with our approach, such as the uncinate or the arcuate fasciculi which form specific temporoparieto-frontal connections (Burns et al., 2003; Kubicki et al., 2005b; Jones et al., 2006).

In conclusion, altered cortical white-matter microstructure in schizophrenia has been replicated in this large study, particularly in frontal and temporoparietal lobes. Hypothetically, abnormal myelination due to oligodendrocyte dysfunction might account for these findings. This might potentially affect intrahemispheric communication and ultimately lead to the cognitive disturbances seen in people with schizophrenia.

ACKNOWLEDGEMENTS

We thank Dr. Sarah Tosato, MD, for helpful comments on the earlier version of this manuscript. This work was partly supported by grants from the American Psychiatric Institute for Research and Education (APRE/AstraZeneca Young Minds in Psychiatry Award) and from the Italian Ministry for Education, University and Research (PRIN 2005068874) to PB.

REFERENCES


Liu, Y., Ford, B., Mann, M. A., et al. (2001) Neuregulins increase alpha7 nicotinic acetylcholine receptors and
enhance excitatory synaptic transmission in GABAergic interneurons of the hippocampus. Journal of Neuroscience. 21, 5660–5669.


Correspondence: Dr Paolo Brambilla, Dipartimento di Patologia e Medicina Clinica e Sperimentale, Cattedra di Psichiatria, Policlinico Universitario, Via Colunga 50, 33100 Udine, Italy. Tel: +39 0432 55 9494; fax: +39 0432 54 5526; email: paolo.brambilla@uniud.it

(First received 21 December 2005, final revision 12 January 2007, accepted 24 January 2007)

N. ANDREONE, MD, M. TANSELLA, MD, Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology, University of Verona, and Verona–Udine Brain Imaging and Neuropsychology Programme, Inter-University Centre for Behavioural Neurosciences, University of Verona; R. CERINI, MD, Department of Morphological and Biomedical Sciences, Section of Radiology, GB Rossi Hospital, University of Verona; A. VERSACE, MD, G. RAMBOLDI, MS, C. PERLINI, PsychD, N. DUSI, MD, L. PELIZZA, PsychD, Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology, University of Verona, and Verona–Udine Brain Imaging and Neuropsychology Programme, Inter-University Centre for Behavioural Neurosciences, University of Verona; M. BALESTRIERI, MD, PhD, Verona–Udine Brain Imaging and Neuropsychology Programme, Inter-University Centre for Behavioural Neurosciences, and Department of Pathology and Experimental and Clinical Medicine, Section of Psychiatry, University of Udine; C. BARBUI, MD, M. NOSÉ, MD, Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology, University of Verona; A. GASPARINI, MD, Department of Morphological and Biomedical Sciences, Section of Radiology, GB Rossi Hospital, University of Verona; P. BRAPBILLA, MD, PhD, Verona–Udine Brain Imaging and Neuropsychology Programme, Inter-University Centre for Behavioural Neurosciences, and Department of Pathology and Experimental and Clinical Medicine, Section of Psychiatry, University of Udine and Scientific Institute IRCCS I Medea, Udine, Italy

and Neuropsychology Programme, Inter-University Centre for Behavioural Neurosciences, University of Verona; A. VERSACE, MD, G. RAMBOLDI, MS, C. PERLINI, PsychD, N. DUSI, MD, L. PELIZZA, PsychD, Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology, University of Verona, and Verona–Udine Brain Imaging and Neuropsychology Programme, Inter-University Centre for Behavioural Neurosciences, and Department of Pathology and Experimental and Clinical Medicine, Section of Psychiatry, University of Udine; C. BARBUI, MD, M. NOSÉ, MD, Department of Medicine and Public Health, Section of Psychiatry and Clinical Psychology, University of Verona, and Verona–Udine Brain Imaging and Neuropsychology Programme, Inter-University Centre for Behavioural Neurosciences, and Department of Pathology and Experimental and Clinical Medicine, Section of Psychiatry, University of Udine and Scientific Institute IRCCS I Medea, Udine, Italy
Frontal release signs and cognition in people with schizophrenia, their siblings and healthy controls

THOMAS M. HYDE, TERRY E. GOLDBERG, MICHAEL F. EGAN, MARC C. LENER and DANIEL R. WEINBERGER

Background Frontal release signs, a subset of neurological soft signs, are common in schizophrenia.

Aims To explore the relationship between frontal release signs and neuropsychological tests of frontal lobe function in people with schizophrenia, their siblings and healthy controls.

Method Neuropsychological tests and frontal release signs were measured in a cohort of index cases (n = 302), their siblings (n = 240) and healthy controls (n = 346).

Results The mean total score of frontal release signs was 1.5 (s.d. = 1.58) in the schizophrenia group, 0.54 (s.d. = 0.92) for siblings and 0.42 (s.d. = 0.77) for controls. Schizophrenia group scores were greater than healthy control or sibling cohort scores (P < 0.0001), which did not differ. In all three cohorts, right grasp reflex scores positively correlated with number of perseverative errors on the Wisconsin Card Sort Task (P < 0.05). In the schizophrenia group, frontal release signs showed an inverse correlation with IQ (R = -0.199, P < 0.0005).

Conclusions Our findings of relationships between frontal release signs and cognitive assays of cortical dysfunction and the increased frequency of these signs in people with schizophrenia implicate a cortical origin for these clinical signs and evidence of frontal lobe dysfunction in this disorder.

Declaration of interest None.

An association between schizophrenia and subtle neurological abnormalities has been reported for nearly a century, with extensive study over the past 40 years (Woods et al., 1986). People with schizophrenia have a high frequency of abnormal subtle neurological findings, also known as neurological soft signs. There may be a genetic origin to neurological abnormalities in schizophrenia, as siblings without the disorder demonstrate a greater incidence of these signs (Kinney et al., 1986; Ismail et al., 1998; Niethammer et al., 2000). Frontal release signs (also known as primitive reflexes) are a subset of neurological soft signs: they consist of a group of involuntary motor responses which are normally found early in postnatal development and are subsequently inhibited, but may be ‘released’ from inhibition by cerebral, usually frontal, damage (Paulson & Gottlieb, 1968; Schott & Rossor, 2003). Frontal release signs are common in the general population, occurring in roughly a quarter of young healthy adults, and are more common with advancing age (Gladstone & Black, 2002). Although a single sign is of limited clinical significance, multiple signs tend to correlate with brain pathology (Isakov et al., 1984; Schott & Rossor, 2003). In this study we examined a large cohort of people with schizophrenia, their healthy siblings and a healthy non-psychiatric control comparison group to determine the frequency with which frontal release signs occur in these three groups, and specifically to assess the familiarity of these signs by comparing the schizophrenia and sibling cohorts. Second, we sought to investigate the association between frontal release signs and cognitive impairment, with a focus on measures of frontal lobe function as well as more general cognitive measures such as full-scale IQ.

METHOD

Data concerning neurological soft signs were collected from 302 people with a diagnosis of schizophrenia or schizoaffective disorder, 240 of their full siblings and 346 members of a healthy control group who had participated in a study of neurobiological phenotypes associated with schizophrenia – the Clinical Brain Disorders Branch/National Institute of Mental Health (NIMH) Sibling Study (Egan et al., 2000). Siblings were excluded if they were diagnosed with schizophrenia or a schizophrenia-spectrum disorder; otherwise, they were included irrespective of their psychiatric status. In a previous analysis of neurological soft signs in our laboratory, 407 out of the 888 participants in this study overlapped (43.8%; Egan et al., 2001b). The NIMH institutional review board approved all procedures and testing. Details of recruitment and exclusion criteria have been described previously (Egan et al., 2000). Briefly, families were recruited from local and national sources; the comparison group was recruited from the National Institutes of Health volunteer office and was matched to the sibling group on age, gender, education and performance on the Wide Range Achievement Test (WRAT; Jastak & Wilkinson, 1984). In order to reduce the sample heterogeneity, only White participants (who were Caucasians of European ancestry) were included in this analysis. All participants were screened to exclude those with a premorbid full-scale IQ below 70, those with recent (within 1 year) significant drug or alcohol misuse or more than 5 years of prior dependence, and those with significant medical or neurological conditions. The healthy control group was selected with the additional requirement that they should not have a first-degree relative with a schizophrenia-spectrum disorder (Egan et al., 2000). All patients were clinically stable, with no psychiatric hospitalisation within 6 months of entry into the study. Written informed consent was obtained after complete description of the study to the participants.

Procedure All participants were interviewed by a research psychiatrist (masked to family status) using the Structured Clinical Interview for DSM–IV Axis I Disorders (SCID; First et al., 1996). A second research psychiatrist reviewed all diagnostic data and the consensus diagnosis was used. All participants had a thorough medical evaluation, including magnetic resonance imaging of the brain. They were administered an
extensive battery of neuropsychological tests as described in detail elsewhere (Weickert et al., 2000). These included the short form of the Wechsler Adult Intelligence Scale – Revised (WAIS–R; Wechsler, 1981), a measure of full-scale IQ. The WRAT was administered to assess premorbid IQ. Participants were tested also on a variety of cognitive measures that assess prefrontal function: working memory/executive function was tested with the Wisconsin Card Sorting Test (WCST; Heaton, 1981) and Letter fluency (Benton et al., 1983); psychomotor speed and oculomotor scanning were tested with Trail making test part B (Reitan, 1986); and working memory/updating was tested with the n-back task (Goldberg et al., 2003) (one-back and two-back) (Egan et al., 2001a).

Demographic data are presented in Table 1. Participants underwent a detailed neuropsychological examination by one of two research neurologists who were formally masked to diagnosis and familial relationships. Inter-rater reliability was assessed on 10 participants and revealed that all ratings were significantly correlated (intraclass correlation coefficients 0.54–0.90, P < 0.02). The examination included the Neurological Evaluation Scale (NES; Buchanan & Heinrichs, 1989) scored as previously described (Sanders & Keshavan, 1998). The examiners followed the previously published clinical procedures for assessing each frontal release sign (Paulson & Gottlieb, 1968; Ovsiew, 1997).

Data analyses

The primary outcome measures were the individual and summed frontal release sign scores from the NES (Buchanan & Heinrichs, 1989). For these measures, which are not independent, P=0.05 was accepted as significant. Data analyses were performed using SAS (version 9.1 for Windows). Mean scores were contrasted by mixed model analysis of variance (ANOVA), treating family status as a random effect. In order to investigate the relationships between frontal release signs and cognitive function, the summed frontal release sign scores were calculated from the individual tests of frontal lobe function from the NES: these included ratings of glabellar, suck, snout, and right and left grasp reflexes. Correlations between averages of the total and individual frontal release sign scores were obtained for each diagnostic group (schizophrenia, sibling and healthy control), using Pearson correlation coefficients. When shared environmental factors are not considered causative of a shared trait in family members, relative risk is commonly thought to reflect shared genetic factors. Relative risk was assessed by comparing the proportion of affected individuals in the sibling cohort with the proportion of affected individuals in the control cohort (‘affected’ was defined by a summed frontal release sign score greater than 1 standard deviation above the control group mean). A chi-squared analysis was performed to test the significance of the relative risk. As an additional test of possible heritability in sibships, an intraclass correlation coefficient was calculated for total frontal release sign scores.

RESULTS

In the schizophrenia cohort the maximum individual frontal release sign score (on a scale of 0–10) was 7, with an average of 1.50 (s.d.=1.58). In the control cohort, the maximum score was 4, with a mean of 0.42 (s.d.=0.77). In the sibling cohort the maximum score was 6, with a mean of 0.54 (s.d.=0.92) (Fig. 1). Using a mixed model ANOVA, contrasting total frontal release sign scores with diagnostic group, the schizophrenia cohort had significantly higher average scores than the control and sibling cohorts (d.f. = 746, F = 76.26, t = 10.77 and t = 10.52 respectively; P<0.0001). The control and sibling cohorts did not differ (t = 1.23; P=0.22). The relative risk was 1.40 (P=0.25 by χ² analysis), suggesting that frontal release signs are not a strongly familial characteristic in schizophrenia. An alternative method of testing heritability, the intraclass correlation coefficient, was not significant at 0.14 for total frontal release sign score (P=0.98).

The schizophrenia cohort had the greatest number of significant correlations between total and individual frontal release

### Table I  Demographic characteristics of the sample

<table>
<thead>
<tr>
<th></th>
<th>Study group</th>
<th>Post hoc analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Index (n=302)</td>
<td>Siblings (n=240)</td>
</tr>
<tr>
<td>Age, years: mean (s.d.)</td>
<td>36.5 (9.2)</td>
<td>37.0 (9.8)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male, %</td>
<td>76.8</td>
<td>50.0</td>
</tr>
<tr>
<td>Education, years: mean (s.d.)</td>
<td>14.0 (2.3)</td>
<td>15.8 (2.4)</td>
</tr>
<tr>
<td>WAIS full-scale IQ: mean (s.d.)</td>
<td>93.0 (12.1)</td>
<td>106.1 (10.4)</td>
</tr>
<tr>
<td>WRAT (premorbid IQ): mean (s.d.)</td>
<td>102.5 (11.3)</td>
<td>107.4 (9.9)</td>
</tr>
</tbody>
</table>

NS, not significant; WAIS, Wechsler Adult Intelligence Scale; WRAT, Wide Range Achievement Test.

1. Includes participants enrolled full time in college or graduate school.

**Fig. 1** The participants with schizophrenia had significantly higher total frontal release sign (FRS) scores than either their siblings or a non-related control group of healthy individuals (error bars represent standard deviations).
sign scores and performance on the neuropsychological testing battery, although all were weak (Table 2). In particular, total and individual frontal release sign scores had highly significant inverse correlations with performance on the WAIS–R full-scale IQ. In fact, if the results are Bonferroni-corrected, the only result that meets the corrected criteria of P < 0.0006 is the correlation between total FRS score and WAIS–R full-scale IQ in the schizophrenia group. Within the uncorrected data-set, in the control group there was a trend towards an inverse correlation between total frontal release sign scores and WAIS–R full-scale IQ (Table 3). For percentage of perseverative errors on the WCST, a number of significant correlations were noted across all three cohorts. In the schizophrenia cohort the presence of a right grasp reflex positively correlated with number of perseverative errors, whereas in the control cohort, total and several individual frontal release sign scores positively correlated with number of perseverative errors. In the sibling cohort, right grasp reflex scores positively correlated with number of perseverative errors (Table 4). Finally, in the sibling cohort there was a positive correlation between left grasp reflex scores and time to complete the Trails B test.

**DISCUSSION**

Classical clinical–pathological correlations have suggested that frontal release signs in adults are one of the few bedside indices of prefrontal cortical dysfunction. Participants with schizophrenia in our study had a much higher number of frontal release signs on average than controls or their unaffected siblings. This finding is largely in agreement with a number of previous studies (Taylor & Abrams, 1984; Woods et al, 1986; Liddle, 1987; Ismail et al, 1998; Sanders & Keshavan, 1998; Egan et al, 2001b; Lawrie et al, 2001; Cuesta et al, 2002; Gourion et al, 2004). In this study, both individual and summed frontal release sign scores showed weak inverse correlations with several neuropsychological measures, including full-scale IQ, and a positive correlation with number of perseverative errors on the WCST (a test of executive function that reliably engages the dorsolateral prefrontal cortex). These findings were most apparent in the schizophrenia group, perhaps as a result of a greater dynamic range in frontal release sign and cognitive scores.

The greater number of frontal release signs in people with schizophrenia compared with siblings and normal controls was reported previously (Ismail et al, 1998). In fact, a grouping of neurological soft signs that included frontal release signs, abnormalities in eye movements and short-term memory deficits differentiated people with schizophrenia from a healthy control group better than any other sub-scale from the Neurological Evaluation Score (Arango et al, 1999). Both genetic and environmental factors have been cited as the cause of neurological soft signs in schizophrenia; in these studies frontal release signs were summed within a larger set of clinical measures (neurological soft signs) and were
not examined independently (Ismail et al., 1998; Egan et al., 2001b).

Neurological soft signs previously have been associated with cognitive impairment (Taylor & Abrams, 1984; Liddle, 1987; Schonfeld et al., 1989; Cuesta et al., 2002), including a propensity towards lower IQ (Obiols et al., 1999; Fellick et al., 2001) and a poor long-term functional outcome in first-episode patients (Johnstone et al., 1990). Patients with schizophrenia who had a higher number of soft signs had lower IQ (Kennard, 1960; Mosher et al., 1971; Marcus et al., 1985) and deficits on measures of executive function such as working memory, learning and attention (Saykin et al., 1991; Franke et al., 1992; Braff, 1993; Paulsen et al., 1995). Neuroanatomically, soft signs have been associated with regional grey-matter volume changes that may be an index of perturbed cortical–subcortical connectivity (Dazzan et al., 2004). Although not specific to schizophrenia, neurological soft signs appear to be an intrinsic element of this disorder and have a negative connotation with respect to illness severity.

Heritability of frontal release signs

There is a suggestion of the heritability of neurological soft signs in general, when studying people with schizophrenia and their non-affected siblings (Egan et al., 2001b; Gourion et al., 2003; 2004). Somewhat unexpectedly, frontal release signs do not appear to share the same characteristic, at least when siblings are compared with healthy controls using a chi-squared analysis of relative risk. Moreover, the intraclass correlation coefficient also was not significant. This suggests that environmental factors may have a significant role in the development of frontal release signs, and by extension, some aspects of frontal pathology in schizophrenia. It should be noted, however, that our study rigorously excluded siblings with schizophrenia and schizophrenia-spectrum disorders.

The notion that neurological abnormalities in schizophrenia might be due to genetic factors came from studies that found a higher incidence of these abnormalities in family members without schizophrenia of patients with this disorder (Ismail et al., 1998; Niethammer et al., 2000). However, the genetic contribution to the liability towards the development of neurological abnormalities in schizophrenia is modest at best (Kinney et al., 1986; Rossi et al., 1990; Ismail et al., 1998; Niethammer et al., 2000; Egan et al., 2001b). In our study we also found very modest evidence of the heritability of frontal release signs, as measured by relative risk or intraclass correlation. The greater incidence of frontal release signs in people with schizophrenia compared with their non-affected siblings suggests a significant role for environmental factors. Alternatively, in any given individual, genetic susceptibility to schizophrenia is probably secondary to the interaction of polymorphisms in several genes with small interacting effects, acting in conjunction with the effects of environmental factors (Wildenauer et al., 1996; Harrison & Weinberger, 2005). People with schizophrenia most probably have both a much greater genetic load and a greater exposure to predisposing environmental factors than their non-affected siblings. Hence, the lack of heritability of frontal release signs may reflect the effects of these genetic–environmental interactions.

Frontal release signs and cognitive impairment

In general our correlation between the presence of frontal release signs and poor performance on neuropsychological tests of prefrontal function in schizophrenia agrees with previous reports. However, previous studies differ in some of the details of the findings from our study. One group found that frontal release sign scores correlated with a higher number of random errors but not with perseverative errors on the WCST in people with schizophrenia (Wong et al., 1997). In two studies, global scores of neurological soft signs correlated with perseverative errors on the WCST in schizophrenia, but frontal release signs were not specifically examined (Braun et al., 1995; Mohr et al., 2003). Poor performance with people with schizophrenia on the WCST (as measured by achieved categories) directly correlated with a greater number of neurological soft signs (Bersani et al., 2004). Most studies of frontal release signs have relied upon much smaller samples and therefore have less statistical power than our study. Moreover, most studies of the relationship between neurological abnormalities and prefrontal dysfunction in schizophrenia have not examined frontal release signs separately from other neurological soft signs.

The inverse correlation between full-scale IQ and frontal release signs in schizophrenia is the clearest result in our data. In fact, this is the only correlation that withstands a rigorous Bonferroni correction. This suggests that at least in this sample frontal release signs implicate a more widespread pathology of higher cortical systems. Barnes et al. (1995) measured frontal release sign scores and performance on the WAIS-R in people with schizophrenia: no significant correlation was found, but the total sample size of 48 gave limited power. Two studies have reported an inverse relationship between high scores on more broad-based measures of neurological soft signs and low scores on IQ tests. Neither of these studies evaluated the participants for the presence of frontal release signs; instead, they relied upon other neurological soft signs (Obiols et al., 1999; Fellick et al., 2001). Mohr et al. (2003) found that overall neuropsychological performance inversely correlated with the total score on a battery of neurological soft signs. However, that study did not assay frontal release signs. They did note that higher soft sign scores inversely correlated with a number of sub-tests from the WAIS-R. In general, our findings agree with these previous reports.

Side-effects of antipsychotics might account for our findings. However, similar neurological abnormalities have been noted in schizophrenia for nearly a century, decades before the introduction of antipsychotic therapy (Bleuler, 1950; Kraepelin, 1971). Other studies support the notion that neurological soft signs are an intrinsic part of schizophrenia rather than a direct or indirect consequence of treatment (Johnstone et al., 1990; Gupta et al., 1995; Browne et al., 2000; Mohr et al., 2003; Dazzan et al., 2004). Neurological soft signs and frontal release signs are common in first-episode schizophrenia (Browne et al., 2000); in fact it has been reported that frontal release signs are more common in people with schizophrenia who have never been treated with antipsychotics than in treated patients (Gupta et al., 1995). In addition, unmedicated participants at high risk of schizophrenia have more neurological soft signs than healthy control individuals (Lawrie et al., 2001). These findings suggest that exposure to antipsychotics is not necessary for the appearance of neurological soft signs in general and frontal release signs in particular. In addition, both typical (Mishara & Goldberg, 2004) and atypical (Weickert et al., 2003) antipsychotics may sometimes improve cognitive performance,
militating against antipsychotics as a cause. As a whole, the preponderance of the findings in the psychiatric literature makes it highly unlikely that the findings in our study are solely directly attributable to the deleterious effects of antipsychotic medications. In summary, people with schizophrenia had more frontal release signs than their siblings or the control group. In the schizophrenia group both individual and summed frontal release sign scores inversely related with several neuropsychological measures, including full-scale IQ and the number of perseverative errors on the WCST (a relatively selective test of prefrontal function). Although these findings were most apparent in the participants with schizophrenia, perhaps as a result of a greater range in frontal release sign scores, a trend for similar relationships was also seen in the control group. This suggests that frontal release signs are at best a weak index of prefrontal cognitive dysfunction, particularly in schizophrenia, but also in healthy individuals.

**ACKNOWLEDGEMENTS**

This research was supported in its entirety by the intramural research programme of the National Institute of Mental Health. The authors would also like to thank the clinical staff of the Clinical Brain Disorders Branch, Genes, Cognition and Psychosis Program, for their efforts in patient recruitment and characterisation, and Dr Llewellyn Bigelow in particular for his efforts in clinical diagnosis.

**REFERENCES**


Genetic influence on laterality in schizophrenia?

Textbook of Neuropsychiatry


Emotion recognition and genetic vulnerability to schizophrenia

BENOIT BEDIOU, FATIMA ASRI, JEROME BRUNELIN, PIERRE KROLAK-SALMON, THIERRY D'AMATO, MOHAMED SAOUD and IMANE TAZI

Background  Epidemiological studies of schizophrenia suggest that this disorder has a substantial genetic component. Cognitive and social abilities, as well as the volumes of brain regions involved in emotion processing, have been found to be distributed along a continuum when comparing patients, siblings and controls, with siblings showing intermediate scores.

Aims  To establish whether facial expression recognition is impaired in unaffected siblings of patients.

Method  Emotion and gender recognition were evaluated in a three-group pre–post study design in drug-naive patients with first-episode schizophrenia (n=40) and their unaffected siblings (n=30) compared with controls (n=26).

Results  Patients and their healthy siblings showed impaired emotion recognition but normal gender recognition compared with controls. Patients’ performance did not improve despite effective clinical stabilisation.

Conclusions  Impaired performance in healthy siblings and time stability in patients provides evidence of impairment of facial emotion recognition as an actual phenotype of schizophrenia.

Declaration of interest  None.

A key component of social cognition, facial expression recognition, is one of the hallmark deficits of schizophrenia (Schneider et al, 2006). Although people with schizophrenia may show deficits in other aspects of face perception (Hall et al, 2004), only expression recognition performance has been related to social dysfunction (Hooker & Park, 2002). Studies have implicated interaction between the amygdala and prefrontal and temporal cortex in facial expression processing (Krolak-Salmon et al, 2004). In schizophrenia this deficit has been related to amygdala–hippocampal complex (AHC) dysfunction (Gur et al, 2002), although the frontal and temporal cortex may also have a role (Streit et al, 2001). Many of these regions, including the amygdala and the prefrontal and superior temporal cortices, are known to be reduced in volume in schizophrenia – see reviews by Lawrie & Abukmeil (1998) and Wong & Van Tol (2003) and the meta-analysis by Wright et al (2000). Interestingly, despite a recent contradictory result (McDonald et al, 2006), reduced volume of the AHC has been consistently reported in healthy siblings of patients (Lawrie et al, 1999; Aleman & Kahn, 2005). However, some methodological issues (inclusion criteria, regions of interest, medial temporal lobe considered as a whole) could account for these discrepancies. Moreover, siblings show poor perception of non-verbal social–emotional cues (Toomey et al, 1999). Facial expression recognition is one of the most crucial vectors of non-verbal communication relying on amygdala integrity and may therefore be impaired in healthy relatives of patients, although this has never been evidenced.

Our hypothesis was that facial expression recognition performance, like AHC volume, would be distributed along a continuum between controls and patients with schizophrenia, with healthy relatives of patients having intermediate values, as has been shown for several cognitive (Saoud et al, 2000; Brunelin et al, 2007) and social functions (Toomey et al, 1999). Moreover, as previously reported with other cognitive deficits, patients are unlikely to perform as well as controls even when the disorder is effectively treated (Brunelin et al, 2007). To test this hypothesis we measured facial emotion recognition in patients experiencing their first schizophrenic episode before and after first antipsychotic treatment and in their healthy siblings compared with controls. The method used in the current study is original in using morphing controlling for performance in another task requiring facial feature analysis, that is gender recognition (Bediou et al, 2005). Our previous report of patients’ preserved performance in this task suggests that it does not rely on the same dysfunctional structures.

METHOD

Forty drug-naive patients (mean age 26.9 years, s.d.=5.4) experiencing their first episode of schizophrenia diagnosed according to DSM–IV criteria (American Psychiatric Association, 1994), 30 of their non-affected first-degree relatives (mean age 30.4 years, s.d.=10.4) and 26 healthy controls (mean age 24.3 years, s.d.=3.3) were enrolled in this study. All participants were male. Their vision was normal or corrected to normal. Groups differed significantly in age (F(2,91)=5.12, P<0.01) and in years of full-time education (F(2,91)=5.12, P<0.001). All participants received detailed information regarding the study and gave their written informed consent. The study was conducted in accordance with the latest version of the Declaration of Helsinki and its design was approved by the standing ethics committee. Recruitment took place between September 2003 and September 2005. Patients with schizophrenia were recruited at the Ibn Nafis psychiatric hospital in Marrakech. All participants were Moroccans, born, raised and currently living in Morocco. The group of healthy relatives consisted of unaffected siblings (brothers) of patients participating in the study; they had no past or current history of any Axis I DSM–IV disorder as verified by the Mini International Neuropsychiatric Interview (MINI; Sheehan et al, 1998) semi-standardised evaluation, version 4.4. Participants in the healthy control group also had no personal or familial past or current history of Axis I DSM–IV disorder
as tested by MINI evaluation; they were recruited from the hospital staff and among students from the Medical and Pharmaceutical University of Marrakech. None had ever received psychotropic medication and none reported any misuse of drugs.

To address time stability and treatment effect of facial emotion recognition deficit in schizophrenia, participants in the schizophrenia group were tested twice, at baseline and after clinical stabilisation of symptoms under haloperidol treatment (mean dose 10.0 mg, s.d. = 1.6). To control for potential learning or training effects, the other two study groups were also tested twice over a similar period. The mean interval between baseline and follow-up testing was 30.2 days (s.d. = 5.3) in each group. Testing sessions consisted of each participant performing the expression and gender recognition task described below, the order being counterbalanced across participants and, for each participant, between testing sessions. At the time of each testing, symptoms were assessed by a trained clinician masked to participants’ status. Patients’ schizophrenic symptoms were assessed with the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987). Since depression may affect emotion recognition performance, depressive symptoms were also controlled for in the sibling and schizophrenia groups with the Calgary Depression Scale for Schizophrenia (CDSS; Addington et al., 1990) and in the control group with the abbreviated version of the Beck Depression Inventory (BDI; Beck et al., 1961). Schizotypal personality traits were evaluated in the sibling and control groups with the Schizotypal Personality Questionnaire (SPQ; Raine, 1991). Demographic and clinical data are summarised in Table 1.

### Tasks

We used sensitive measures of emotional facial expression recognition and a control measure of facial gender recognition, both using morphed faces. A detailed description of the method can be found in our previous work (Bediou et al., 2005).

#### Emotional facial expression recognition

Briefly, 132 faces from two male and two female faces morphed between a neutral and an emotional expression (happiness, disgust, fear, anger) in 10% steps were each randomly presented on a computer screen for 1000 ms, preceded by a fixation cross (250 ms) and followed by the choice list (neutral, happiness, disgust, fear or anger), which was maintained on the screen until the participant responded. Participants were required to match each face with the word that best described the emotional expression displayed by pressing the corresponding key.

#### Facial gender recognition

In the gender recognition task 132 faces, morphed between an average face with ‘no gender’ (obtained by averaging 20 male and 20 female faces in equal proportion) and one of six male or one of six female faces in 10% steps, were each randomly presented for 1000 ms, preceded by a fixation cross (250 ms) and followed by the choice list (male or female). Participants were instructed to match each face with the word that best described its gender by pressing the corresponding key.

### Statistical analyses

Dependant variables were the percentage of correct responses in the emotion and gender recognition tasks. Independent variables were the group (schizophrenia, siblings or control); the phase (baseline or follow-up); the task (expression or gender); the intensity of emotional expression or gender (0–100% in 10% increments); the expression (happiness, disgust, fear or anger); and the gender (male or female). The first variable is an intra-individual factor whereas the others are inter-individual factors.

Performance of the schizophrenia group, time stability/treatment effect and performance of the siblings group were tested by means of repeated-measures analyses of variance; a main 3 group × 2 phase × 10 task × 10 intensities multivariate analysis of variance (MANOVA) was followed by a 3 group × 2 phase × 4 expression × 10 intensities and a 3 group × 2 phase × 2 gender × 10 intensities MANOVA. To correct for multiple comparisons, Bonferroni correction led to a level of significance retained at 0.01 for analyses of variance (ANOVAs). Significant effects and interactions were then tested by means of planned comparisons using single-factor ANOVAs for inter-individual factors and contrast analyses.

### Table 1 Characteristics of the sample (values are means and standard deviation)

<table>
<thead>
<tr>
<th></th>
<th>Schizophrenia group (n=40)</th>
<th>Siblings group (n=30)</th>
<th>Control group (n=26)</th>
<th>P*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Follow-up</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td>27.3 (5.8)</td>
<td>31.2 (10.6)</td>
<td>24.3 (3.3)</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Educational level, years</td>
<td>9.6 (2.1)</td>
<td>9 (6.6)</td>
<td>16.2 (1.1)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>PANSS score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Positive (max. 49)</td>
<td>29.5 (7.1)</td>
<td>10.2 (6.7)</td>
<td></td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Negative (max. 49)</td>
<td>27.2 (7.6)</td>
<td>11.4 (3.8)</td>
<td></td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>General (max. 112)</td>
<td>39 (6.6)</td>
<td>20 (4)</td>
<td></td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>CDSS (max. 27)</td>
<td>1.4 (2.2)</td>
<td>0.1 (0.7)</td>
<td>0.6 (2.0)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>SPQ (max. 74)</td>
<td>14.1 (13.5)</td>
<td>10.8 (11.7)</td>
<td></td>
<td>NS</td>
</tr>
<tr>
<td>BDI (max. 39)</td>
<td>1.9 (3.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BDI, Beck Depression Inventory; CDSS, Calgary Depression Scale for Schizophrenia; max., maximum; SPQ, Schizotypal Personality Questionnaire.

* Between-group comparisons were performed with single-factor analyses of variance for age and education level and with Student's t-tests for SPQ scores. Within-group comparisons for PANSS and CDSS scores were assessed with paired t-tests.
using the full data set for intra-individual factors (e.g., phase effects: baseline vs. follow-up). Student’s *t*-tests were used to probe group differences in socio-demographic characteristics. Given the significant group differences in age and education, and since performance did not correlate with either of these variables, repeated-measures analyses of covariance (MANCOVA) were also performed to control for a potential influence of these variables on significant effects. Since the same significant effects and interactions were found with MANOVAs and corresponding MANCOVAs, only results from the MANOVAs are reported.

**RESULTS**

Positive, negative and general schizophrenia symptoms as well as depressive symptoms were significantly reduced between baseline and follow-up in the patient group (Table 1).

**Emotion and gender recognition performance**

The main MANOVA revealed a main effect of group \(F(12,186) = 28.97, P < 0.001\), task \(F(1,186) = 777.86, P < 0.001\) and intensity \(F(9,1674) = 917.27, P < 0.001\) and significant group × task \(F(2,186) = 31.75, P < 0.001\), group × intensity \(F(11,1674) = 10.51, P < 0.001\), task × intensity \(F(9,1674) = 105.14, P < 0.001\) and group × task × intensity interactions \(F(11,1674) = 6.86, P < 0.001\). There was no main effect of phase nor any interaction of phase with group, task or intensity. Consistent with our hypothesis, performance did not differ significantly between baseline and follow-up in any of the groups. There was a difference between overall emotion and gender recognition performance in all groups, suggesting that the emotion task was more difficult: emotion and gender recognition in the control group were 61% (s.d. = 9) and 78% (s.d. = 7) respectively. However, this effect was modulated by group. The schizophrenia group performed significantly worse than the control group on emotion recognition at baseline \(F(1,186) = 38.76, P < 0.001\) and the difference remained significant at follow-up \(F(1,186) = 38.16, P < 0.001\) despite clinical stabilisation of symptoms. Moreover, the siblings group performed significantly worse than the control group but significantly better than the schizophrenia group in overall facial expression recognition (Fig. 1; scores averaged across all emotions and phases). No difference was found between groups on overall gender recognition \(F(1,186) = 1.06, NS\). This absence of group difference was confirmed at all intensities of the gender morphs.

Analysis by emotion category revealed significant group × expression \(F(6,558) = 3.69, P < 0.001\) and group × expression × intensity interactions \(F(14,9222) = 1.50, P = 0.01\). Performance differed between the schizophrenia and control groups in the recognition of all basic emotions and between the sibling and control groups in the recognition of disgust and fearful faces (Fig. 2). Correlations between facial expression recognition and age \((r = -0.029, P = 0.88)\) and between facial expression recognition and educational level \((r = 0.073, P = 0.71)\) were not significant. Likewise, emotion recognition performance did not correlate with mood (assessed with the BDI) in the control group, nor with depression scores (assessed with the CDSS) in the sibling and schizophrenia groups.

**DISCUSSION**

Consistent with our hypothesis, emotion recognition performance was distributed along a continuum between the control group and the schizophrenia group, with the siblings group midway between, whereas gender recognition was highly preserved in all groups. Such impairment may reflect a specific deficit in emotion processing rather than a general impairment in face processing. Thus, the continuum in emotional facial expression recognition parallels the one observed in cognition (Saoud et al., 2000; Sitskoorn et al., 2004; Brunelin et al., 2007) and in perception of other non-verbal social-emotional cues (Toomey et al., 1999) as well as in AHC volume (Lawrie et al., 1999). However, there have been few published research studies in which participants with schizophrenia were not found to differ from healthy controls in their emotion recognition skills (Flack et al., 1997; Gur et al., 2002). These data and our results suggest that emotion recognition skills may occur on a continuum, and that individuals with schizophrenia, in general, tend to score lower on this continuum than their healthy siblings or healthy unrelated individuals.

Our results also revealed that facial expression recognition was already impaired at the onset of the illness, before treatment, and that performance remained impaired after 1 month of low-dose haloperidol treatment despite symptom stabilisation, suggesting trait-like features. Although age and educational level differed significantly between groups, it is unlikely that these factors account for the group differences observed here. Indeed, neither correlation nor comparison between MANOVA and MANCOVA results suggested an effect of age or education on emotion recognition. In addition, most previous studies failed to find any influence of these variables on performance.

Abnormal visual scanning of emotional faces may indirectly underlie some aspects of face recognition impairments in schizophrenia (Loughland et al., 2002). It has been related to amygdala function (Taylor et al., 2002; Adolphs et al., 2005). Evidence for the stability of these disturbances over time and illness progression points to their trait-like nature (Streit et al., 1997). Moreover,
abnormal visual scanning of facial expressions has also been documented in unaffected siblings of people with schizophrenia (Loughland et al., 2004), reinforcing the assumption that emotion recognition deficit can be considered as a potential marker of familial vulnerability to schizophrenia. Unlike expression recognition, gender recognition may rely on a more holistic face analysis that is achieved in posterior occipito-temporal areas despite abnormal visual scanning of face parts.

In contrast to the overall deficit in facial expression recognition in the schizophrenia group, unaffected siblings were only impaired in fear and disgust recognition, suggesting that the emotion-specific deficit of healthy relatives may generalise with illness onset. However, a comparable deficit in fear and disgust recognition had already been observed in participants whose schizophrenia was in remission in our previous study (Bediou et al., 2005). Atypical anti-psychotic treatment, used in the latter study, may therefore be more efficient than haloperidol, used in the current study, in improving emotion recognition in schizophrenia (Kee et al., 1998). However, performance was not completely restored by either of these treatments and the deficit remains significant at all phases of the illness, supporting the trait-like hypothesis. Interestingly, relatives’ performance appeared similar to that of patients when emotional expression was more subtle (30%), whereas it was more like that of the control group at higher intensities of emotional expression (80–90%). This may relate to the efficiency of social cognition processes (Toomey et al., 1999), facial scanning (Loughland et al., 2002), or both. However, this original result must be interpreted with caution. Indeed, even if the sibling and control groups did not differ significantly in anger recognition, lack of power might account for this result.

In social situations inaccurate decoding of emotional expression could be a source of stress and a barrier to deep social interactions and communication. It has been shown that metabolic and probably physical and social stress could cause an acute exacerbation of schizophrenia symptoms (van Os & McGuftin, 2003). Thus, one can easily conceive that abnormal interpretation of emotional expressions could increase the level of emotional stress and thus participate in precipitating vulnerable people into schizophrenia. However, the neurochemistry of emotion recognition remains to be specified. There is growing evidence that dopamine has a central role in emotion recognition (Salgado-Pineda et al., 2005). Direct evidence comes from studies using pro-dopamine (Delaveau et al., 2005) or anti-dopamine (Lawrence et al., 2005) agents in healthy volunteers. Also, indirect evidence is provided by studies of neuropsychiatric patients, specifically in schizophrenia (Bediou et al., 2005) and Parkinson’s disease (Lachenal-Chevallet et al., 2006). Increasing serotonin transmission, on the other hand, can improve emotion recognition performance in healthy individuals (Hamer et al., 2003) and in people with major depressive disorder (Hamer et al., 2004). Abnormalities in the genes involved in dopamine transmission have been associated with poor cognitive functioning, reduced frontal lobe activation and high risk of schizophrenia (Goldberg et al., 2003), and variation in the serotonin transporter gene has been associated with variation in amygdala response to facial expressions (Harrir et al., 2002). This first report of emotion recognition as a marker of familial vulnerability to schizophrenia may encourage studies addressing its association with genetic polymorphism affecting dopamine or serotonin neurotransmission as well as the correlation between emotion recognition performance and amygdala volume. Moreover, given the strong relationship between emotion processing ability and functional outcome, specific remediation programmes should be encouraged (Wolwer et al., 2005). Likewise, the efficacy of emotion processing remediation programmes in vulnerable individuals as a putative preventive strategy should be evaluated.

REFERENCES


American Psychiatric Association (1994) Diagnostic and Statistical Manual of Mental Disorders (4th edn) (DSM-IV), APA.


Long-acting injectable risperidone v. olanzapine tablets for schizophrenia or schizoaffective disorder

Randomised, controlled, open-label study

NICHOLAS A. KEKS, MICHAEL INGHAM, AKBAR KHAN and KEITH KARCHER

Background The efficacy and safety of long-acting injectable risperidone have not been compared with those of an oral atypical antipsychotic.

Aims To compare long-acting risperidone and oral olanzapine in 377 patients with DSM–IV schizophrenia or schizoaffective disorder.

Method Patients were randomised to receive long-acting risperidone (25 mg or 50 mg every 14 days) or olanzapine (5–20 mg/day).

Results In the 13-week phase, long-acting risperidone was at least as effective as (not inferior to) oral olanzapine. In the 12-month phase, significant improvements in the Positive and Negative Syndrome Scale (PANSS) total and factor scores from baseline to month 12 and end-point were seen in both groups of patients. Few patients discontinued treatment because of an adverse event.

Conclusions Both treatments were efficacious and well tolerated.

Declaration of interest N.K. has received support from or been a consultant for AstraZeneca, Bristol-Meyers Squibb, Janssen Pharmaceutical, Eli Lilly, Sanofi-Synthelabo, Pfizer and Wyeth. M.I., A.K. and K.K. are employees of Johnson & Johnson.

The efficacy and safety of long-acting injectable risperidone have been evaluated in several trials of patients with schizophrenia or schizoaffective disorder, including a 12-week, double-blind, placebo-controlled study (n=554; Kane et al, 2003) and a 12-month open-label trial (n=613; Fleischhacker et al, 2003). The effectiveness of long-acting risperidone has also been demonstrated in patients switched from typical and atypical oral antipsychotic medication (Lindenmayer et al, 2004; Chue et al, 2005) and from conventional depot antipsychotics (Turner et al, 2004). In the 12-week double-blind study by Chue et al (2005), long-acting risperidone was compared with oral risperidone in patients with schizophrenia. Both treatments were efficacious and well tolerated. According to a non-inferiority analysis, the two treatments showed comparable efficacy in Positive and Negative Syndrome Scale (PANSS; Kay et al, 1987) total scores over the short term. Long-acting risperidone, however, has not been compared with an oral formulation of another atypical antipsychotic such as olanzapine. In this 53-week, open-label, randomised controlled international study (registered with the US National Institutes of Health at http://clinicaltrials.gov as NCT00236457) we compared long-acting risperidone with olanzapine tablets in patients with schizophrenia or schizoaffective disorder.

The objectives of the study were, first, to demonstrate that in the short term long-acting injectable risperidone was at least as effective as (not inferior to) oral olanzapine in patients with schizophrenia or schizoaffective disorder. Non-inferiority would be demonstrated if, at the end of the initial 13-week treatment period, the upper limit of the confidence interval for the difference in mean change from baseline in PANSS total scores was not more than 8 points in favour of oral olanzapine. The second objective was to examine the long-term efficacy and safety of long-acting risperidone and oral olanzapine in these patients.

METHOD

The study protocol and amendments were reviewed by independent ethics committees or institutional review boards. The study was conducted in accordance with the recommendations guiding physicians in biomedical research involving humans contained in the 1989 version of the Declaration of Helsinki and according to the guidelines of the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2001. All patients or their legal representatives gave their written informed consent to participate in the trial.

Participants Patients with schizophrenia or schizoaffective disorder were recruited at 48 centres (in Australia, Belgium, France, Germany, Greece, Luxembourg, Poland, Russia, Spain, The Netherlands and the UK). Inclusion criteria included a diagnosis of schizophrenia or schizoaffective disorder (DSM–IV; American Psychiatric Association, 1994); PANSS total score 50 or over at randomisation; age at least 18 years; body mass index (BMI) not exceeding 40 kg/m²; and the requirement that within the previous 2 months the patient had been hospitalised or required medical intervention for an acute exacerbation of psychosis and had experienced an additional acute exacerbation during the previous 2 years. Exclusion criteria were prior treatment with clozapine or with a depot antipsychotic within one treatment cycle before screening, and resistance or sensitivity to risperidone or olanzapine. Also excluded were women who were pregnant or breast-feeding or, if of child-bearing age, not using adequate contraception.

Protocol deviations that warranted exclusion from the primary efficacy analysis were:

(a) patients who discontinued before week 8 of treatment (to meet International Conference on Harmonisation guidelines);

(b) patients who received additional antipsychotic treatment (other than oral risperidone for patients in the risperidone arm or olanzapine in the olanzapine arm) between the end of run-in and the end of the first 13-week period;
(c) patients who started treatment with a depot antipsychotic within one treatment cycle before the randomisation visit or who started another depot treatment during the initial 13-week period.

Randomisation

The patients were randomised to receive long-acting risperidone or olanzapine, with stratification factors of psychopathology (PANSS total scores), number of previous psychiatric hospitalisations, BMI and inpatient or outpatient status, using a central dynamic randomisation procedure. Randomisation was based on a minimisation algorithm that used a probability of assignment other than 0.5 to maintain balance of treatment groups within levels of each stratification factor. Constraints on imbalance were defined within each level of each factor; violation of a constraint resulted in adjustment to the treatment assignment probabilities. Randomisation numbers were allocated by an interactive voice response system (IVRS). When a participant was ready to be randomised, the investigator called the IVRS by telephone and entered the person’s stratification information. Based on the minimisation algorithm, the IVRS returned the randomisation number of the appropriate box of study medication at the site.

Dosing and delivery of long-acting risperidone

According to the original study protocol, patients in the long-acting risperidone group received 25, 50 or 75 mg of long-acting risperidone. After study initiation, the original clinical trial programme revealed that the 75 mg dose provided no greater benefit than the lower doses and the protocol was amended to restrict doses to 25 or 50 mg of long-acting risperidone. The 64 patients who had already received 75 mg of long-acting risperidone completed the end-point visit and were then withdrawn from this study and invited to enrol in an open-label extension study. Thus, patients receiving 25 or 50 mg of long-acting risperidone were the focus of the analyses reported here.

During week 1 of the study previous antipsychotic treatments were discontinued and replaced with oral risperidone. The dose of oral risperidone was adjusted to 2, 4 or 6 mg according to each patient’s clinical response. The initial dose of long-acting risperidone was determined by a protocol-specified conversion scheme: patients who had received 2–4 mg of oral risperidone during week 1 received 25 mg per 14 days of long-acting risperidone and patients who had received 6 mg of oral risperidone during week 1 received 50 mg per 14 days of long-acting risperidone. The dosage of long-acting risperidone could be adjusted during the trial according to each patient’s clinical response. Before the protocol was amended and doses were restricted to 25 and 50 mg of long-acting risperidone, patients who had received 6 mg of oral risperidone during week 1 received 75 mg of long-acting risperidone.

Oral risperidone at the week 1 dosage was continued for 3 weeks after the first injection of long-acting risperidone. Oral risperidone supplementation was given when necessary after the initial 3 weeks.

Dosages of oral olanzapine

During week 1 previous medications were discontinued and oral olanzapine was introduced and adjusted to the patients’ optimal dosage of 5–20 mg/day at daily increments of 5 mg. During weeks 2–53 the patients received flexible dosages of 5–20 mg/day of olanzapine.

Assessments of efficacy and safety

The primary measure of efficacy was the change in total score on the PANSS (Structured Clinical Interview) from baseline to end-point (last observation carried forward, LOCF) in the initial 13-week period (short-term outcome). The secondary measures (long-term outcomes) were the changes in PANSS total scores from baseline to month 12 and end-point; changes in PANSS factor scores (positive symptoms, negative symptoms, disorganised thoughts, uncontrolled hostility/excitement and anxiety/depression; Marder et al., 1997); and changes in scores on the Clinical Global Impression – Severity (CGI–S; Guy, 1976) scale. Quality of life was evaluated by means of the Wisconsin Quality of Life Index (Becker et al., 1993). The Wisconsin test was designed for patients with severe mental illness and comprises nine dimensions: life satisfaction, occupational activities, psychological well-being, physical health, social relations, economics, activities of daily living, symptoms and the patient’s own goals.

Clinical improvement was defined as a 20% or greater reduction in PANSS total scores. Maintenance of effect was assessed by determining the time to significant deterioration in the psychotic condition, defined as hospitalisation for symptom exacerbation; the need for an increased level of care and an increase in CGI–S scores of 2 points or more over a 2-week period; or self-injury, suicidal or homicidal ideation or violent behaviour. Significant psychotic deterioration was assessed in the total group and in patients who were rated as stabilised after 13 weeks of treatment. A patient was considered stabilised if he or she had been on the same dosage for 4 weeks or more, the PANSS total score at week 13 did not exceed 70 and the CGI–S score at weeks 9 and 13 was 3 or below and did not increase between weeks 9 and 13.

Assessments were completed at baseline (randomisation), weeks 5, 9, 13, 25, 37 and 53 and at end-point (last observation carried forward, LOCF). The CGI–S was also completed at weeks 1 and 3 and psychotic deterioration was evaluated at week 3. Adverse events were recorded at each visit. Severity of movement disorders was assessed by means of the Simpson–Angus Rating Scale (SARS; Simpson & Angus, 1970) at baseline, at weeks 13, 25, 37 and 53 and at end-point.

Statistical analysis

Differences in changes in PANSS total scores from baseline to end-point (LOCF) in the 13-week study between the two treatment groups were evaluated by an analysis of covariance (ANCOVA) model. Factors included in the model were baseline PANSS score as covariate, randomisation group, the stratification variables (excluding the PANSS factor since it was included as covariate) and investigator nested in country. Because some investigators had only a few patients, pooling of some investigators and countries was required for the fixed-effects ANCOVA model. To avoid pooling, an additional ANCOVA model was performed (for 13 weeks, month 12 visit and end-point) in which country and investigator were treated as random effects. The number and proportion of patients who achieved clinical improvement were tabulated at each assessment point and at end-point. The 95% CI of the odds ratios of the two treatment groups was obtained at month 12 and at end-point. A logistic model (with logit link function and binomial error structure) was applied with the stratification variables as fixed effects and investigator as random effect.
The number and proportion of participants who experienced a significant deterioration were tabulated at each assessment point and at end-point. A Cox proportional hazards model (controlling for the four stratification variables and stratified by country) was used to obtain the 95% CI of the ratio of the hazards in both treatment groups.

RESULTS

Of the 618 patients who were randomised and treated (318 to long-acting risperidone and 300 to olanzapine), 64 were excluded from the short-term (weeks 1–13) analysis of efficacy because they received injections of 75 mg of long-acting risperidone; 66 (38 risperidone and 28 olanzapine patients) were excluded because of major protocol deviations and 110 (52 risperidone and 58 olanzapine patients) were excluded because of non-adherence to Good Clinical Practice standards at one study site. The principal protocol deviations were use of unapproved concomitant medications and inadequate duration of treatment. Thus the per-protocol short-term sample included 164 patients in the long-acting risperidone group and 214 in the olanzapine group (Fig. 1). For the analysis of long-term treatment (months 1–12) a further 2 patients were excluded because they received 75 mg of long-acting risperidone and 14 patients (7 risperidone and 7 olanzapine) were excluded because of major protocol deviations. Thus the per-protocol long-term sample for the evaluation of efficacy included 155 patients in the long-acting risperidone group and 207 in the olanzapine group (Fig. 1). Safety was evaluated in all randomised participants who received at least one dose of study medication and did not receive a 75 mg injection during the entire trial: 247 in the long-acting risperidone group and 300 in the olanzapine group.

Patient characteristics and disposition

Background characteristics of the two patient groups were similar (Table 1). Of the 547 patients who were randomised, received at least one dose of study medication and did not receive a 75 mg injection during the entire trial, 347 (63%) completed the 12-month trial. These included 160 (65%) of the long-acting risperidone group and 187 (62%) of the oral olanzapine group (Fig. 1, Table 2).
Changes in Positive and Negative Syndrome Scale (PANSS) total scores (a) and scores on the five PANSS factors ((b), positive symptoms; (c), negative symptoms; (d), disorganised thoughts; (e) hostility/excitement; (f) anxiety/depression) from week 5 to end-point.

**Table 2** Study completion and reasons for discontinuation

<table>
<thead>
<tr>
<th>Reason</th>
<th>Long-acting risperidone (n=300)</th>
<th>Olanzapine (n=247)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed</td>
<td>65</td>
<td>62</td>
</tr>
<tr>
<td>Discontinued</td>
<td>35</td>
<td>38</td>
</tr>
<tr>
<td>Consent withdrawal</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td>Insufficient response</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Non-adherence</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Adverse events</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Other</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

**Medication dosage and duration**

The patients received a mean of 20.3 injections (s.d.=8.8) of long-acting risperidone. The mean modal dosage was 40.7 mg per 14 days (s.d.=12.1) and the mean duration of treatment with long-acting risperidone was 274 days (s.d.=124.2). A modal dosage of 25 mg per 14 days was received by 37% of the patients and 50 mg per 14 days was received by 63%. Oral risperidone supplementation was received by 98% of the patients during the 3 weeks after the first injection of long-acting risperidone (as per protocol). Oral risperidone supplementation was received by 54 patients (25%) during months 2–3 at a mean modal dosage of 2.2 (s.d.=0.7) mg/day, and by 14–16% patients during the remainder of the trial. When a dosage increase in long-acting risperidone was deemed necessary, additional coverage with oral risperidone was required during the first 3 weeks of the higher dosage.

The mean dose of olanzapine during months 1–12 was 14.6 mg/day (s.d.=4.6) and the duration of treatment was 285 days (s.d.=119.7). Most patients (62%) received modal doses of 15 mg (24%) or 20 mg (38%).

**Concomitant medications**

Concomitant medications were received by 85% of patients in the long-acting risperidone group and 80% of patients in the olanzapine group. These included sedatives or hypnotics, taken by 65 and 53% respectively; antidepressants, taken by 43 and 34%; antiparkinsonian drugs, taken by 37 and 18%; anticonvulsants, taken by 21 and 19%; and muscle relaxants, taken by 11 and 10% respectively.

**Medication adherence**

Medication adherence was high. In the risperidone group the mean injection interval was 14.2 days (range 13–16) and in the olanzapine group the mean time off drug per patient was 0.7 days (s.d.=3.7, range 0–52).

**Efficacy**

**Short-term outcome (weeks 1–13)**

The upper limit of the PANSS 95% CI (score of 3.0) was well below the non-inferiority margin (score of 8), demonstrating the primary end-point that long-acting risperidone was at least as effective as olanzapine (Table 3).

**Long-term outcomes (months 1–12)**

Significant improvements in PANSS total and factor scores from baseline to month 12 and end-point were seen in both groups of patients (Table 3, Fig. 2). Among the patients who completed the long-term trial, significantly greater improvement on one PANSS factor score (disorganised thoughts, P<0.05) was seen in patients receiving long-acting risperidone than in those receiving oral olanzapine (Table 3). At end-point, significantly greater improvement in anxiety/depression was seen in the olanzapine group.

Clinical improvement (20% minimum reduction in PANSS total scores) was achieved by significantly more patients receiving long-acting risperidone than those receiving oral olanzapine at month 12 (91 v. 79%; P<0.001), based on a logistic
LONG-ACTING RISPERIDONE V. ORAL OLANZAPINE

Table 3  Positive and Negative Syndrome Scale total and factor scores in patients receiving long-acting risperidone or olanzapine

|                          | Long-acting risperidone | Olanzapine | LSM of the differences1
|--------------------------|-------------------------|------------|------------------------
|                          | n | Mean (s.d.) | n | Mean (s.d.) | (95% CI) |
| **Short-term outcome (weeks 1–13)2,3** |   |            |   |            |          |
| PANSS total score        |   |            |   |            |          |
| Baseline                 | 164 | 78.7 (14.8) | 213 | 78.6 (14.3) |          |
| Change at end-point      | 164 | −16.9 (15.5) | 213 | −17.8 (15.4) | 0.2 (−2.7 to 3.0) |
| **Long-term outcome (months 1–12)2,3** |   |            |   |            |          |
| PANSS total score        |   |            |   |            |          |
| Baseline                 | 155 | 78.7 (14.4) | 206 | 78.5 (14.4) |          |
| Change at month 12       | 116 | −25.8 (14.4) | 148 | −23.7 (18.2) | −2.3 (−5.6 to 1.0) |
| Change at end-point      | 155 | −20.4 (18.8) | 206 | −20.5 (20.3) | 0.2 (−3.4 to 3.8) |
| Positive symptoms        |   |            |   |            |          |
| Baseline                 | 155 | 22.0 (5.1)  | 206 | 22.3 (6.4)  |          |
| Change at month 12       | 116 | −8.1 (5.0)  | 148 | −7.3 (6.1)  | −0.9 (−2.0 to 0.2) |
| Change at end-point      | 155 | −6.8 (5.8)  | 206 | −6.5 (6.9)  | −0.4 (−1.5 to 0.7) |
| Negative symptoms        |   |            |   |            |          |
| Baseline                 | 155 | 20.2 (6.2)  | 206 | 19.8 (5.7)  |          |
| Change at month 12       | 116 | −6.1 (5.9)  | 148 | −5.7 (6.4)  | −0.5 (−1.6 to 0.6) |
| Change at end-point      | 155 | −4.7 (6.6)  | 206 | −4.8 (6.6)  | 0.3 (−0.7 to 1.4) |
| Disorganised thoughts    |   |            |   |            |          |
| Baseline                 | 155 | 18.1 (4.6)  | 206 | 18.0 (4.2)  |          |
| Change at month 12       | 116 | −5.5 (3.7)  | 148 | −4.7 (5.0)  | −0.9 (−1.7 to −0.0)* |
| Change at end-point      | 155 | −4.3 (4.8)  | 206 | −4.0 (5.2)  | −0.2 (−1.1 to 0.7) |
| Hostility/excitement     |   |            |   |            |          |
| Baseline                 | 155 | 7.7 (2.9)   | 206 | 8.1 (3.0)   |          |
| Change at month 12       | 116 | −2.4 (2.7)  | 148 | −2.4 (3.0)  | −0.3 (−0.8 to 0.1) |
| Change at end-point      | 155 | −1.6 (3.7)  | 206 | −1.8 (4.0)  | −0.1 (−0.8 to 0.6) |
| Anxiety/depression       |   |            |   |            |          |
| Baseline                 | 155 | 10.8 (3.4)  | 206 | 10.3 (3.5)  |          |
| Change at month 12       | 116 | −3.8 (3.1)  | 148 | −3.6 (3.6)  | 0.3 (−0.3 to 0.9) |
| Change at end-point      | 155 | −3.1 (3.6)  | 206 | −3.4 (3.7)  | 0.6 (0.1 to 1.2)* |

1 LSM, least squares mean; PANSS, Positive and Negative Syndrome Scale.
2 Least squares means. Short-term analysis: analysis of covariance with factors randomisation group, body mass index, number of previous hospitalisations, patient status and country × investigator and baseline score as covariate (type III sum of squares, SS). Long-term analysis: mixed-effects model with fixed effects of randomisation group, body mass index, number of previous hospitalisations, patient status, random effects for country and investigator, and baseline score as covariate (type III SS).
3 All changes in PANSS total and factor scores from baseline to month 12 and end-point were significant (P < 0.0001; paired t-test).
*P < 0.05.

regression model controlling for in-patient/ out-patient status, BMI, number of previous hospitalisations and investigator. At end-point, 79% of patients in the long-acting risperidone group and 73% in the olanzapine group achieved clinical improvement (P = 0.057; Fig. 3). Similar reductions in the overall severity of illness (CGI–S score) were seen in the long-acting risperidone and olanzapine groups: mean CGI–S scores at baseline were 3.1 (s.d. = 0.8) in the long-acting risperidone group and 3.3 (s.d. = 0.9) in the olanzapine group, and mean changes at end-point were −1.1 (s.d. = 1.2) and −1.3 (s.d. = 1.2) respectively. The proportions of patients who were rated as ‘not ill’ or ‘mildly ill’ increased respectively from 19 and 17% at baseline to 82 and 76% at month 12 and 66% and 67% at end-point. Mean scores on the patient version of the Wisconsin Quality of Life Index were similar in the two treatment groups at baseline: 0.40 (s.d. = 0.85) and 0.38 (s.d. = 0.92). Patients’ quality of life improved from baseline to end-point on all sub-scale ratings. Clinically meaningful improvements (score changes > 0.5 points) were seen in three domains in both treatment groups: occupational activities, psychological well-being and symptoms/outlook.

Maintenance of effect
The proportion of patients with significant deterioration in psychotic symptoms was 0.6% in the long-acting risperidone group and 2.0% in the olanzapine group at week 3; these respective proportions rose to 3% and 2% at week 13 and at month 12 and to 10% and 9% at end-point. The time to first deterioration was comparable in the
two groups (hazard ratio 1.38, 95% CI 0.82–2.33). Among the 179 patients who were stabilised at week 13, significant deterioration was noted in 3% of both the long-acting risperidone group and the olanzapine group at month 12, and in 5% and 6% respectively at end-point. The time to first deterioration was comparable in the two groups (hazard ratio 1.37, 95% CI 0.47–3.99).

**Safety**

**Adverse events**

Treatment-emergent adverse events reported by 5% or more of patients in either group are listed in Table 4. Adverse events resulted in treatment discontinuation for 7 patients in the long-acting risperidone group (3%) and 11 patients in the olanzapine group (4%). Serious adverse events were reported by 23% of the patients in the long-acting risperidone group and 21% of the olanzapine group (Table 4).

Adverse events related to extrapyramidal symptoms were reported by 25% of the long-acting risperidone group and 15% of the olanzapine group (P < 0.05; Table 5). Only one patient (in the long-acting risperidone group) discontinued treatment because of an extrapyramidal adverse event (hypokinesia). Severity of extrapyramidal symptoms was mild in both treatment groups. Median scores on the SARS – scores range from 0 (no symptom) to 4 (extreme) – were 0 at all time points in both treatment groups. At end-point, SARS total scores ranged from 0 to 1.5 in the long-acting risperidone group and from 0 to 1.7 in the olanzapine group. New-onset tardive dyskinesia was reported in two patients in each treatment group.

Treatment-emergent sexual side-effects were reported by 3% of the patients in each treatment group. The most common of these were non-puerperal lactation (in five patients in the long-acting risperidone group and two patients in the olanzapine group) and impotence (in two patients in each group). One patient in each group discontinued because of a sexual side-effect. Glucose-related adverse events were reported in 2% of patients in both the long-acting risperidone and olanzapine groups. These included diabetes mellitus in one patient in each group; hyperglycaemia in four patients in each group; and hypoglycaemia in one patient in the olanzapine group. No clinically relevant change in mean laboratory test values was seen in either treatment group.

**Deaths**

Eight patients died during the study or soon after its termination, two in the long-acting risperidone group and six in the olanzapine group. Causes of death were accident (n = 1) and oesophageal cancer (n = 1) in the long-acting risperidone group, and cardiac insufficiency/circulatory insufficiency (n = 1), status epilepticus/myocardial ischaemia, (n = 1) myocardial infarction (n = 1), pneumonia (n = 1), and suicide (n = 2) in the olanzapine group.

**Body weight**

Body weight increased by 1.7 kg in the long-acting risperidone group and by 4.0 kg in the olanzapine group (P < 0.05; Fig. 4). Body weight increases of 7% or more were seen in 20% of the long-acting risperidone group and 36% of the olanzapine group; decreases of 7% were seen in 6% of patients in both groups. Body mass index increased by 0.6 kg/m² in the long-acting risperidone group and by 1.4 kg/m² in the olanzapine group (P < 0.05). Six patients discontinued because of weight gain, one in the risperidone group and five in the olanzapine group.

**Patients receiving 75 mg long-acting risperidone**

The PANSS total and factor scores and adverse events in patients receiving 75 mg of long-acting risperidone are reported in Tables 6 and 7. In these patients, who had

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Adverse events reported by at least 5% of patients and serious adverse events reported by at least 2% of patients in either group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Long-acting risperidone</td>
</tr>
<tr>
<td>Adverse events</td>
<td>%</td>
</tr>
<tr>
<td>Psychosis</td>
<td>29</td>
</tr>
<tr>
<td>Insomnia</td>
<td>22</td>
</tr>
<tr>
<td>Depression</td>
<td>20</td>
</tr>
<tr>
<td>Anxiety</td>
<td>14</td>
</tr>
<tr>
<td>Agitation</td>
<td>10</td>
</tr>
<tr>
<td>Headache</td>
<td>8</td>
</tr>
<tr>
<td>Hyperkinesia</td>
<td>8</td>
</tr>
<tr>
<td>Extrapyramidal disorder</td>
<td>7</td>
</tr>
<tr>
<td>Rhinitis</td>
<td>7</td>
</tr>
<tr>
<td>Weight increase</td>
<td>6</td>
</tr>
<tr>
<td>Somnolence</td>
<td>5</td>
</tr>
<tr>
<td>Tremor</td>
<td>5</td>
</tr>
<tr>
<td>Injury</td>
<td>5</td>
</tr>
<tr>
<td>Serious adverse events</td>
<td>23</td>
</tr>
<tr>
<td>Psychosis</td>
<td>12</td>
</tr>
<tr>
<td>Suicide attempt</td>
<td>4</td>
</tr>
<tr>
<td>Anxiety</td>
<td>3</td>
</tr>
<tr>
<td>Injury</td>
<td>2</td>
</tr>
</tbody>
</table>

1. Adverse events reported by ≥5% of patients (risperidone group n = 218, olanzapine group n = 294).
2. Serious adverse events reported by ≥2% of patients (risperidone group n = 247, olanzapine group n = 300).

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Adverse events related to extrapyramidal symptoms reported in the two patient groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Long-acting risperidone (n = 247)</td>
</tr>
<tr>
<td>Hyperkinesia</td>
<td>8</td>
</tr>
<tr>
<td>Extrapyramidal disorder</td>
<td>8</td>
</tr>
<tr>
<td>Tremor</td>
<td>7</td>
</tr>
<tr>
<td>Hypertonia</td>
<td>4</td>
</tr>
<tr>
<td>Tardive dyskinesia</td>
<td>1</td>
</tr>
<tr>
<td>Dystonia</td>
<td>1</td>
</tr>
<tr>
<td>Involuntary muscle contractions</td>
<td>1</td>
</tr>
<tr>
<td>Oculogyric crisis</td>
<td>1</td>
</tr>
<tr>
<td>Dyskinesia</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Tetany</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Hypokinesia</td>
<td>0</td>
</tr>
</tbody>
</table>
Changes in body weight from baseline to month 12 and end-point in patients receiving long-acting risperidone before week 13.

Before week 13, some patients who received 75 mg of risperidone did not receive 75 mg long-term because the exact number of patients changes from the short-term to the long-term outcome. Baseline PANSS scores were not available for 1 patient.

### Table 6: Positive and Negative Syndrome Scale total and factor scores in patients receiving 75 mg of long-acting risperidone

<table>
<thead>
<tr>
<th>Short-term outcome (weeks 1–13)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PANSS total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>58</td>
<td>83.9 (16.2)</td>
</tr>
<tr>
<td>Change at end-point</td>
<td>58</td>
<td>−13.9 (20.0)</td>
</tr>
<tr>
<td>Long-term outcome (months 1–12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PANSS total</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>61</td>
<td>84.1 (15.8)</td>
</tr>
<tr>
<td>Month 12</td>
<td>16</td>
<td>63.9 (20.9)</td>
</tr>
<tr>
<td>Change at end-point</td>
<td>60</td>
<td>−12.3 (22.8)</td>
</tr>
<tr>
<td>Positive symptoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>61</td>
<td>24.9 (6.1)</td>
</tr>
<tr>
<td>Month 12</td>
<td>16</td>
<td>17.8 (8.6)</td>
</tr>
<tr>
<td>Change at end-point</td>
<td>60</td>
<td>−4.8 (8.2)</td>
</tr>
<tr>
<td>Negative symptoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>61</td>
<td>20.2 (7.1)</td>
</tr>
<tr>
<td>Month 12</td>
<td>16</td>
<td>16.8 (6.6)</td>
</tr>
<tr>
<td>Change at end-point</td>
<td>60</td>
<td>−2.4 (7.4)</td>
</tr>
<tr>
<td>Disorganised thoughts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>61</td>
<td>20.0 (4.6)</td>
</tr>
<tr>
<td>Month 12</td>
<td>16</td>
<td>16.1 (4.2)</td>
</tr>
<tr>
<td>Change at end-point</td>
<td>60</td>
<td>−2.8 (5.0)</td>
</tr>
<tr>
<td>Hostility/excitement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>61</td>
<td>8.6 (3.1)</td>
</tr>
<tr>
<td>Month 12</td>
<td>16</td>
<td>5.9 (2.5)</td>
</tr>
<tr>
<td>Change at end-point</td>
<td>60</td>
<td>−0.6 (4.7)</td>
</tr>
<tr>
<td>Anxiety/depression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>61</td>
<td>10.5 (2.9)</td>
</tr>
<tr>
<td>Month 12</td>
<td>16</td>
<td>7.3 (2.9)</td>
</tr>
<tr>
<td>Change at end-point</td>
<td>60</td>
<td>−1.7 (4.1)</td>
</tr>
</tbody>
</table>

PANSS, Positive and Negative Syndrome Scale

1. Baseline PANSS scores were not available for 1 patient. The number of patients changes from the short-term to the long-term outcome because some patients who received 75 mg of risperidone after week 13 did not receive 75 mg before week 13.

### Table 7: Adverse events reported in at least 5% of patients receiving 75 mg of long-acting risperidone (n = 71)

<table>
<thead>
<tr>
<th>Adverse event</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychosis</td>
<td>44</td>
</tr>
<tr>
<td>Insomnia</td>
<td>28</td>
</tr>
<tr>
<td>Anxiety</td>
<td>25</td>
</tr>
<tr>
<td>Depression</td>
<td>14</td>
</tr>
<tr>
<td>Suicide attempt</td>
<td>10</td>
</tr>
<tr>
<td>Agitation</td>
<td>7</td>
</tr>
<tr>
<td>Extrapyramidal disorder</td>
<td>11</td>
</tr>
<tr>
<td>Hyperkinesia</td>
<td>11</td>
</tr>
<tr>
<td>Headache</td>
<td>10</td>
</tr>
<tr>
<td>Tremor</td>
<td>6</td>
</tr>
<tr>
<td>Rhinitis</td>
<td>10</td>
</tr>
<tr>
<td>Pharyngitis</td>
<td>7</td>
</tr>
<tr>
<td>Constipation</td>
<td>6</td>
</tr>
<tr>
<td>Back pain</td>
<td>6</td>
</tr>
<tr>
<td>Flu-like symptoms</td>
<td>6</td>
</tr>
<tr>
<td>Weight increase</td>
<td>6</td>
</tr>
<tr>
<td>Amenorrhoea</td>
<td>6</td>
</tr>
</tbody>
</table>

Risperidone and olanzapine have been shown to be effective and generally well tolerated both in short-term (Marder & Meibach, 1994; Tolleson et al., 1997) and long-term (Csernansky et al., 2002; Beasley et al., 2003) trials of patients with schizophrenia and schizoaffective disorder. Three recent large studies have compared the oral formulations of the two agents. Conley & Mahmoud (2001) reported that the efficacy and safety of risperidone and olanzapine were generally similar in their double-blind, 8-week study. The only significant between-group differences were the greater improvements in the risperidone-treated patients on two of the five PANSS factors (positive symptoms and anxiety/depression) among patients who completed the trial, and the greater weight gain in the olanzapine-treated patients. Risperidone and olanzapine were among the five atypical

**Fig. 4**: Changes in body weight from baseline to month 12 and end-point in patients receiving long-acting risperidone or olanzapine. *P < 0.05 v olanzapine.
antipsychotics evaluated in the double-blind Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) trial of patients with schizophrenia (Lieberman et al, 2005). Time to treatment discontinuation (the primary outcome measure) was significantly longer in patients receiving olanzapine than risperidone (9.2 vs. 4.8 Kaplan–Meier estimated median months, \( P < 0.01 \)). However, similar improvements in PANSS total scores were seen in patients treated with risperidone and olanzapine at month 18, both in the total group (Lieberman et al, 2005) and in patients whose treatment had been switched to one of these two antipsychotics after discontinuing their previous treatments (Stroup et al, 2006). There was some suggestion that olanzapine was not as well tolerated as risperidone: substantial differences were noted in the proportions of patients who discontinued treatment because of intolerability (10% of the risperidone patients vs. 19% of the olanzapine patients; \( P < 0.05 \)) and 2% of the risperidone group vs. 9% of the olanzapine group discontinued because of weight gain or metabolic effects (\( P < 0.001 \), comparing all five treatment groups) (Lieberman et al, 2005). The 12-month results of the large, international open-label Intergroup Schizophrenia Outpatient Health Outcomes (IC-SOHO) study have been published recently (Dossenbach et al, 2005): similar proportions of patients in the risperidone and olanzapine groups responded to treatment during the 12 months (74% and 81%) or had relapsed (9% and 8%) (response and relapse were defined according to patient scores on the Clinical Global Impression–Schizophrenia scale).

In their meta-analysis of studies of atypical antipsychotics, Davis et al (2003) reported that the effect sizes of risperidone and olanzapine (compared with conventional antipsychotics) were similar (0.25 and 0.21 respectively) and highly significant (\( P < 0.001 \)). This analysis included data from 22 risperidone trials and 14 olanzapine trials.

The primary efficacy result of our trial was that in the short term (weeks 1–13) long-acting injectable risperidone was as effective as oral olanzapine in the treatment of patients with schizophrenia or schizoaffective disorder, an expected outcome given the previous findings of short-term studies of the oral formulations of the two agents.

Efficacy of the two treatments

Significant reductions in PANSS total and factor scores were seen in the analyses of the short-term and long-term data in both treatment groups. Patients receiving long-acting risperidone demonstrated significant benefits over treatment with olanzapine on two outcomes: clinical improvement (at least 20% reduction in PANSS total score) at month 12 and at end-point, and improvement on a PANSS factor at month 12 (disorganised thoughts). According to the patients’ ratings in both treatment groups, quality of life was improved from baseline to end-point.

Long-term outcomes

Figure 2 shows that the improvements with long-acting risperidone and olanzapine in PANSS total scores and scores on three of the five factors start to diverge at months 9–12, suggesting more positive long-term responses to long-acting risperidone than to olanzapine. A similar trend was evident in the data on clinical improvement (at least 20% reduction in PANSS total score). These results seem to be in line with those of a previous study (Hogarty et al, 1979), which reported comparable relapse rates with depot and oral antipsychotics (fluphenazine decanoate and fluphenazine hydrochloride) during the first year of treatment (39 and 35% respectively), but substantially lower rates with the depot medication than with the oral formulation during the second treatment year (8 and 42% respectively). The high medication adherence rates in this 1-year controlled study are noteworthy. The mean time off drug was 0.7 days (s.d.=3.7) in the oral olanzapine group, a substantially higher rate than reported in 1-year and 2-year studies of adherence rates in patients with schizophrenia receiving oral antipsychotics (Gilmer et al, 2004; Weiden et al, 2004). Thus, application of our findings to the real-world effectiveness of the two medications will need to take into account the impact of medication adherence rates on treatment outcome.

Tolerability

A high proportion of the patients completed the 1-year trial (65% of the long-acting risperidone group and 62% of the olanzapine group). Both treatments were safe and well tolerated. Few patients (7 in the risperidone group and 11 in the olanzapine group) withdrew from treatment because of an adverse event. The incidence of extrapyramidal adverse events was higher in the long-acting risperidone group than in the olanzapine group at baseline, but by months 9–12 the rates were comparable in the two groups (this does not appear to be a result of differential withdrawal rates). New-onset tardive dyskinesia (reported in two patients in each treatment group) was a rare event. Increases in body weight and BMI were significantly lower in the long-acting risperidone group than in the olanzapine group.

Implications

In patients with schizophrenia or schizoaffective disorder, long-acting risperidone and olanzapine tablets were efficacious and well tolerated over the 12-month duration of this study. The efficacy results suggest that in the long term patients might benefit more from treatment with long-acting risperidone than with oral olanzapine, but longer-term comparative data will help to confirm these observations.

ACKNOWLEDGEMENTS

The study was supported by Johnson & Johnson Pharmaceutical Research and Development. The authors thank Ilse Van Hove, MSc (Johnson & Johnson Pharmaceutical Research and Development, Beerse, Belgium) for completing the statistical analyses.

REFERENCES


Long-acting injectable risperidone: efficacy and safety of the first long-acting second-generation antipsychotic.


NICHOLAS A. KEKS, MB, BS, MPM, PhD, FRANZCP, Mental Health Research Institute of Victoria and Monash University, Victoria, Australia; MICHAEL INGHAM, MSc, AKBAR KHAN, MD, KEITH KARCHER, MSc, Johnson & Johnson Pharmaceutical Research and Development, Titusville, New Jersey, USA

Correspondence: Professor Nicholas A. Keks, Delmont Hospital, 298 Warrigal Road, Glen Iris, VIC 3146, Australia. Email: Nicholas.keks@med.monash.edu.au


Deployment-related stress and trauma in Dutch soldiers returning from Iraq

Prospective study

IRIS M. ENGLERHARD, MARCEL A. VAN DEN HOUT, JOS WEERTS, ARNOUD ARNTZ, JOOP J. C. M. HOX and RICHARD J. McNALLY

Background Some questionnaire studies have shown increased mental health problems, including probable post-traumatic stress disorder (PTSD), in soldiers deployed to Iraq.

Aims To test prospectively whether such problems change over time and whether questionnaires provide accurate estimates of deployment-related PTSD compared with a clinical interview.

Methods Dutch infantry troops from three cohorts completed questionnaires before deployment to Iraq (n=479), and about 5 months (n=382, 80%) and 15 months (n=331, 69%) thereafter. Post-traumatic stress disorder was evaluated by questionnaire and clinical interview.

Results There were no group changes for general distress symptoms. The rates of PTSD for each cohort were 21, 4 and 6% based on questionnaires at 5 months. The deployment-related rates of PTSD based on the clinical interview were 4, 3 and 3%.

Conclusions There was a specific effect of deployment on mental health for a small minority. Questionnaires eliciting stress symptoms gave substantial overestimations of the rate of PTSD.

Declaration of interest None. Funding detailed in Acknowledgements.

On 22 May 2003 the United Nations called on its member states to help reconstruct Iraq and over 30 countries have contributed soldiers to the coalition. As of 21 November 2006, there were 3113 deaths of coalition soldiers from 19 nations (http://edition.cnn.com/SPECIALS/2003/iraq/forces.casualties). Concerns have been raised about the mental health costs for service personnel. Hoge et al, 2004 estimated that 12.9% of US soldiers involved in combat operations in Iraq had post-traumatic stress disorder (PTSD), which was higher than in a sample tested before their deployment (5%). Hotopf et al (2006) estimated a 4% incidence of PTSD among UK armed forces returning from Iraq, which was similar to a non-deployed sample. Hacker Hughes et al (2005) revealed a lower score for mental health problems after deployment, compared with before, for the UK Air Assault Brigade but the response rate was low.

These studies provide clues about the impact of deployment to Iraq, but two studies were cross-sectional and did not include the participants’ health status pre-deployment. This may lead to an overestimation of the effects of deployment on stress symptoms (Hotopf & Wessely, 2005; Hedeker & Gibbons, 2006). Two recent prospective studies showed that PTSD was uncommon (<3.2%) for UK armed forces (Rona et al, 2006) and PTSD symptoms increased modestly for US armed forces (Vasterling et al, 2006) after deployment. These prior cross-sectional and prospective studies investigated PTSD using questionnaires, but the DSM–IV diagnostic criteria for PTSD require that symptoms interfere in important ways with the individual’s functioning, which is routinely checked in diagnostic interviews but not questionnaires. Failing to take this into account may result in overestimated rates of deployment-related PTSD (see Regier et al, 1998; Frueh et al, 2000; Ismail et al, 2002; Wessely, 2004; McNally, 2006). This was recently found in a re-analysis of PTSD among Vietnam veterans. In 1988 the estimated lifetime prevalence rate for PTSD was 30.9% and the current rate (11–12 years after the war) 15.2%. Dohrenwend et al (2006) consulted archival data and eliminated PTSD which was unrelated to war events and PTSD without impairment. This decreased estimates of lifetime and current (late 1980s) PTSD to 18.7 and 9.1% respectively, thereby confirming the suspicion of critics who believed the original rates to be implausibly high (McNally, 2007).

This paper reports a prospective study of deployment-related mental health problems in three Dutch infantry cohorts stationed in the Iraqi province of Al-Muthanna under British command. Mental health measures were collected before deployment, and 3 months and 15 months thereafter. We tested for individual changes in these variables over time as well as potential predictors for changes. To compare assessment methods, we established PTSD rates by questionnaire and clinical interview.

METHOD

Participants About 6 weeks before their deployment, 481 Royal Netherlands Army troops were asked to participate in this study. They were from three infantry battalions that rotated successively in three deployment phases, each lasting about 4 months, designated Stabilisation Force Iraq (SFIR) 3, 4 and 5. These took place from 15 March 2004 to 15 March 2005. During this period, about 4990 Dutch soldiers were deployed to Iraq. They sustained two casualties. An Armoured Infantry Battalion was deployed on SFIR 3, and two battalions of the Air Assault Brigade were deployed on SFIR 4 and 5. Their main duties were to create and maintain stability and peace, and assist in reconstruction. At various sites, troops available during their preparation programme were told about the aim and general procedures of the study by their commanding officers. They met the principal investigator (I.M.E.) or research assistant a few days later, who gave full (oral and written) information about the study. Participation was voluntary without financial compensation. Participants were told that commanders would be informed only about pooled results. Two soldiers refused and 479 agreed to participate.
including 214 SFIR 3 soldiers, 169 SFIR 4 soldiers and 96 SFIR 5 soldiers.

**Procedures**

About 5 months after their deployment, questionnaires about potentially traumatic events in Iraq and current mental health problems were administered. The Structured Clinical Interview for DSM-IV (SCID; First et al., 1996) was then administered by a trained clinical psychologist (about 20 interviews were carried out by graduate students), who did not know the responses to the questionnaires. At about 15 months, measures for current problems were administered as well as the face-to-face SCID. At 5 months, most questionnaires were given to small groups at a base, and at 15 months 31% of questionnaires were sent by post. Non-response was partly a result of soldiers being on leave, attending a training course, or being posted to new units. The institutional review board of Maastricht University approved the study.

**Measures**

Symptoms of common mental health problems were measured with the 90-item Symptom Checklist (SCL–90; Arrindell & Eterna, 2003). Each item was rated on a 1 (not at all) to 5 (very much) scale. We focused on sub-scales of anxiety (10 items), depression (17 items), somatic complaints (12 items) and sleeping problems (3 items), and used the SCL–90 score for general distress. Prior life events were assessed with a 17-item checklist that included road accidents, sudden death of a loved one, fire and being robbed. A score was compiled of all endorsed items.

Events in Iraq were assessed with the 21-item Potentially Traumatizing Events Scale (Litz et al., 1997; Maguen et al., 2004), which is derived partly from the Combat Exposure Scale, and measures the frequency of exposure to war-zone-related stressors. The scale was adjusted for use in Iraq by deleting one item (patrolling areas where there were land mines) and adding two (being informed of a Dutch soldier who got killed and having injured civilians due to own action). For each item experienced, individuals rated how negative it was for them on a Likert scale. We calculated the number of reported events as well as the number of events appraised as moderate to extremely negative, and used both in the analyses.

Symptoms of PTSD were measured with the 17-item PTSD Symptom Scale (PSS), which has proven to be effective for screening for PTSD (Foa et al., 1993; Wohlfarth et al., 2003; Coffey et al., 2006). Each symptom was rated from 0 (not at all) to 3 (very much) for the past month. We used two case definitions for PTSD: a broad definition that follows diagnostic symptom criteria (a minimal number of symptoms had to be rated at least ‘some of the time’), and a stricter definition for which a cut-off score of 14 was used (Coffey et al., 2006). Both scoring methods have been used previously. After completing the PSS, participants were asked to rate their distress and functional impairment in different areas of their lives (work, home, interpersonal relationships) on a 4-point scale (0 not at all, 3 very much). Self-reports of impairment were compared with SCID-based assessments which include the DSM-IV symptoms and questions about subjective distress and functional impairment caused by these symptoms.

**Statistical analysis**

Analyses were performed with SPSS (version 11.5) and HLM (version 5), both for Windows. Missing items were estimated by observed item means if no more than two items per scale were missing. Demographic variables were compared between the cohorts. We tested whether there were systematic differences between individuals who did or did not drop out on characteristics from earlier measurement occasions. The rates of potentially traumatic events were assessed and the mean scores on mental health scales before deployment and at 5 and 15 months were calculated. These variables were non-normally distributed and therefore statistical tests were based on robust standard errors (sandwich estimates; White, 1982). A within-class hierarchical linear model was used to test whether the level of mental health symptoms varied across the three assessments for the three cohorts. The slopes of the time variables were allowed to vary across individuals, to test whether the variance components for the intercept and the regression slope for the time variables were significant. If these were significant, we sought to explain this variance by running separate between-cohort models to examine the effects of demographic and background factors (age, gender, partner status, education, temporary v. permanent contract, number of previous missions, previous life events, rank, cohort), number of potentially traumatic events in Iraq, and number of events rated as negative. The model was run again without non-significant predictors to reduce error. Rates of PTSD were calculated based on the questionnaire and SCID. We tested to what extent the PTSD rates were predicted by pre-deployment PTSD and general distress symptoms, prior life events, and number of events in Iraq by logistic regression analysis. Odds ratios with 95% confidence intervals were generated. All statistical tests of significance were two-tailed at the $\alpha=0.05$ level.

**RESULTS**

Table 1 shows the characteristics of the three cohorts. The SFIR 3 and 5 cohorts were older than SFIR 4 and had more prior deployments. The SFIR 5 cohort had more officers. About 5 months after deployment, 382 (80%) soldiers completed questionnaire and SCID data at 15 months are not considered. Soldiers who did not complete questionnaires at 5 months had slightly more prior missions than those who did. There were no significant differences on other pre-deployment measures. Responders and non-responders at 15 months did not differ on variables assessed earlier. We therefore treated the drop-outs as occasions missing at random.

Table 2 shows that cohorts SFIR 3 and 4 reported more potentially traumatic events in Iraq than SFIR 5 ($P(2, 370)=60.73, P<0.001$). The number of events rated as negative was higher for SFIR 3 and 5 than for SFIR 4 ($\chi^2(2)=14.65, P=0.001$).

Table 3 indicates scores on mental health scales before and after deployment. The mean levels of anxiety, depression, somatic complaints, sleeping problems and general distress did not vary over time for the cohorts. The variance components for the intercept and the regression slope for the time variables were significant for all mental health scales, which means that individuals had different initial states as well as different rates of change. We sought to explain this variance. The final model for this analysis showed that prior life events and the SFIR 3 cohort (Armoured Infantry
both

naire rates dropped (to 17, 4 and 0% using the stricter cut-off score, the questionnaires who completed the SCID rates. This was similar for mate was about 2 times higher than the un-mated cohort compared with the Armoured Infantry cohort. 

Stabilisation Force Iraq.

1. Some data were missing for 1 person from SFIR 3 and 2 persons from SFIR 5.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>SFIR 3 (n = 214)</th>
<th>SFIR 4 (n = 169)</th>
<th>SFIR 5 (n = 96)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years: mean (s.d.)</td>
<td>23.06 (4.49)</td>
<td>21.63 (3.54)</td>
<td>22.73 (4.02)</td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>204 (95)</td>
<td>169 (100)</td>
<td>94 (98)</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>167 (78)</td>
<td>132 (78)</td>
<td>73 (78)</td>
</tr>
<tr>
<td>Married/cohabiting</td>
<td>47 (22)</td>
<td>37 (22)</td>
<td>21 (22)</td>
</tr>
<tr>
<td>Education, n (%)</td>
<td>14 (7)</td>
<td>3 (2)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>Primary school</td>
<td>194 (91)</td>
<td>162 (96)</td>
<td>88 (92)</td>
</tr>
<tr>
<td>Secondary school</td>
<td>5 (2)</td>
<td>3 (2)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>College</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rank, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-commissioned officer</td>
<td>15 (7)</td>
<td>16 (10)</td>
<td>16 (17)</td>
</tr>
<tr>
<td>Commissioned officer</td>
<td>2 (1)</td>
<td>2 (1)</td>
<td>4 (4)</td>
</tr>
<tr>
<td>Other rank</td>
<td>196 (92)</td>
<td>151 (89)</td>
<td>76 (79)</td>
</tr>
<tr>
<td>Prior missions, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>131 (61)</td>
<td>121 (72)</td>
<td>51 (53)</td>
</tr>
<tr>
<td>1</td>
<td>52 (24)</td>
<td>28 (17)</td>
<td>29 (30)</td>
</tr>
<tr>
<td>2–4</td>
<td>31 (15)</td>
<td>20 (11)</td>
<td>16 (17)</td>
</tr>
<tr>
<td>Contract status, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Permanent</td>
<td>16 (8)</td>
<td>6 (4)</td>
<td>17 (18)</td>
</tr>
<tr>
<td>Temporary</td>
<td>197 (92)</td>
<td>163 (96)</td>
<td>79 (82)</td>
</tr>
</tbody>
</table>

SFIR, Stabilisation Force Iraq.

DISCUSSION

Main findings

The main findings of the study were that (a) levels of general distress remained relatively stable from before to after deployment; (b) the Armoured Infantry cohort showed a higher PTSD estimate by questionnaire than the Air Assault Brigade cohorts owing to more pre-deployment symptoms and more reported harmful exposure in Iraq; (c) unadjusted PTSD rates from diagnostic interview were 41% lower than estimates from the questionnaire; (d) interview rates of PTSD were nearly halved after adjustment for PTSD which was unrelated to deployment and functional impairment; (e) a small minority of soldiers showed full-blown deployment-related PTSD.

The deployment-related rates of PTSD were much lower when the SCID was used as a diagnostic tool. The questionnaire could have led to inflated rates of symptoms because several individuals endorsed symptoms stemming from traumatic events which were unrelated to deployment (e.g., death of a relative), or from stressful but seemingly non-catastrophic events. Other studies have also linked non-traumatic life events to PTSD symptoms (see McNally, 2003). Moreover, PTSD questionnaires assess symptoms during the past month and fail to control for pre-existing stress and psychopathology. The PTSD arousal symptoms (e.g., difficulty sleeping, irritability, concentration problems) are not specific for the disorder, and may very well have been present before deployment (Clark et al, 1994).

Other studies

The questionnaire-based estimate of PTSD after deployment to Iraq has been documented in a few previous studies, but comparison is limited by differences in populations studied, sampling and response rates. Hoge et al (2004) reported high levels of combat exposure in US infantry soldiers and used a broad symptom-based definition that resulted in a PTSD estimate of 18% 3–4 months after their return from Iraq. The PTSD estimate in our SFIR 3 (Armoured Infantry) cohort is similar (17–21%). Hotopf et al (2006) reported a rate of 4% in a random UK military sample that reported less trauma exposure. Hacker Hughes et al (2005) found that the PTSD estimate was 2% in a sample of the UK Air Assault Brigade. This is in the range we found for
Various reasons have been proposed for the different outcomes of US and UK studies, including differences in trauma severity and healthcare systems (Hotopf et al., 2006).

Implications
Can the present findings be extrapolated to other types of trauma and civilian populations? We do not know whether our PTSD rates may be generalised to the military populations studied: although our response rates were exceptionally high, the convenience sampling method was less desirable than random sampling. However, there are no empirical or theoretical reasons to assume that the pattern of results would be different for different samples or populations. Population-level screening for PTSD is important to identify healthcare needs.

Table 2  Potentially traumatic events experienced by soldiers in Iraq.

<table>
<thead>
<tr>
<th>Item</th>
<th>Item experienced, %</th>
<th>Item rated as moderately to extremely negative, %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SFIR 3 (n=170)</td>
<td>SFIR 4 (n=140)</td>
</tr>
<tr>
<td>Fear of being ambushed or attacked</td>
<td>97</td>
<td>96</td>
</tr>
<tr>
<td>Being informed of a Dutch soldier who got killed</td>
<td>95</td>
<td>97</td>
</tr>
<tr>
<td>Fear of having unit fired on</td>
<td>94</td>
<td>96</td>
</tr>
<tr>
<td>Going on patrols or performing other dangerous duties</td>
<td>90</td>
<td>99</td>
</tr>
<tr>
<td>Witnessing violence</td>
<td>83</td>
<td>89</td>
</tr>
<tr>
<td>Witnessing an explosion</td>
<td>75</td>
<td>73</td>
</tr>
<tr>
<td>Fear that you might be taken hostage</td>
<td>72</td>
<td>69</td>
</tr>
<tr>
<td>Disarming civilians</td>
<td>68</td>
<td>59</td>
</tr>
<tr>
<td>Needing to manage civilians in chaotic conditions</td>
<td>64</td>
<td>79</td>
</tr>
<tr>
<td>Being shot at</td>
<td>62</td>
<td>66</td>
</tr>
<tr>
<td>Patrolling through the zone of separation</td>
<td>55</td>
<td>54</td>
</tr>
<tr>
<td>Seeing dead or injured Dutch soldiers</td>
<td>50</td>
<td>51</td>
</tr>
<tr>
<td>Seeing dead or injured civilians</td>
<td>42</td>
<td>49</td>
</tr>
<tr>
<td>Having injured civilians by own action</td>
<td>34</td>
<td>29</td>
</tr>
<tr>
<td>Seeing human remains</td>
<td>29</td>
<td>23</td>
</tr>
<tr>
<td>Seeing dead or injured NATO (non-Dutch) soldiers</td>
<td>27</td>
<td>26</td>
</tr>
<tr>
<td>Having to aid in the removal of unexploded ordnance</td>
<td>20</td>
<td>15</td>
</tr>
<tr>
<td>Locating unexploded land mines</td>
<td>19</td>
<td>24</td>
</tr>
<tr>
<td>Being injured because of an accident</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>Being injured because of an assault/attack</td>
<td>14</td>
<td>11</td>
</tr>
<tr>
<td>Having to aid in the removal of human remains</td>
<td>13</td>
<td>16</td>
</tr>
<tr>
<td>Experienced sexual harassment during the deployment</td>
<td>9</td>
<td>8</td>
</tr>
</tbody>
</table>

Air Assault Brigade cohorts (0–6%). For some, psychological symptoms may actually decrease, which has been shown previously in a UK study (Hacker Hughes et al., 2005). The lack of change over time after deployment in our study is in line with other UK research (Hotopf et al., 2006), but not with the study of Hoge et al. (2006) which suggested that rates of PTSD increase in the months after deployment.

Table 3  Scores on mental health scales for the three cohorts

<table>
<thead>
<tr>
<th>Item</th>
<th>Before deployment</th>
<th>5 months after deployment</th>
<th>15 months after deployment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SFIR 3 (n=214)</td>
<td>SFIR 4 (n=169)</td>
<td>SFIR 5 (n=96)</td>
</tr>
<tr>
<td></td>
<td>SFIR 3 (n=170)</td>
<td>SFIR 4 (n=140)</td>
<td>SFIR 5 (n=72)</td>
</tr>
<tr>
<td></td>
<td>SFIR 3 (n=152)</td>
<td>SFIR 4 (n=120)</td>
<td>SFIR 5 (n=59)</td>
</tr>
<tr>
<td>SCL–90 Anxiety score: mean (s.d.)</td>
<td>11.6 (3.0)</td>
<td>10.9 (1.7)</td>
<td>10.7 (1.6)</td>
</tr>
<tr>
<td></td>
<td>11.2 (3.3)</td>
<td>10.7 (1.8)</td>
<td>10.6 (1.4)</td>
</tr>
<tr>
<td></td>
<td>11.4 (3.1)</td>
<td>11.0 (2.6)</td>
<td>10.2 (0.5)</td>
</tr>
<tr>
<td>Depression score: mean (s.d.)</td>
<td>18.9 (4.8)</td>
<td>17.7 (3.1)</td>
<td>17.6 (2.6)</td>
</tr>
<tr>
<td></td>
<td>19.0 (5.9)</td>
<td>18.0 (4.7)</td>
<td>17.4 (2.9)</td>
</tr>
<tr>
<td></td>
<td>19.2 (6.8)</td>
<td>18.1 (5.8)</td>
<td>16.7 (1.4)</td>
</tr>
<tr>
<td>Somatisation score: mean (s.d.)</td>
<td>14.4 (3.5)</td>
<td>13.3 (2.3)</td>
<td>13.6 (2.2)</td>
</tr>
<tr>
<td></td>
<td>14.0 (3.2)</td>
<td>13.4 (3.0)</td>
<td>13.4 (2.1)</td>
</tr>
<tr>
<td></td>
<td>14.3 (4.5)</td>
<td>13.5 (3.0)</td>
<td>12.8 (1.7)</td>
</tr>
<tr>
<td>Sleeping problems score: mean (s.d.)</td>
<td>3.9 (1.9)</td>
<td>3.4 (1.1)</td>
<td>3.4 (0.8)</td>
</tr>
<tr>
<td></td>
<td>4.2 (2.2)</td>
<td>3.4 (1.2)</td>
<td>3.6 (1.5)</td>
</tr>
<tr>
<td></td>
<td>4.2 (2.3)</td>
<td>3.3 (0.7)</td>
<td>3.2 (0.7)</td>
</tr>
<tr>
<td>PSS symptom severity score: mean (s.d.)</td>
<td>3.0 (3.8)</td>
<td>N/A</td>
<td>3.0 (4.4)</td>
</tr>
<tr>
<td></td>
<td>6.5 (7.1)</td>
<td>3.7 (4.4)</td>
<td>2.1 (3.0)</td>
</tr>
<tr>
<td></td>
<td>5.0 (6.9)</td>
<td>2.8 (5.0)</td>
<td>1.2 (2.1)</td>
</tr>
</tbody>
</table>

SCL–90, 90-item Symptom Checklist; PSS, Post-traumatic Stress Disorder Symptom Scale; NA, not available.
and self-report measures such as the PSS have reasonable sensitivity and specificity (see Coffey et al., 2006). However, when PTSD rates are low, as in the present cohort, it would be unwise to implement questionnaire-based screening (see also Rona et al., 2006). For clinical purposes, the rate of false-positives is less of a concern because the initial questionnaires can be followed by a comprehensive diagnostic interview. Such a two-step approach has been recommended to identify PTSD (Shrout et al., 1986). However, in epidemiological studies, the impact of traumatic events on mental health is often determined by merely examining whether or not participants meet symptom criteria for PTSD on a questionnaire. Our findings suggest that many people screened positive for PTSD may actually not have the disorder. Other studies have also shown that questionnaires provide higher estimates of ill health than clinical assessment: symptoms may signify disorder, but then again they might not (see Frueh et al., 2000; Wessely, 2004; Hotopf & Wessely, 2005). Healthcare needs might be much lower than expected on the basis of questionnaires.

A striking finding was that for 36.8% of soldiers showing the full PTSD symptom pattern on the SCID, symptoms did not produce more than slight impairment in their lives. This is in line with a re-analysis of PTSD among Vietnam veterans, in which individuals exposed to traumatic events and who experienced minimal symptoms for the diagnosis might very well have qualified as having PTSD despite living well-adjusted, productive lives (Dohrenwend et al., 2006). It also fits well with a recent major re-analysis of two large US community surveys in which previously unused data on the clinical significance of symptoms were used to recalculate prevalence rates of mental disorder. Prevalence rates of any disorder were lowered by 17 and 32% (depending on the survey; Narrow et al., 2002). Hoge et al. (2004) also recalculated the PTSD rates after combat in Iraq on the basis of functional impairment or greater severity, and found that this decreased symptom-based PTSD rates by nearly 30% (from 18 to 12.9%). The DSM–IV classification system requires this functional impairment to differentiate symptoms from disorder, but population-based studies typically do not consider this criterion. Unfortunately, the system lacks objective criteria to determine impairment, and more work should be done to define when impairment becomes clinically significant.

**Limitations**
The present study does not provide information about the natural course of PTSD and functional impairment extending over 15 months. Despite considerable effort, the sample interviewed at follow-up was small, owing largely to the high turnover of personnel. This was also responsible for reduced sample sizes in previous studies (Hotopf & Wessely, 2005). Longitudinal evaluations of civilian populations suggest that PTSD symptoms decrease substantially within the first year, but little is known about the predictive validity of mild and sub-threshold forms of PTSD for soldiers on active duty. There is evidence from the National Comorbidity Study that a proportion of people with ‘mild’ mental disorders had worse clinical outcomes up to 10 years later (Kessler et al., 2003). Similar longitudinal studies are needed for military populations. Some active soldiers may not experience functional impairment until they leave the military. If such predictive validity were found, the need for interventions might be considered on the basis of functional impairment as well as the risk of progression from a mild to a more severe disorder. This could have great public health importance. Clearly, these issues await future research.

There are some issues of sample size and power in our study. The use of a complex multilevel analysis and the presence of attrition is an obstacle for estimating power for the analysis method used. However, a one-way analysis of variance at the first measurement (before deployment) with a sample size of 479 achieves a power of 0.99 for medium effects and 0.48 for small effects. Analysis at 15 months with a sample size of 331 achieves a power of 0.98 for medium effects and 0.33 for small effects. The multilevel analysis should have at least this much power and the power to detect medium size effects should be very high. However, these results suggest that the study has only weak power for detecting small effects.

**Future directions**
Understanding PTSD from an epidemiological perspective is vital for estimating the likely need for healthcare services and information. This study shows that some individuals meet the PTSD symptom criteria but lead productive lives despite stress, and that some PTSD is triggered by causes unrelated to deployment. Not considering these aspects leads to inflated rates of deployment-related PTSD.

**ACKNOWLEDGMENTS**
We thank Mariëtte van Baar and Latte Barmels for assistance with data collection, Erik Schouten for help with analyses, and the commanders and troops for their time and effort. This study was supported by a grant from the Veterans Institute (Doorn, The Netherlands) and an award by the Netherlands Organisation for Scientific Research to Iris P. Engelhard (Innovational Research Incentive VENI scheme 016.045.06). We thank (representatives of) the Netherlands Ministry of Defense for their cooperation, and in particular the Afdeling Individuele Hulpverlening (AIH) and Col MD Kees Ijzerman.
REFERENCES


Social identification and post-traumatic stress symptoms in post-conflict Northern Ireland†

ORLA T. MULDOON and CIARA DOWNES

Background Understanding of the psychological impact of politically motivated violence is poor.

Aims To examine the prevalence of post-traumatic symptoms subsequent to the ‘troubles’ in Northern Ireland.

Method A telephone survey of 3000 adults, representative of the population in Northern Ireland and the border counties of the Irish Republic, examined exposure to political violence, post-traumatic stress disorder (PTSD) and national identity.

Results Ten per cent of respondents had symptoms suggestive of clinical PTSD. These people were most likely to come from low-income groups, rate national identity as relatively unimportant and have higher overall experience of the ‘troubles’ than other respondents.

Conclusions Direct experience of violence and poverty increase the risk of PTSD, whereas strong national identification appears to reduce this risk.

Declaration of interest None. Funding detailed in Acknowledgements.

Although war and political conflict have grave consequences, increased national identification and community solidarity during wartime appear to protect mental health. Moreover epidemiological studies in Northern Ireland have indicated comparatively good mental health of the population during the 35-year period of political violence that has affected the region (Cairns et al, 2003), colloquially known as the ‘troubles’. Given the chronic nature of the conflict, the scale of casualties in terms of total population (3300 fatalities from a population of 1.68 million between 1969 and 1998), the effects of the ‘troubles’ have been widely felt (Hayes & McAllister, 2001), and like other conflicts, the impact has not been distributed evenly across the population (Cairns, 1996). Worldwide, those most likely to be affected by conflict are the poorest (World Health Organization, 2002) and within affected countries, those reporting the most experience of violence tend also to be the most socially disadvantaged (Bryce et al., 1989; Muldoon & Trew, 2000).

The most common psychological consequence of war and conflict is post-traumatic stress disorder (PTSD). To date only a limited number of epidemiological studies have examined the prevalence of PTSD post-conflict (De Girolamo & McFarlane, 1996). However, these prevalences are often higher than those in countries where conflict is ongoing (De Jong et al., 2003). The course of PTSD may well be linked to community and group identity, as is the stress process (Haslam & Reicher, 2006). In particular, the very high variability in levels of post-traumatic stress in referred and clinical samples in Northern Ireland might in part be attributable to social identity. For instance, Wilson et al (1997) found an incidence of 5% of probable PTSD in police officers exposed to life-threatening incidents during the ‘troubles’ whereas Daly & Johnston (2002) reported 67% among those held at gunpoint in a bar towards the end of the ‘troubles’. This comparatively low rate among police officers indicates the value of a consolidated identity to preserving mental health. The Royal Ulster Constabulary (RUC), the police force in Northern Ireland during the ‘troubles’, was strongly identified with one community and officers were highly committed to its identity (Mulcahy, 2006). However, those exposed in the bar incident were bystanders and the 1994 ceasefire had led many to believe the conflict was over.

The ability to cope with stress is intrinsically related to psychological and material resources (Lazarus & Folkman, 1984), which are likely to be adversely affected by repeat traumatisation experienced during politically motivated conflict. Experience and appraisal of trauma tends to be related to both poverty (Muldoon, 2003) and social identity (Haslam et al., 2004).

The aims of this study were first to examine the population prevalence of PTSD in Northern Ireland post-conflict and to examine the relationship between PTSD and the strength of national identification. Second, although a comparatively affluent society, deprivation within the region remains a significant social issue and therefore we examined PTSD across socio-economic groups. Finally, lifetime experience of violence was assessed to determine the relationship between chronic traumatisation and PTSD.

METHOD Sample
A random sample of household telephone numbers was drawn from domestic listings for Northern Ireland and the Republic of Ireland. These numbers were matched with the relevant postal address and a letter was sent to selected households, explaining the nature and purpose of the study. Each household was then contacted by telephone. Where more than one adult resided in a household, the last birthday technique was used to randomise the selection of respondents included in the sample.

The survey was carried out using computer-assisted telephone interviewing, which facilitates interview monitoring via listening in facilities. A quota control mechanism controlled the number of respondents by location based on adult population statistics from the latest census (2001 Northern Ireland, 2002 Republic of Ireland).
experience of the ‘troubles’ were included:
one asked whether respondents viewed
themselves as a victim of the ‘troubles’
(Cairns et al, 2003); a final question asked
whether they had used alcohol, prescription
or other drugs to cope with their experi-
ences (Bleich et al, 2003).

Ethical considerations
Participants were given details of the re-
search in writing when invited to partici-
pate. The confidentiality and anonymity
of all responses was assured; participants
were also given the opportunity to refuse
to participate and/or to withdraw at any
time. A free-phone number where trained
counsellors were available to discuss issues
arising from the interview was provided at
the end of all interviews. This service was
active for 6 months from the start of the
project. No calls were received at this
number and no participant requested
counselling via this system.

RESULTS
Of the 3000 respondents, 1269 (42%) re-
ported experience of a distressing event as
a result of the ‘troubles’ and thus were as-
sessed for PTSD with the PTSD Checklist.
Based on standard cut-off scores (Walker
et al, 2002) 10% of respondents (n=299)
had symptoms severe enough to warrant a
diagnosis of PTSD. Of these 299 people,
239 were from Northern Ireland (12% preva-
ence) and 60 from the border coun-
ties of the Irish Republic (6% prevalence).
This difference was significant (χ²=13.92,
d.f.=1, P<0.01). No gender or religious
differences were observed.

Characteristics of those with PTSD
Those classified as having PTSD were less
likely to have third-level education (20 v.
31%; χ²=19.4, d.f.=7, P<0.01) and were
more likely to be unemployed owing to
job loss (4.3 v. 1.6%) or unable to work
owing to illness (6.7 v. 1.3%; χ²=29.4,
d.f.=8, P<0.01). People with PTSD were
more likely to be in unskilled, partly skilled
or manual occupations (10.6 v. 6.1%, 16.6
v. 13.5%, 13.8 v. 11.1% respectively; χ²
=14.2, d.f.=5, P<0.01) (6.1%, 13.5% and
11.1% respectively). People with PTSD
also reported lower average household
incomes. In Northern Ireland, 33% of re-
pondents with probable PTSD had a
household income of less than £20 000
and 14% had an income of less than
£10 000 per annum. In comparison, 24%
of households overall reported an income
of less than £20 000, with only 6% with
an income less than £10 000. In the Repub-
lic, 32% of people with PTSD lived in a

Table 1 Sample profile according to gender, age and jurisdiction

<table>
<thead>
<tr>
<th></th>
<th>Northern Ireland</th>
<th>Republic of Ireland</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>848</td>
<td>42</td>
<td>459</td>
</tr>
<tr>
<td>Female</td>
<td>1152</td>
<td>58</td>
<td>541</td>
</tr>
<tr>
<td>Age group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–24 years</td>
<td>130</td>
<td>7</td>
<td>67</td>
</tr>
<tr>
<td>25–44 years</td>
<td>744</td>
<td>37</td>
<td>309</td>
</tr>
<tr>
<td>45–64 years</td>
<td>737</td>
<td>37</td>
<td>456</td>
</tr>
<tr>
<td>65+ years</td>
<td>389</td>
<td>20</td>
<td>168</td>
</tr>
<tr>
<td>Total</td>
<td>2000</td>
<td>100</td>
<td>1000</td>
</tr>
</tbody>
</table>

Table 2 Substance use to help with experiences related to the ‘troubles’ among those classified with and without probable PTSD according to the PTSD Checklist

<table>
<thead>
<tr>
<th>Substance</th>
<th>Probable PTSD (n=60)</th>
<th>Without PTSD (n=1029)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>(%)</td>
</tr>
<tr>
<td>Alcohol</td>
<td>38</td>
<td>(12.7)</td>
</tr>
<tr>
<td>Prescribed drugs</td>
<td>47</td>
<td>(15.7)</td>
</tr>
<tr>
<td>Other drugs</td>
<td>13</td>
<td>(4.3)</td>
</tr>
</tbody>
</table>

PTSD, post-traumatic stress disorder.
DISCUSSION

Main findings

The prevalence of probable PTSD in Northern Ireland after a period of protracted political conflict is approximately 10%. This is higher than that observed in police officers exposed to life-threatening incidents during the same ‘troubles’, who have been reported to have a strong sense of shared identity (Mulcahy, 2006). Similarly, the weaker national identities of people with probable PTSD suggests that social identity can protect mental health in situations of violence, in accordance with the integrated social identity model of stress (Haslam & Reicher, 2006). In conflict situations, identities underpin the conflict (Kelman, 1999) and consequently deliberate attempts to reduce the salience of these identities post-conflict (MacGinty et al, 2007) might inadvertently affect mental health. Clearly, longitudinal studies are needed to explore any such effects more fully.

People with PTSD more frequently reported using alcohol to cope with their experience of the ‘troubles’ (12.7% v. 2.3%; \( \chi^2 = 48.785, \) d.f. = 1, \( P < 0.01 \)). Similarly, 15.7% reported using prescribed medication to cope with the ‘troubles’ compared with 1.5% of the other respondents (\( \chi^2 = 95.801, \) d.f. = 1, \( P < 0.01 \)). Finally six times as many people with PTSD reported the use of other drugs to cope with the ‘troubles’ (4.3% v. 0.8%; \( \chi^2 = 11.361, \) d.f. = 1, \( P < 0.01 \)).

People with PTSD reported more direct (\( F(1, 1267) = 149, \ P < 0.001 \)) and indirect experiences (\( F(1, 1267) = 85, \ P < 0.001 \)) of the ‘troubles’ (Fig. 1). The importance attached to national identity was also related to PTSD. People with PTSD rated their national identity as less important (\( F(1, 1138) = 6.78, \ P < 0.01 \)). Perceived victimhood was also related to PTSD symptoms (\( \chi^2 = 171, \) d.f. = 4, \( P < 0.001 \)). Only 9% of respondents often or very often considered themselves victims of the ‘troubles’; however, 24% of those with PTSD stated that they often or very often considered themselves to be a victim of the ‘troubles’. On the other hand and perhaps more surprisingly, 46% of those with symptoms severe enough to suggest clinically significant PTSD never or rarely considered themselves victims of the ‘troubles’ (Fig. 2).

The observed prevalence of PTSD is similar to that in other regions affected by long-term conflict, such as Israel and Sri Lanka, and higher than that observed subsequent to acute incidents such as the 9/11 attacks in the USA (De Jong et al, 2003). Although there are clear differences between both the situations and the studies, overall the incidence of PTSD would appear to be higher in situations of ongoing or chronic political violence rather than subsequent to acute incidents. Similarly, we found that those with PTSD were more likely to report multiple direct and indirect experiences, with direct experience appearing to have a more powerful impact. This provides further evidence that previous exposure needs to be considered when evaluating the relative impact of traumatic events in situations of war and violence, not least because of the resource-depleting effects of multiple traumaisation.

However, half of our respondents reported that they had encountered no particularly distressing incident during the ‘troubles’. The impact of conflict is therefore not distributed evenly – some have suffered not at all and others have suffered greatly. Respondents of lower socio-economic status were disproportionately affected by PTSD. Although symptoms might contribute to disadvantage (as a result of disability and unemployment), the fact that many had low educational status suggests that social disadvantage increases the risk of developing PTSD. Of course, social disadvantage might also increase the risk of engaging with the conflict (Cairns, 1996), thereby increasing the risk of exposure to trauma. In reality, the coincidence of deprivation and multiple traumatisations in situations of political violence are likely to be twin, inextricably linked risks. That said, those identified as having probable PTSD represent a particularly vulnerable and disadvantaged group in terms of financial, psychological and social capital.

Methodological limitations

Although our sample was comparable to the general population in Northern Ireland, no details are available regarding the mental health status of non-respondents. Reluctance to participate is reflective of the ‘whatever you say, say nothing’ approach to engaging in any contentious discourse which is evident in many societies with conflict (Cairns, 1996). Indeed a similar Israeli study achieved a 57% response rate (Bleich et al, 2003). The limited verification of respondents’ accounts of distressing events is also important to the interpretation of the findings. Although the greater prevalence of direct and indirect experience in people with probable PTSD provides a form of verification through triangulation, it is possible that respondents did not actually experience life-threatening events personally. Although these limitations may act to alter overall patterns, a 10% prevalence rate is consistent with findings from Israel (Bleich et al, 2003) and estimates following explosions (25%; Hayes & McAllister, 2001) and being a victim of violence in Northern Ireland (14%).

Fig. 1 Mean number of direct and indirect experiences of the ‘troubles’ according to classification with the Post-Traumatic Stress Disorder Checklist. PTSD, post-traumatic stress disorder; ◼ direct experiences; □ indirect experiences.

Fig. 2 Percentage of total sample and subsample with post-traumatic stress disorder (PTSD) who consider themselves to be ‘victims’ of the ‘troubles’. Never; ◼ rarely; ◼ sometimes; ◼ often; □ very often (the percentage answering ‘Don’t know’ is too small to show).
Clinical implications

Many of those with symptoms suggestive of PTSD do not consider themselves victims of the ‘troubles’ and hence it is not surprising that some have resorted to self-medication instead of seeking professional help: our evidence shows a higher reported misuse of substances. Current government policy is targeting services towards ‘victims of the troubles’. Our findings suggest that advertising or targeting resources towards ‘victims’ might act as a barrier to those who have been most adversely affected. Finally, holistic approaches that consider previous traumatic experiences and socio-economic background are crucial to understanding the impact of any specific incident in conflict situations.

ACKNOWLEDGEMENTS

This research was supported by an award to OTM from the Cross-Border Consortium under the EU Peace II Programme and was financed in part by the UK and Irish governments.

REFERENCES


Prevalence of dementia in intellectual disability using different diagnostic criteria

A. STRYDOM, G. LIVINGSTON, M. KING and A. HASSIOTIS

Background Diagnosis of dementia is complex in adults with intellectual disability owing to their pre-existing deficits and different presentation.

Aims To describe the clinical features and prevalence of dementia and its subtypes, and to compare the concurrent validity of dementia criteria in older adults with intellectual disability.

Method The Becoming Older with Learning Disability (BOLD) memory study is a two-stage epidemiological survey of adults with intellectual disability without Down syndrome aged 60 years and older, with comprehensive assessment of people who screen positive. Dementia was diagnosed according to ICD–10, DSM–IV and DC–LD criteria.

Results The DSM–IV dementia criteria were more inclusive. Diagnosis using ICD–10 excluded people with even moderate dementia. Clinical subtypes of dementia can be recognised in adults with intellectual disability. Alzheimer’s dementia was the most common, with a prevalence of 8.6% (95% CI 5.2–13.0), almost three times greater than expected.

Conclusions Dementia is common in older adults with intellectual disability, but prevalence differs according to the diagnostic criteria used. This has implications for clinical practice.

Declaration of interest None.

METHOD

The Becoming Older with Learning Disability (BOLD) memory study is a two-stage epidemiological survey of dementia in the total population of adults with intellectual disability without Down syndrome aged 60 years and older living in five London Boroughs; this area had a total adult population aged 60 years and older of 177,544 people in the UK 2001 census (http://www.statistics.gov.uk/census2001/census2001.asp). The protocol received approval from the Thames Valley Multi-centre Research Ethics Committee and from the Research and Development offices of all participating National Health Service (NHS) organisations.

Participants We identified potential participants from social services’ electronic databases of past and present intellectual disability service users combined with lists of past or present users of local intellectual disability health teams. We also contacted all residential and day services for adults with intellectual disability to ensure that all known older adults with such disability had been identified. Participants included those resident in their own homes, family homes, residential homes of all types, nursing homes and hospitals. In two of the boroughs we also contacted all geriatricians, old age psychiatrists, mental health teams for older people, and all residential and nursing homes caring for people without intellectual disability: this resulted in the identification of only one additional participant with intellectual disability, and so this extension of the sampling frame was not implemented in the other three boroughs. Participants received accessible information written in simple language with pictures. A capacity assessment was undertaken to determine whether the person was able to provide consent; if this was the case, written informed consent was obtained. For those who did not have capacity to consent, assent was given by carers, provided the person did not show unwillingness to participate. Written informed consent was also gained from informants for their own participation. We sought historical information to cover at least the preceding 2 years for those who screened positive.

Intellectual disability was defined according to ICD–10 criteria for mental retardation (World Health Organization, 1992).
Dementia in Intellectual Disability

1993): that is, a reduced level of intellectual functioning (an IQ below 70) which first manifested during the developmental period and results in diminished ability to adapt to the daily demands of the normal social environment. Those in whom the diagnosis was uncertain underwent an assessment and were excluded if they did not meet the ICD–10 criteria. Adults with Down syndrome were identified from records of chromosomal analysis or by their characteristic features, and excluded from the study.

Screening stage

All participants who were able and all informants completed a screen for symptoms of dementia or cognitive decline. Informants completed the Dementia Questionnaire for Persons with Mental Retardation (DMR; Evenhuis, 1996), an established screening tool for dementia with good psychometric properties in this population (Strydom & Hassiotis, 2003). They also completed a brief activities of daily living schedule based on the Adaptive Behavior Scale (Nihira et al., 1992) and the Activities for Daily Living Schedule (Lawton & Brody, 1969). We recorded collected information about level of functioning in early life and decline in activities of daily living over the past 2 years from informants. Participants with intellectual disability who had sufficient communication ability completed a three-item object memory task based on the Shoe Box Test (Burr & Aylward, 2000; Silverman et al., 2004). Screening criteria were designed for maximum sensitivity so that no person with dementia would be missed. Screen positives fulfilled any of the following conditions: a score at or above the cognitive score thresholds for dementia provided by Evenhuis (1996) for severe, high-moderate or mild intellectual disability on the DMR; an unexplained decline in activities of daily living; or a delayed recall of fewer than two items in the memory task. Participants who screened negative on these criteria were presumed not to have dementia.

Assessment of people who screened positive

Participants who screened positive completed a full assessment to elicit symptoms of dementia as described below.

All screening tests and assessments were completed by a qualified intellectual disability psychiatrist (A.S.).

Neuropsychological assessment

Basic neuropsychological assessment consisted of the Test for Severe Impairment (Albert & Cohen, 1992), additional memory items from the Severe Impairment Battery (Saxton & Swihart, 1989), the Tower of London test (Shalllice, 1982), the Supermarket Fluency task (Troyer, 2000), the British Picture Vocabulary Scale (Dunn & Dunn, 1997) and the Luria three-stage command (Hodges, 1994). Informants also completed a questionnaire based on a modification of the Cambridge Mental Disorders Examination (CAMDEX) informant questionnaire to elicit a history of changes in memory, personality, general cognitive function and confusion (Ball et al., 2004).

Physical examination

A structured physical examination was conducted to record neurological symptoms and signs associated with dementia and to identify other physical disease such as thyroid disease, neurological conditions and cardiovascular disorders, based on memory clinic assessments (Hassiotis et al., 2003). This included a vision and hearing screen. Informants provided details of current health and medications, and medical records were reviewed to obtain information on previous health status and recent investigations. We recorded the results of neuroimaging undertaken in the preceding 2 years.

Mental state examination

Mental disorders and psychiatric symptoms were screened for with the mini Psychiatric Assessment Schedule for Adults with Developmental Disability (PAS–ADD), a tool for assessing adults with intellectual disability (Moss, 2002).

Diagnosis

All the above information was compiled in an anonymised summary, which was presented to two of three psychiatrists (A.H., G.L. or A.S.) for independent diagnostic review. Two were intellectual disability psychiatrists and one (G.L.) was an old age psychiatrist. Any disagreement in ratings was settled by discussion with the third psychiatrist. A specially developed tick list with operationalised criteria was used to produce a differential diagnosis. We applied the following diagnostic principles:

(a) The key to dementia diagnosis in this population is decline in cognitive function from an individual baseline, not change from a normal level (Aylward et al., 1997).

(b) We followed a hierarchical process, consistent with diagnostic systems such as DC–LD (Royal College of Psychiatrists, 2001), whereby developmental mental retardation syndrome, autistic disorders, physical illness and medication effects, sensory loss, environmental change or life events, or mental illness had to be considered sequentially as possible reasons for screening positive.

(c) General dementia criteria had to be met first before moving on to subtyping. However, since criteria for Lewy body dementia and frontotemporal dementia were designed as stand-alone criteria outside of the ICD–10 or DSM–IV criteria, these disorders were not subjected to the two-stage process.

(d) The list included the following criteria for dementia: ICD–10 Research Diagnostic Criteria (World Health Organization, 1993), DSM–IV–TR (American Psychiatric Association, 2000) and DC–LD criteria (Royal College of Psychiatrists, 2001), which are compared in Table 1; ICD–10 (World Health Organization, 1993), DSM–IV (American Psychiatric Association, 2000) and the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer’s Disease and Related Disorders Association (NINCDS–ADRDA); McKhann et al., 1984) criteria for Alzheimer’s disease; ICD–10, DSM–IV and National Institute of Neurological Disorders and Stroke–Association Internationale pour la Recherche et l’Enseignement en Neurosciences (NINDS–AIREN (Roman et al., 1993) criteria for vascular dementia; the Consortium on Dementia with Lewy Bodies (DLB) criteria (McKeith et al., 1996); and the Work Group on Frontotemporal Dementia and Pick’s Disease criteria for frontotemporal dementia (FTD; McKhann et al., 2001).

(e) Dementia is an organic disorder and should therefore trump mental illnesses such as depression in hierarchical systems; instead, it is often defined as a diagnosis of exclusion in the diagnostic systems. We made the diagnosis of dementia in the presence of depressive symptoms if these were deemed not to account for the cognitive decline, but the final judgement
Table I  Criteria for dementia in the classification systems

<table>
<thead>
<tr>
<th>Impaired domain/ symptoms</th>
<th>DSM–IV</th>
<th>ICD–10</th>
<th>DC–LD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Short- and/or long-term</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Higher cortical functions²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive function</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Thinking</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Judgement</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Other cognitive skills</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Information processing</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Aphasia/language skills</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apraxia</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Agnosia</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioural and emotional function¹</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Emotional lability</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Irritability</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Apathy</td>
<td>O</td>
<td>O</td>
<td></td>
</tr>
<tr>
<td>Social behaviour</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other criteria</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change from premorbid state/decline in level of functioning¹</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Duration of at least 6 months</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Exclusions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not caused by delirium</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Not caused by mental illness</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>

*, required for diagnosis.

1. The ICD–10 classification requires a decline in memory and other cognitive function, but does not have a separate criterion for change or deterioration in function.

2. At least one of the circled is required.

For the purpose of this analysis the participants were divided into two groups: those with dementia (if they met any of the above diagnostic criteria) and those who did not meet the criteria.

Statistical analysis

Data were entered into the Statistical Package for the Social Sciences version 11 for Windows. Prevalence rates are presented in percentages, rounded to one decimal place. Symmetrical exact binomial 95% confidence intervals were derived using a calculator available at http://statpages.org/confint.html. Chi-squared tests were used to analyse categorical variables with continuity correction for 2 × 2 tables; Fisher’s exact tests were used if 50% or more cells had expected values of less than 5. Significance level was set at P < 0.01 owing to the number of tests; t-tests were used to analyse differences in mean age. Correlation between sets of criteria was calculated with Spearman’s rho.

Prevalence rates for Alzheimer’s disease and vascular dementia in the general population were obtained from the most recent European collaborative study of population-based cohorts (4.4% for Alzheimer’s disease and 1.6% for vascular dementia; Lobo et al, 2000). These rates were used to calculate expected counts for this study. The observed count divided by the expected count provided standardised morbidity ratios (SMRs) for comparison of rates between populations (Page et al, 1995). Confidence intervals for SMRs were obtained with a calculator providing exact 95% Poisson confidence intervals (http://home.clara.net/sisa/smr.htm).

RESULTS

We identified 258 potential participants from health or social services. An additional 23 (8.2%) were identified through other providers. All 281 potential participants were contacted. Of these, 24 (8.5%) were ineligible for the study because of unrecorded Down syndrome status, being too young, having died recently, not having an intellectual disability, or not residing at the given address. Of the remaining 257 individuals, 35 (13.6%) refused participation, or their carers refused on their behalf; 222 (86.4%) participated. The prevalence of eligible participants in the total population of all adults aged 60 years and older living in these boroughs was 0.15%. Participants did not differ from non-participants in terms of mean age (68.8 vs. 67.9 years; t = -0.776, P = 0.439) or gender (Pearson χ² = 0.14, P = 0.708). The proportion of male to female participants was 52.7% to 47.3%. With regard to severity of disability, 123 (55.4%) participants were rated to have mild intellectual disability and 99 (44.6%) had moderate or more severe disability.

Participants who screened positive

Overall, 60 people screened positive; 29 of these met at least one set of dementia criteria (including DLB and FTD criteria). Of these, 13 (45%), or 5.9% of the total) already had the diagnosis of probable or possible dementia recorded in their clinical notes. ‘False’ positives (i.e. those who screened positive but did not meet dementia criteria) were younger (mean age 70.9 v. 76.4 years; t = -2.667, P = 0.01) and more likely to have severe intellectual disability (41.9 v. 3.4%; χ² = 10.349, P = 0.001), but the true positives and false positives did not differ significantly with
regard to gender, health problems, mental illness or sensory disabilities.

**Mental illness**

The prevalence rates of current mental illness (as reported by informants or extracted from medical records) are given in Table 2; this table also includes the numbers with scores above the mini PAS–ADD thresholds. The proportions of those with mental illness who were also diagnosed with dementia are given in the last column. Since depression is an important differential diagnosis of dementia and may be difficult to distinguish from dementia in older adults, we examined all the cases with a history or mini PAS–ADD threshold score of depression that also met the criteria for dementia. Six adults with a recent history of depression were deemed to have dementia. Only two of them had scores above the depression threshold of the mini PAS–ADD; the rest had fully recovered or had remission of most symptoms, and their cognitive declines were deemed not to relate to the depressive episode. Three of them met all three sets of dementia criteria; one met only the DSM–IV criteria because she did not have a history of behavioural or social decline. She was diagnosed with dementia due to Parkinson’s disease. Of the adults who reached the mini PAS–ADD threshold for depression, one was a 69-year-old woman with mild intellectual disability and a long history of cognitive decline, considerable loss of function and emergence of other neuropsychiatric symptoms. She was diagnosed by her local intellectual disability psychiatrist as having Alzheimer’s disease 2 years prior to participating in the study, and was treated with donepezil for 6 months. She was rated to have depression symptoms secondary to dementia and met the dementia criteria of ICD–10, DC–LD and DSM–IV. The other person was a 75-year-old man with mild intellectual disability and a history of psychotic illness with depressive episodes since early adulthood. He had a 2-year history of gradual decline in cognitive function and activities of daily living, personality and behavioural changes, episodes of confusion and falls. He had memory deficits on psychometric testing and met the ICD–10 criteria for dementia, but not those of DSM–IV or DC–LD because the raters were not unable to exclude the possibility that his symptoms were related to his mental illness.

**Dementia symptoms**

There were 26 participants with dementia for whom the informants could identify the initial symptoms. The most common initial symptom was general deterioration in functioning (n=13; 50% of those with dementia), followed by behavioural or emotional change (n=4; 15.4%). Deterioration in memory (n=2; 7.7%) or other cognitive functions (n=2; 7.7%) was rarely noticed to be prominent in the early stages of the disorder. Other early symptoms (n=5) included episodes of confusion (n=3).

We compared the current dementia symptoms reported by informants in those who screened positive by diagnostic group (any dementia compared with no dementia) (Table 3). The most common symptoms for those with dementia were decline in self-care (90% of those with dementia), decline in instrumental activities of daily living (72%), memory decline (73%), episodes of confusion (52%) and the development of muddled thinking (62%). Symptoms that significantly discriminated between those with and without dementia in those who screened positive were deterioration in self-care ability, deterioration in instrumental activities of daily living, change in memory, development of muddled thinking, development of problems with thinking ahead and planning, and newly developed perseveration. None of the behavioural and emotional symptoms were discriminative of dementia.

**Table 2 Mental illness and dementia diagnoses**

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Prevalence (%)</th>
<th>With dementia n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recent history of mental illness (n=222)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>24</td>
<td>10.8</td>
<td>6 (25)</td>
</tr>
<tr>
<td>Bipolar disorder</td>
<td>10</td>
<td>4.5</td>
<td>2 (20)</td>
</tr>
<tr>
<td>Anxiety disorder</td>
<td>16</td>
<td>7.2</td>
<td>4 (25)</td>
</tr>
<tr>
<td>Schizophrenia</td>
<td>2</td>
<td>0.9</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Psychosis NOS</td>
<td>15</td>
<td>6.8</td>
<td>7 (47)</td>
</tr>
<tr>
<td>Any mental disorder (including behavioural problems)</td>
<td>93</td>
<td>41.9</td>
<td>18 (19)</td>
</tr>
<tr>
<td>Mini PAS–ADD cases (n=60)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>2</td>
<td>2 (100)</td>
<td></td>
</tr>
<tr>
<td>Mania</td>
<td>2</td>
<td>1 (50)</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>9</td>
<td>6 (67)</td>
<td></td>
</tr>
<tr>
<td>Psychosis</td>
<td>7</td>
<td>5 (71)</td>
<td></td>
</tr>
<tr>
<td>Unspecified disorder</td>
<td>7</td>
<td>7 (100)</td>
<td></td>
</tr>
</tbody>
</table>

PAS–ADD, Psychiatric Assessment Schedule for Adults with Developmental Disabilities; NOS, not otherwise specified.

1. Percentage of group with that disorder.
2. Completed for those who screened positive only; prevalence rates therefore not available.

**Overall dementia and subtype prevalence rates**

Prevalence rates for dementia and subtype criteria are given in Table 4. Criteria for Alzheimer’s disease (ICD–10, DSM–IV or NINCDS–ADRDA) were met in 66% of those with dementia. The second most common subtype was Lewy body dementia (possible and probable cases) followed by frontotemporal dementia and then vascular dementia. Frontotemporal dementia was the most common subtype after Alzheimer’s disease if possible cases of Lewy body dementia are discounted. Alzheimer’s and vascular dementias diagnosed by DSM–IV criteria were almost twice as common as the corresponding ICD–10 rates (Table 4). The prevalence rates for those aged 65 years or over who met any criteria for Alzheimer’s or vascular dementia were used to make comparisons with the general population rates. The 17 observed cases of Alzheimer’s disease among those aged 65 years or over compared with 6.25 expected cases resulted in a standardised morbidity ratio (SMR) of 2.72 (95% CI 1.58–4.35). The corresponding observed v. expected count for vascular dementia was 5 v. 2.27 (SMR=2.20, 95% CI 0.72–5.14).

**Dementia criteria**

Twenty-eight people met any of the ICD–10, DSM–IV or DC–LD criteria for dementia; 27 of these (12.2% of the total sample)
met the criteria for DSM–IV dementia, 22 (9.9%) met the criteria for ICD–10 dementia and 23 (10.4%) the criteria for DC–LD dementia. The overlap between these criteria is shown in Fig. 1: this demonstrates that 21 participants (75%) met all three sets of criteria, those meeting DC–LD criteria were a subset of those meeting DSM–IV criteria, and there were 5 participants who met one set of diagnostic criteria only (ICD–10 or DSM–IV). The criteria are therefore correlated as follows: DSM–IV × ICD–10 $r = 0.772 \ (P \leq 0.005)$; DSM–IV × DC–LD $r = 0.872 \ (P \leq 0.005)$; DC–LD × ICD–10 $r = 0.894 \ (P \leq 0.005)$.

The raters made clinical ratings of severity of dementia for all 29 meeting at least one set of criteria: 12 (41%) were rated as having mild dementia, 16 (55%) as having moderate dementia and 1 (3%) as having severe dementia. Those diagnosed according to ICD–10 and DSM–IV dementia criteria were compared according to severity of dementia. Ten (83%) of the 12 rated as having mild dementia met DSM–IV criteria compared with 8 (66.7%) who met ICD–10 criteria. Six people met criteria for DSM–IV dementia but not ICD–10, and one met the criteria for ICD–10 but not DSM–IV. Of the six diagnosed by DSM–IV but not by ICD–10, half were rated clinically to have dementia of moderate severity. These were

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Dementia symptoms reported by informants (screen-positive cases; n = 60)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Memory</td>
</tr>
<tr>
<td></td>
<td>Change in memory**</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Executive function</td>
</tr>
<tr>
<td></td>
<td>Decision difficulty</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Thinking ahead/planning problems**</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Other cognitive functions</td>
</tr>
<tr>
<td></td>
<td>Keep mind on things/concentration</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Muddled thinking**</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Talking more or less</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Word-finding difficulty</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Perseveration**</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Behavioral and emotional functions</td>
</tr>
<tr>
<td></td>
<td>More impulsive</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Character change</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Other symptoms</td>
</tr>
<tr>
<td></td>
<td>Hallucinations</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Episodes of confusion</td>
</tr>
<tr>
<td></td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Yes</td>
</tr>
</tbody>
</table>

ADL: activities of daily living.

**P < 0.01; χ² tests with continuity correction and 1 degree of freedom.
Participants with dementia were reported by their carers to have had initial deterioration in functional ability rather than changes in memory and other cognitive functions. Non-cognitive symptoms such as personality changes were also common early symptoms, but did not differentiate between those with and without dementia who screened positive. Deterioration in self-care ability and instrumental activities of daily living were both discriminative of dementia in people who screened positive.

Because dementia may present differently in this population compared with the general population, criteria for the disorder may also perform differently. This is the first study to report the prevalence of dementia criteria in older adults with intellectual disability. We have demonstrated that correlations between the ICD–10, DSM–IV and DC–LD dementia criteria were good, but there were important differences. The DSM–IV criteria diagnosed a larger number of participants with mild dementia than ICD–10 criteria and were therefore more inclusive. The ICD–10 criteria excluded not only those with mild dementia, but also a considerable proportion of those with moderate-to-severe dementia.

**DISCUSSION**

This is the first study to report the prevalence of subtypes of dementia, including frontotemporal and Lewy body dementia, in older adults with intellectual disability. We have demonstrated that the symptoms associated with all dementia subtypes can be recognised in older adults with such disability. As in their general population counterparts, Alzheimer’s disease was the most common diagnosis, but with a prevalence of almost three times higher than expected. Lewy body and frontotemporal dementias were common, as in the general population (Stevens *et al.*, 2002). However, these dementias were more common than vascular dementia, which is unexpected since vascular dementia is usually the second most common type in the general population (Fratiglioni *et al.*, 2000). This may be due to the criteria for frontotemporal dementia we have used, which are broad and expected to be more sensitive than other criteria (Neary *et al.*, 2005), but may also be due to the low prevalence of some (but not all) vascular risk factors such as smoking in the population with intellectual disability (Janicki *et al.*, 2002).

**Limitations**

This study is the largest cross-sectional survey of dementia in the intellectual disability population to date; our sample represents approximately 1% of the estimated 26,000 adults aged 60 years and over known to have intellectual disability in England (Emerson & Hatton, 2004). We employed epidemiological sampling methods and achieved high participation rates. We identified all older adults known to have intellectual disability. Participants underwent a very sensitive screening strategy, and were fully assessed with established assessment methods and tools if screened positive, before we applied a rigorous diagnostic procedure, which incorporated the main diagnostic criteria for dementia.

Despite the comprehensive recruitment strategy, it is possible that we have missed some older adults with intellectual disability who were unknown to social or health services. However, we believe this number

---

**Table 4** Prevalence rates for dementia subtypes

<table>
<thead>
<tr>
<th>Dementia subtype</th>
<th>Age ≥ 60 years</th>
<th>Age ≥ 65 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=222)</td>
<td>(n=142)</td>
</tr>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Alzheimer’s dementia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any criteria</td>
<td>19</td>
<td>8.6</td>
</tr>
<tr>
<td>Specific criteria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICD–10</td>
<td>8</td>
<td>3.6</td>
</tr>
<tr>
<td>DSM–IV</td>
<td>14</td>
<td>6.3</td>
</tr>
<tr>
<td>NINCDS–ADRDA</td>
<td>12</td>
<td>5.5</td>
</tr>
<tr>
<td>Vascular dementia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any criteria</td>
<td>6</td>
<td>2.7</td>
</tr>
<tr>
<td>Specific criteria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICD–10</td>
<td>3</td>
<td>1.4</td>
</tr>
<tr>
<td>DSM–IV</td>
<td>6</td>
<td>2.7</td>
</tr>
<tr>
<td>NINDS–AIREN</td>
<td>6</td>
<td>2.7</td>
</tr>
<tr>
<td>Dementia of Lewy body type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any criteria</td>
<td>13</td>
<td>5.9</td>
</tr>
<tr>
<td>DLB Consortium criteria</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possible</td>
<td>9</td>
<td>4.1</td>
</tr>
<tr>
<td>Probable</td>
<td>4</td>
<td>1.8</td>
</tr>
<tr>
<td>Frontotemporal dementia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FTD Work Group criteria</td>
<td>7</td>
<td>3.2</td>
</tr>
<tr>
<td>Other dementias (e.g. head trauma and Parkinson’s disease)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any criteria</td>
<td>3</td>
<td>1.4</td>
</tr>
<tr>
<td>Any dementia</td>
<td>29</td>
<td>13.1</td>
</tr>
</tbody>
</table>

DLB, dementia with Lewy bodies; FTD, frontotemporal dementia.

---

excluded from ICD–10 criteria either because informant history of memory decline was absent (as opposed to other evidence of such decline, which is acceptable for DSM–IV diagnosis) or by the absence of behavioural and emotional symptoms. The extra ICD–10 case was rated to have mild dementia. The reason this did not meet DSM–IV criteria was that depressive symptoms were present and therefore one of the DSM–IV exclusion criteria was met.

---

Fig. 1 Venn diagram of participants diagnosed with dementia on different diagnostic criteria.
to be small because older adults with such disability are likely to need assistance owing to the functional problems associated with ageing. This is more likely to be provided by agencies outside the family because informal support networks decrease as people grow older. Furthermore, the comprehensive care system in the UK promotes formal assistance. A small number of potential participants unknown to any service might reduce the increased prevalence of Alzheimer’s disease when compared with the general population, but is unlikely to change our main findings about the relative prevalence of subtypes, presentation of dementia or performance of diagnostic criteria. We excluded adults with Down syndrome recognised by their clinical features, but did not undertake chromosomal analysis; it is therefore possible that some of these excluded adults did not have trisomy 21.

Another limitation is that cross-sectional assessments are less reliable than sequential assessments. We therefore supplemented our data with historical information from informants and medical records. Nevertheless, for a proportion of participants we were unable to decide whether or not they had dementia owing to insufficient data; these were included in the group without dementia. Our study might therefore have underestimated the true prevalence of subtypes of dementia.

**Dementia symptoms and concurrent validity of dementia criteria**

Because of diagnostic difficulties in this population, clinical diagnosis cannot be used as the gold standard for comparison. We therefore determined the correlation of different dementia criteria and demonstrated their utility, but also highlighted particular issues. Cognitive deficits are difficult to demonstrate in adults with limited verbal and functional ability (Burt & Aylward, 1999); clinicians therefore often rely on informant reports of change. Our data confirm that change in memory and higher functions are not noticed early in people with intellectual disability, and because these changes are required for dementia diagnosis, adults with both intellectual disability and dementia may be diagnosed later in the course of the disorder when these changes have become more apparent.

Dementia criteria differ considerably and therefore yield widely differing prevalence rates in the general population (Ballard & Bannister, 2005). The ICD–10 criteria are more specific but less sensitive than DSM–III–R or DSM–IV criteria (Erkinjuntti et al, 1997). We have shown that this is also the case in older adults with intellectual disability. One of the reasons for this is that ICD–10 criteria are more demanding to apply because they are more dependent on reliable information from informants (Henderson et al, 1994). Another limitation of the ICD–10 and DC–LD criteria is that behavioural and emotional changes are an additional required symptom for ICD–10 and DC–LD dementia, but not for DSM–IV dementia. These were reported to have occurred early in a small but significant number of adults with intellectual disability and dementia. However, these symptoms were not good at discriminating between those with and without dementia, and limit the number of people diagnosed with ICD–10 criteria. Even those clinically rated to have moderate severity of dementia did not meet ICD–10 criteria. This was contrary to the expectation of an international consensus group (Aylward et al, 1997).

The ‘false’ screen positives need special mention. Those with severe intellectual disability were more likely to meet screening criteria but not diagnostic criteria for dementia. The proportion of false screen positives may seem high, but a recent study in an elderly population noted that of 96 people with confirmed cognitive and functional impairment, only 55 satisfied the DSM–IV criteria for dementia (Shaji et al, 2005). These authors felt that the DSM–IV prevalence of dementia is possibly an underestimation; this might also be the case in the population with intellectual disability, because the ‘false’ screen positive group might contain cases of dementia that did not meet criteria owing to lack of informant or medical history, or to the difficulty of making this diagnosis in a group with severe disability.

**Clinical implications**

We found that more than double the number of older adults with intellectual disability meet dementia criteria than is recognised by their carers or health professionals. Functional decline was reported to be more common than memory decline early on in the presentation; perhaps the potential for pathological causes underlying such decline is not recognised in adults with lifelong deficits. Dementia should always be considered as a possible diagnosis when investigating reports of decline in older adults with intellectual disability. Our findings also give credence to screening approaches that rely on functional change (Prasher et al, 2004).

We preferred the DSM–IV criteria for dementia in this population. They are clearly set out and easy to interpret. They do not rely exclusively on informant report of memory and cognitive change like the ICD–10 criteria, which allows the clinician to use other sources of information such as sequential cognitive assessments and medical records. Furthermore, they do not require behavioural or emotional change but focus on functional change, which is important in this population. This has important implications for patients, since the use of DSM–IV criteria may enable earlier diagnosis of dementia in larger numbers of older adults with intellectual disability, which could gain them timely access to appropriate interventions.

**Future research**

Our findings raise questions about the aetiology of dementia in older adults with intellectual disability but without Down syndrome. It is important to establish why Alzheimer’s dementia may be more common in these adults than in the general population; we have estimated an SMR of 2.72 (95% CI 1.58–4.35). Possibilities include genetic causes such as apolipoprotein E4 alleles, or environmental causes such as brain damage during birth and early life, which is associated with intellectual disability but may also in the long term be associated with Alzheimer’s disease.

The incidence and presentation of dementia and validity of diagnoses should be confirmed longitudinally. It is also important to confirm subtype diagnoses with post-mortem studies, and to investigate the aetiology of dementia in this population. This will enable appropriate interventions and illuminate our understanding of dementia presentation and progression throughout the intellectual range.

**ACKNOWLEDGEMENTS**

This study was funded by the Medical Research Council (UK) with a Training Fellowship Grant to AS. Additional support was provided by the Penrose Society in the form of a Jancar Traveling Fellowship.


Associations between common mental disorders and the Mental Illness Needs Index in community settings

Multilevel analysis*

DAVID L. FONE, FRANK DUNSTAN, ANN JOHN and KEITH LLOYD

Background The relationship between the Mental Illness Needs Index (MINI) and the common mental disorders is not known.

Aims To investigate associations between the small-area MINI score and common mental disorder at individual level.

Method Mental health status was measured using the Mental Health Inventory of the Short Form 36 instrument (SF–36). Data from the Caerphilly Health and Social Needs population survey were analysed in multilevel models of 10,653 individuals aged 18–74 years nested within the 2001 UK census geographies of 110 lower super output areas and 33 wards.

Results The MINI score was significantly associated with common mental disorder after adjusting for individual risk factors. This association was stronger at the smaller spatial scale of the lower super output area and for individuals who were permanently sick or disabled.

Conclusions The MINI is potentially useful for small-area needs assessment and service planning for common mental disorder in community settings.

Declaration of interest None.

Despite the high prevalence and public health importance of the common mental disorders (Weich, 1997) there is no small-area index specific to mental health for health needs assessment and planning the appropriate provision of services in primary care settings. For severe mental illness, the Mental Illness Needs Index (MINI; Glover et al., 1998, 2004) can be used to estimate need for specialist psychiatric services in areas definable by electoral wards. In this study we investigated the small-area ecological relationship between MINI scores and common mental disorders, whether MINI was associated with these disorders after accounting for individual risk factors, and whether any observed associations varied in magnitude with geographical scale and population subgroup, characterised by age, gender, social class and employment status.

METHOD

Data source

We analysed data from the Caerphilly Health and Social Needs Survey, a community study of health inequality set in Caerphilly county borough, Wales, UK, described in detail elsewhere (Fone et al., 2006). The borough is one of the 22 local government areas in Wales, situated between the cities of Cardiff and Newport in the south and the Brecon Beacons National Park to the north. Briefly, we carried out a cross-sectional postal questionnaire survey of the resident adult population aged 18 years and over in autumn 2001 and obtained a representative data-set on 12,408 residents of the borough (adjusted response 63%). The random sample was stratified by census ward and drawn from the general practitioner age-gender register held by the health authority. The survey included questions on a wide range of demographic and socio-economic factors, and the 36-item Short Form Health Survey (SF–36) version 2 health status questionnaire (Ware et al., 2000a).

Mental health outcome measure

The five-item Mental Health Inventory (MHI–5) sub-scale of the SF–36 version 2 health status questionnaire was used as the measure of common mental disorder (Ware et al., 2000a). The validity and reliability of the MHI–5 is well established and reflects the continuously distributed nature of mental health status in the population (Ware & Gandek, 1998; Ware et al., 2000b). Studies have shown that the MHI–5 is at least as good a measure of common mental disorder as the commonly used 12-item General Health Questionnaire (Weinstein et al., 1989; Berwick et al., 1991; McCabe et al., 1996).

The MHI–5 used in the SF–36 version 2 comprises five questions relating to the past 4 weeks: ‘Have you been very nervous?’ ‘Have you felt down in the dumps that nothing could cheer you up?’ ‘Have you felt calm and peaceful?’ ‘Have you felt downhearted and depressed?’ ‘Have you been happy?’ Each of the five questions has five response categories which are scored from 1 to 5: ‘all of the time’ 1; ‘most of the time’ 2; ‘some of the time’ 3; ‘a little of the time’ 4; or ‘none of the time’ 5. Thus each respondent could achieve a total score within the range 5–25. For the third and fifth questions the scoring was reversed so that lower scores indicated worse mental health status on a continuous scale. We transformed the response scores and imputed missing data to a scale of range 0 to 100 using the standard method (Ware et al., 2000b).

Survey population for analysis

We restricted the analysis to respondents aged less than 75 years because the SF–36 is less reliable in UK elderly populations (Hayes et al., 1995; Hill et al., 1996) and the proportion of missing mental health and socio-demographic response data in the data-set increased markedly for those over the age of 75 years. The mental health score was available for 10,653 (97.8%) of the 10,892 respondents aged 18–74 years.

Individual level variables

We selected variables that were significantly associated with the mental health score in univariable analyses. Age was modelled as a continuous variable, centred around the mean to avoid estimation...
problems. Gender, occupational social class, employment status and housing tenure were modelled as categorical variables (Table 1).

**Calculation of the area-level MINI scores**

Two spatial levels are defined by the 2001 census in England and Wales within the study data-set: 110 lower super output areas which are nested within 33 electoral wards. Lower super output areas are built up from around five output areas, the smallest geographical scale used in the census. They are constrained to a minimum population size of 1000 and in Caerphilly borough the mean population was 1541 (range 1010–2141). For wards, the mean population was 5137 (range 1803–11 530).

First, we calculated the MINI score for both of these geographical areas from the six census variables (Table 2) using the original method at ward level described by Glover et al. (1998). Calculation of MINI scores was less straightforward for lower super output areas because detailed tables of census data at this geographical level have not been released by the Office for National Statistics (ONS). However, the Census Key Statistics univariate tables data-set, which is available online from Neighbourhood Statistics (http://www.neighbourhood.statistics.gov.uk), contains selected variables at lower super output area which closely match the definitions required to calculate the MINI score. There are two small differences in these data compared with the data available at ward level which are not likely to make any material difference to the final MINI score: the economically active age range is 16–74 years (instead of the usual age range for the economically active of 16–59 years for women and 16–64 years for men) and the car ownership variable is defined by household instead of by individual. The MINI score is standardised to a mean of 100 in the area of study with a standard deviation of 10.

Second, we wished to follow Glover et al. (2004) and calculate the updated lower super output area MINI score, which is based on the modelled relationship between admission rates for severe mental illness and new population data used in the construction of the Index of Multiple Deprivation in England (Glover et al., 2004). However, the updated MINI is specific to England and the differences in the method of construction between the Welsh and English Indices of Multiple Deprivation mean that an updated MINI for Wales that is comparable to England cannot be calculated. We therefore used the Welsh Index of Multiple Deprivation 2005 (WIMD2005; Local Government Data Unit – Wales, 2005) as the closest approximation for the updated MINI in the analysis.

**Statistical analysis**

We assessed the ecological correlation between the MINI scores and the mean area MHI–5 scores with Spearman’s rank correlation coefficient. We then analysed the data-set in two separate multilevel models. The first model included the 10 653 individuals at level 1 nested within 110 lower super output areas. The second model included the 10 653 individuals at level 1 nested within the 33 census wards at level 2. We used a separate model for each of the geographical levels to avoid the collinearity that would have resulted

---

**Table 1 Univariable associations between mental health and individual risk factors**

<table>
<thead>
<tr>
<th>Gender</th>
<th>Respondents n (%)</th>
<th>Mental health score(1) Mean (s.d.)</th>
<th>Disorder present n (%)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>4770 (44.8)</td>
<td>71.85 (20.79)</td>
<td>1369 (28.7)</td>
<td>27.4–30.0</td>
</tr>
<tr>
<td>Female</td>
<td>5883 (55.2)</td>
<td>67.44 (22.20)</td>
<td>2126 (36.1)</td>
<td>34.9–37.4</td>
</tr>
<tr>
<td>Social class</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I, II</td>
<td>2407 (22.6)</td>
<td>74.35 (19.38)</td>
<td>535 (22.2)</td>
<td>20.6–23.9</td>
</tr>
<tr>
<td>III MN</td>
<td>2103 (19.7)</td>
<td>70.75 (20.84)</td>
<td>634 (30.1)</td>
<td>28.2–32.1</td>
</tr>
<tr>
<td>III M</td>
<td>2171 (20.4)</td>
<td>71.13 (21.31)</td>
<td>645 (29.7)</td>
<td>27.8–31.7</td>
</tr>
<tr>
<td>IV, V</td>
<td>2647 (24.8)</td>
<td>66.45 (22.01)</td>
<td>1039 (39.3)</td>
<td>37.4–41.1</td>
</tr>
<tr>
<td>Other</td>
<td>635 (6.0)</td>
<td>57.38 (25.33)</td>
<td>350 (55.1)</td>
<td>51.2–58.9</td>
</tr>
<tr>
<td>Missing data</td>
<td>690 (6.5)</td>
<td>65.20 (21.72)</td>
<td>292 (42.3)</td>
<td>38.7–46.0</td>
</tr>
</tbody>
</table>

1. Measured using the Mental Health Inventory sub-scale of the 36-item Short Form Health Survey questionnaire.

---

**Table 2 Census variables included in the Mental Illness Needs Index and their weighting within the index, reproduced from Glover et al. (1998) with permission**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proportion of adult population single, widowed or divorced (age 16+ years)</td>
<td>1.8</td>
</tr>
<tr>
<td>Proportion of population resident in households with no car</td>
<td>3.8</td>
</tr>
<tr>
<td>Proportion of population aged 16+ years registered permanently sick</td>
<td>2.5</td>
</tr>
<tr>
<td>Proportion of economically active adults unemployed</td>
<td>0.1</td>
</tr>
<tr>
<td>Proportion of population living in households not self-contained</td>
<td>1.4</td>
</tr>
<tr>
<td>Proportion of population resident in hostels, common lodging houses, miscellaneous establishments or sleeping rough</td>
<td>0.4</td>
</tr>
</tbody>
</table>
from modelling MINI at the two spatial levels simultaneously in one combined model.

The MHI–5 was modelled as a continuously distributed dependent variable in a normal response multilevel model. The modelling strategy was designed to meet the objectives of the study and started with “null” two-level variance components models of random intercepts. Here the variation in the mental health score was modelled by random intercept terms for either lower super output areas or wards, and a random error term for individuals. In model 1, the lower super output area MINI and WIMD2005 and the ward MINI were entered into the respective null models as continuous variables to obtain the unadjusted estimates. We modelled the MINI variables as z-scores so that the parameter estimate represents the change in predicted mental health score for a change in magnitude of the MINI of 1 s.d. Modelling z-scores meant that the MINI estimates could be compared directly between the two geographical levels used in the study. Individual-level variables were then entered in model 2. The categorical variables were modelled so that the reference categories were male, social class I or II, employed and owner-occupier. We modelled missing data for each categorical variable as a dummy term to avoid data loss and to permit direct comparison of each model using the deviance statistic. In this adjusted model the residual lower super output area and ward level random variances were assessed after including the individual-level variables. Finally, in model 3, we assessed whether any association between the mental health score and MINI varied with the age, gender, social class and employment status of individuals by modelling cross-level interactions between MINI and these individual-level variables.

The models were fitted in MLwiN software version 2.02 (Rasbash et al, 2001) and the parameters were estimated using Markov Chain Monte Carlo (MCMC) methods with iterative generalised least squares estimates (IGLS) as the starting values (Goldstein, 2003). Credible estimates for the 2.5th–97.5th quantiles for the random parameters were obtained using MCMC. Preliminary model fitting using IGLS was assessed by change in the deviance statistic. The validity of the final models was assessed using standard diagnostic plots of residuals at each level in the model.

RESULTS

The mean mental health score for all respondents was 69.4 (s.d.=21.7, median 75.0, interquartile range 55.0–85.0). Survey responses were obtained from individuals living in all 110 lower super output areas (mean 97 responses) and from all 33 wards (mean 323 responses). The mean mental health score for lower super output areas varied between 53.8 and 80.7 (mean 69.7, s.d.=4.3) and for wards between 61.4 and 76.1 (mean 69.3, s.d.= 3.1). The range of MINI scores was 71.2–124.0 for lower super output areas and 80.2–120.1 for wards. Figure 1 shows the spatial variation in MINI and mental health scores for lower super output areas, showing poorer mental health and higher MINI scores in the north of the borough. The MINI and mean mental health score were significantly correlated for both lower super output area (MINI: \( r = 0.69, P<0.001 \); WIMD2005: \( r = 0.73, P<0.001 \)) and ward levels \( r = 0.69, P<0.001 \).

Null models

The random variance estimates are shown in Table 3 for each area-level model. The variances at level 2 represent the variation in mental health scores between areas. The majority of the variance occurred at the individual level, with 1.78% (0.96–3.10) at ward level and 2.75% (1.87–3.89) at lower super output area level.

Associations between mental health and the MINI

In model 1, entering the MINI variable to the null models substantially reduced the random variances at both area levels (Table 3). The reduction in the lower super output area variance was greater for WIMD2005 which therefore explained a greater part of the variation in mental health scores. The MINI was significantly associated with individual mental health at both geographical levels in their respective models (Table 4). These associations were greater at the smaller spatial scale of the lower super output area, with some evidence of a stronger effect for WIMD2005 than MINI. We also modelled a quadratic and cubic function of the MINI score to assess the possibility of a non-linear effect, but these terms were not statistically significant. In model 2, addition of the individual-level variables further reduced the random variance at area level, showing the extent to which variation be-

---

**Fig. 1** Spatial variation in mean mental health. Mental Illness Needs Index (MINI) and Welsh Index of Multiple Deprivation 2005 (WIMD2005) scores at lower super output areas in Caerphilly borough.
tween individuals explained the variation in mean mental health scores between areas (see Table 3). The difference between the variance estimates for WIMD2005 and MINI was now smaller. Low mental health scores remained significantly associated with higher levels of each MINI score after adjusting for individual-level variables, with a marginally stronger effect for WIMD2005 compared with MINI (see Table 4).

The magnitude of the association between MINI and mental health can be interpreted by comparison with the association with social class. The raw parameter estimate of the lower super output area MINI score was −0.127. The MINI score ranges from 71.2 to 124.0 (a range of 52.8). Multiplying the raw estimate by the range gives a value of −6.1. The equivalent estimate for WIMD2005 was −6.2 and so both were nearly twice as large as the social class IV/V parameter estimate of −3.4.

Cross-level interactions between MINI and population subgroups

In model 3, both MINI and the cross-level interaction between MINI and the incapacity (defined as permanent sickness or disability) category of economic inactivity were statistically significant at both area levels (see Table 4). Thus, within lower super output areas and wards the gradient of association between mental health and MINI was more steeply negative for people economically inactive from permanent sickness or disability (incapacity) compared with the other categories of employment status (Fig. 2). Other cross-level interactions for gender, social class and tenure categories modelled were non-significant.

Model checking

Owing to the negative skew of the MHI–5 scores, the individual-level residuals were, as expected, negatively skewed. The model residuals at area level were normally distributed. No spatial pattern in these residuals was found and their correlations with the MINI score were not significantly different from zero.

DISCUSSION

We have shown that the small-area MINI score is significantly associated with common mental disorders. First, we found a strong ecological association between the MINI score and the small-area mean MHI–5 score. Second, in a multilevel analysis we found that MINI score was significantly associated with common mental disorders, after controlling for individual risk factors. This association was stronger at the smaller spatial scale of the lower super output area than the larger and more heterogeneous census ward. The association of common mental disorders with MINI score was significantly stronger in people who were economically inactive from permanent sickness or disability (incapacity). This group has the highest prevalence of common mental disorders in Wales (Fone & Dunstan, 2006).

We found little practical difference between MINI and the WIMD2005, used as the nearest proxy to the updated MINI in Wales. The strengths of association were not substantially different, but the WIMD2005 explained a little more of the random variation in mental health status. This suggests that WIMD2005 may be a better predictor of the area mean mental health score. This will be tested in further research.

Strengths and limitations of the study

The Caerphilly Health and Social Needs Survey has the strength of a large sampling fraction, resulting in a response data-set including around one in ten of a socially diverse population living in a geographically defined area, with detailed exposure data linked to the smallest census area level using the postcode of respondents. With a mean of 323 respondents per ward and 97 respondents per lower super output area, it was likely that the data-set was sufficiently large to meet the suggested ‘rules’ on sample sizes for multilevel analyses (Subramanian et al., 2003). Thus we were able to carry out robust analyses at smaller spatial scales than reported in the general multilevel literature (Pickett & Pearl, 2001), and were able to assess the MINI at the small geographical level of the lower super output area. This has the added advantage of being the spatial level at which the WIMD2005 is calculated for use in small-area planning and resource allocation in Wales.

The limitations of the study relate to the potential for bias. We were not able to validate survey responses to the MHI–5 scale with a clinical interview owing to the size of the study. One statistical property of the MHI–5 is that the
distribution of responses is significantly negatively skewed and thus might have violated the assumptions for linear regression. However, in previous research we have found very similar results from modelling the scale using the square transformation or as a binary variable of ‘case’ and ‘non-case’ of common mental disorder. This suggests that the normal response models were robust to departures from normality (Fone & Dunstan, 2006; Fone et al., 2007).

The original MINI (Glover et al., 1998) has the advantage of being calculated from UK census data, allowing comparability throughout the UK. One of the difficulties in using the original MINI in current research and for service planning is that it is less straightforward to calculate using 2001 census data than using the original 1991 data. Calculating the MINI for lower super output areas is hampered as ONS does not release the detailed tables of data required to calculate the MINI using the exact methodology described by Glover et al. (1998). Thus, a small compromise in variable definitions is necessary for using census data for lower super output areas. One advantage of using the MINI calculated from 2001 census data in this study is that the scores were almost exactly contemporaneous with the survey data, thus avoiding bias from temporal mismatch (Buzzelli & Su, 2006).

The updated MINI (Glover et al., 2004) has the advantage of being updatable on a more regular basis as it is derived from the English Index of Multiple Deprivation. However, it is not possible to derive these MINI scores for Wales, Scotland and Northern Ireland because the different versions of the Index of Multiple Deprivation used in the four countries of the UK limit comparability within the UK.

Mental Illness Needs Index scores derived from census data can also be calculated for general practice populations using a weighted proportional allocation methodology (Majeed et al., 1995), based on the distribution of practice populations within wards or lower super output areas. This is considerably less straightforward for the Indices of Multiple Deprivation in view of the more complex methodology used in their construction.

### Association between MINI and Mental Health Scores

<table>
<thead>
<tr>
<th>Model</th>
<th>Description</th>
<th>Estimate (s.e.)</th>
<th>2.5th–97.5th credible estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Lower super output area</strong></td>
<td>MINI²</td>
<td>-2.91 (0.32)</td>
<td>-3.54 to -2.27</td>
</tr>
<tr>
<td></td>
<td>MINI × incapacity²</td>
<td>-3.20 (0.28)</td>
<td>-3.74 to -2.64</td>
</tr>
<tr>
<td></td>
<td>WIMD2005²</td>
<td>-1.27 (0.24)</td>
<td>-1.75 to -0.80</td>
</tr>
<tr>
<td></td>
<td>WIMD2005 × incapacity²</td>
<td>-1.40 (0.58)</td>
<td>-2.52 to -0.27</td>
</tr>
<tr>
<td><strong>Ward</strong></td>
<td>MINI²</td>
<td>-2.51 (0.38)</td>
<td>-3.29 to -1.76</td>
</tr>
<tr>
<td></td>
<td>MINI × incapacity²</td>
<td>-0.93 (0.31)</td>
<td>-1.53 to -0.31</td>
</tr>
<tr>
<td><strong>Model 3</strong></td>
<td>(model 2 + cross-level interaction)</td>
<td>Estimate (s.e.)</td>
<td>2.5th–97.5th credible estimates</td>
</tr>
<tr>
<td></td>
<td>MINI²</td>
<td>-0.97 (0.26)</td>
<td>-1.48 to -0.45</td>
</tr>
<tr>
<td></td>
<td>MINI × incapacity²</td>
<td>-1.95 (0.69)</td>
<td>-3.30 to -0.60</td>
</tr>
<tr>
<td></td>
<td>WIMD2005²</td>
<td>-1.08 (0.25)</td>
<td>-1.57 to -0.59</td>
</tr>
<tr>
<td></td>
<td>WIMD2005 × incapacity²</td>
<td>-1.40 (0.58)</td>
<td>-2.52 to -0.27</td>
</tr>
</tbody>
</table>

**Table 4** Associations between the Mental Illness Needs Index and Welsh Index of Multiple Deprivation scores and individual mental health in two-level multilevel linear regression models

---

**Fig. 2** Model predicted mental health scores on the five-item Mental Health Inventory (MHI–5) × the lower super output area Mental Illness Needs Index (MINI) score, categorised by individual employment status.

---

MINI, Mental Illness Needs Index; WIMD2005, Welsh Index of Multiple Deprivation 2005.

1. The individual-level variables were age, gender, social class, employment status and housing tenure.
2. Modelled as a z-score.
3. The cross-level interaction between the MINI or WIMD2005 score and the incapacity (permanently sick or disabled) category of employment status.
the MINI score was associated with both outcomes (Croudace et al., 2000). One study used MINI to stratify general practices in a cluster randomised controlled trial which evaluated the effect of guidelines for the diagnosis of minor psychiatric morbidity (Croudace et al., 2003). However, this study did not investigate the association between the MINI and the prevalence of common mental disorders. Our study suggests further evidence of the generalisability of MINI as a measure of such disorders.

**Usefulness of the MINI for needs assessment and service planning**

Common mental disorders are highly prevalent in the community and among primary care consulting populations (Weich, 1997; Craig & Boardman, 1997). Yet there is no rational way of allocating resources at local level to support appropriate interventions. In primary care settings, decisions about who should receive treatment for depression, anxiety and other psychological morbidity seem to be made on a patient-by-patient basis and are influenced by the severity of the particular patient’s symptoms (Hyde et al., 2005). Service planning and resource allocation at the population level by primary care organisations require some area-based indication of the likely level of need.

Our results suggest that MINI can be used as a proxy for the prevalence of common mental disorders at small geographical area, practice and primary care organisation level. The MINI can be used for needs assessment and service planning in community settings in the same way that it is used for severe mental illness in secondary care settings. Scores calculated for general practice populations will be useful to primary care organisations in understanding the distribution of need for community mental health services within their defined populations. The establishment of such epidemiological relationships is important in the pursuit of transparent equitable resource allocation and the reconfiguration of mental health services away from the acute sector, allowing those with common mental disorders to be effectively managed in primary care.

**ACKNOWLEDGEMENTS**

We thank the Caerphilly study research team, and participating staff of Caerphilly County Borough Council and Caerphilly Local Health Group (now Caerphilly Local Health Board), for their help in running this collaborative study. We are grateful to Margaret Webber of the National Public Health Service for Wales in help for calculating MINI scores and to Gyles Glover for a personal communication of helpful documents. The study was funded by the Wales Office of Research and Development.

**REFERENCES**


Treatment of depression in primary care
Socio-economic status, clinical need and receipt of treatment

SCOTT WEICH, IRWIN NAZARETH, LOUISE MORGAN and MICHAEL KING

Background Depression is prevalent, costly and often under treated.

Aims To test the hypothesis that people with low socio-economic status are least likely to receive and adhere to evidence-based treatments for depression, after controlling for clinical need.

Method Individuals with an ICD–10 depressive episode in the past 12 months (n=866) were recruited from 7271 attendees in 36 general practices in England and Wales. Depressive episodes were identified using the 12-month Composite International Diagnostic Interview. Treatment receipt and adherence were assessed by structured interview, and rated using evidence-based criteria.

Results We identified 332 individuals (38.3%) who received and adhered to evidence-based treatment. There were few socio-economic differences in treatment allocation. Although those without educational qualifications were least likely to receive psychological treatments (OR=0.55, 95% CI 0.34–0.89, P=0.02), this association was not statistically significant after adjusting for depression severity.

Conclusions We found no evidence of inverse care in the treatment of moderate and severe depression in primary care in England and Wales.

Declaration of interest None. Funding detailed in Acknowledgements.

Depression has a community prevalence of 10% (Singleton et al, 2001), and is associated with physical morbidity and social impairment (Spitzer et al, 1995; Cassano & Fava, 2002). By 2020, depression is expected to become the second highest cause of disease burden worldwide (Murray & Lopez, 1997). The annual cost of depression in England alone was estimated at £9 billion in 2000, of which 90% was attributable to an estimated 110 million lost working days (Thomas & Morris, 2003). Unmet need for treatment (Bebbington et al, 2000; Singleton et al, 2001) is even more apparent when considering only cases of severe disorder (Demyttenaere et al, 2004; Wang et al, 2005). There is a socio-economic gradient in the prevalence of depression (Lorant et al, 2003), and those with the lowest socio-economic status might also be the least likely to receive and/or adhere to effective treatment (Acheson, 1998). The aim of this study was to quantify socio-economic inequalities in the delivery of and adherence to treatments of proven clinical effectiveness. We hypothesised that there would be an ‘inverse care law’ – a statistically significant association between low socio-economic status and (under-) treatment of depression after adjusting for the severity of depressive episode.

METHOD

Study design and setting The study was a nested case-control study which received ethical approval from the London Multi-Centre Research Ethics Committee. General practices in England and Wales belonging to the Medical Research Council’s General Practice Research Framework were approached on the basis of location (London and the South East; Trent, Eastern and West Midlands; South West; North, Yorkshire and North West; and Wales), socio-economic deprivation (using Jarman score for practice, banded as high, medium or low), and practice size (single-handed, 2–3 permanent general practitioner principals and ≥4 principals).

Participants Consecutive attendees aged 18–75 years at participating general practices with an appointment to see a doctor, nurse or other professional about themselves were approached in the waiting room. Exclusion criteria included intellectual disability, cognitive impairment that would prevent completion of the study assessments and inability to communicate in English.

Measures Ascertaining ICD–10 depressive episodes in preceding 12 months Individuals who had experienced an ICD–10 depressive episode in the 12 months before interview were identified using a two-stage procedure. Attendees completed a 10-item screening questionnaire (see data supplement to online version of this paper) containing items from the depression section of the 12-month Composite International Diagnostic Interview (CIDI; World Health Organization, 1997). Two stem questions asked whether in the past 12 months the respondent recalled 2 weeks or longer ‘when nearly every day you have felt sad, empty, or depressed for most of the day?’ and ‘when you lost interest in most things like work, hobbies and other things you usually enjoyed?’ Those who answered ‘yes’ to either question were asked to complete a further eight items concerning ‘the time (or times) in the past 12 months when you felt sad, empty, or depressed or when you lost interest in most things nearly every day for 2 weeks or longer’. These items (with yes/no answers) used CIDI items covering fatigue, appetite, weight loss, insomnia, concentration, and feelings of worthlessness, inferiority and guilt. The screening score (range 0–10) was obtained by counting the number of ‘yes’ responses.

A small pilot study compared questionnaire responses (in a general practice waiting room) with the 12-month CIDI depression section administered by telephone approximately 1 week later. Results suggested that an optimum balance between sensitivity and specificity was likely to be achieved using a cut-off of >4 (out of 10), including at least one positive response to the first...
two (stem) questions. The cut-point was chosen to maximise the positive predictive value and hence minimise false-positive second-stage interviews. Those scoring above this level were invited to participate in an interview with a research nurse. At interview, the nurse established the occurrence of one or more episodes of depression using the depression section from the 12-month CIDI. Severity was rated for the time that most symptoms were present concurrently, using the research version of ICD–10 (World Health Organization, 1993).

Treatments for depression interview
Sources and types of help, general practitioner (GP) consultations, and receipt of and adherence to treatments were assessed using a structured interview designed for this study. This was administered after the depression section of the 12-month CIDI. Since many people do not endorse the term ‘depression’, the latter begins by eliciting core symptoms (low mood, loss of interest and/or fatigue) and then referring to these as ‘problems’. The period about which treatment questions were asked was anchored by identifying the month in the preceding year when the respondent had the ‘largest number of problems [symptoms]’ at the same time’. Using CIDI terminology, participants identified the months and years for the onset and offset of the depressive episode.

Participants were reminded about the ‘problems’ they had described in the preceding year and the month when these were at their worst. They were first asked ‘whom did you turn to for help?’, and up to three responses were recorded. Participants were then asked if they had spoken with their GP about these problems, about the timing of the first consultation in respect of this episode and for an estimate of the number of such consultations prior to interview.

Treatments were enumerated, starting with medication. Using a show card with names of all antidepressants listed in the British National Formulary (http://www.bnf.org), participants were asked to identify up to three drugs that they had been prescribed. For each drug mentioned, participants were asked about dose, duration of adherence and how often they remembered to take this. Participants were asked about psychological and other treatments. They were reminded of the month when the index episode had begun, and were asked if they had been referred (by someone else or themselves) to a counsellor (within or outside the practice), psychiatrist, psychologist, psychotherapist or psychoanalyst, or other mental health professional. Taking each in turn, participants were asked how many sessions they had attended, whether they were still attending and reasons for termination.

Criteria for evidence-based treatments
Pharmacological treatment met evidence-based standards where a participant reported having taken a therapeutic dose of an antidepressant for at least 4 weeks at an average frequency of ≥4 days per week. Therapeutic doses were based on guidance from the British National Formulary: ≥75 mg/day of dothiepin or amitriptyline, ≥20 mg/day of fluoxetine or paroxetine, ≥75 mg/day of venlafaxine and ≥50 mg/day of sertraline. There is little evidence about how many sessions of psychological treatment are minimally sufficient. We therefore ruled that this treatment met evidence-based standards where a participant reported referral to a counsellor, psychologist or psychotherapist, and that either: (a) they had attended ≥3 sessions; (b) the treatment had been completed (according the reason give for termination); or (c) they were still attending.

Assessment of socio-economic status
Socio-economic status was assessed using questions about employment status, housing tenure, car access, education and financial strain. The latter was assessed by means of a widely used survey question, the response to which is highly predictive of both current and future psychiatric morbidity (Weich & Lewis, 1998).

Analysis
Analyses were undertaken using survey commands within Stata which adjust standard errors and χ² statistics for clustering (auto-correlation) within practices. For individuals with a confirmed episode of ICD–10 depression in the 12 months before interview, we use logistic regression to calculate unadjusted odds ratios (with 95% confidence intervals) for the associations between socio-economic variables (employment status, housing tenure, car access, education and financial strain) and receipt of and adherence to evidence-based treatments. These associations were subsequently adjusted for confounding by age, gender, depression severity and other socio-economic variables.

RESULTS
Thirty-six general practices took part; 13 were located in the Trent, Eastern and West Midlands regions, 9 in the North, Yorkshire and the North West, 7 in the South West, 6 in London and the South East, and 1 practice was in Wales. Excluding the ineligible, 7718 individuals were asked to take part in screening and 7271 (94.2%) completed the waiting room questionnaire; 2211 (30.4%) scored above the inclusion threshold for interview and 975 individuals (44.1% of those with positive screen results) were interviewed. No statistically significant difference was found in screening score between those who took part in the interview and those who declined or were not available (mean difference −0.04, 95% CI −0.18 to 0.11, P = 0.61).

We identified 866 individuals (88.8% of those interviewed) who had experienced an ICD–10 depressive episode in the preceding 12 months, of whom 812 (93.8%) endorsed all three ICD–10 ‘core’ depressive symptoms (low mood, anhedonia and fatigue). Twelve individuals (1.4%) had experienced a mild depressive episode, 175 (20.2%) a moderate episode and 679 (78.4%) a severe depressive episode. Results are presented for these individuals, with mild and moderate episodes combined owing to small numbers among the former. Among those with a confirmed episode of ICD–10 depression in the 12 months before interview, 72.9% were women. The mean age of the sample was 46.0 years (s.e. = 0.71), with men (mean age 49.3 years) being slightly older than women (44.8 years, P = 0.001). The characteristics of the study sample are shown in Table 1.

Treatments for depression
There were 391 participants (45.2%) who mentioned consulting a ‘doctor’ or ‘GP’ (excluding psychiatrist or other specialist). About three-quarters (n = 294, 75.2%) of those who spontaneously reported consulting their GP concerning the index episode of depression also reported receiving a prescription for antidepressant medication, whereas 108 (27.6%) were offered psychological treatment. Data about medication dose and adherence were available for 272 (92.5%) of the former, of whom 199
(68.4%) received and adhered to pharmacological treatment in keeping with evidence-based criteria. Of the 108 people who consulted their GP and were offered psychological treatment, 68 (63%) met evidence-based criteria for adherence.

There were 447 of 866 individuals (51.6%) with a confirmed ICD–10 depressive episode in the 12 months before interview who reported receiving at least one prescription for an antidepressant drug. Dose and adherence data were available for 405 of these (90.6%), of whom 294 (72.6%) received and adhered to this treatment in keeping with evidence-based criteria. The dose of medication was judged to be sub-therapeutic for 37 (9.1%) individuals and adherence was unsatisfactory among a further 74 (18.3%). Among all of those with complete data \((n=824)\), the rate of receipt of and adherence to evidence-based pharmacotherapy was 35.7% (294 of 824).

A more conservative estimate based on the total sample of confirmed depressive episodes was 33.9% (294/866).

Of the sample of 866 individuals with an ICD–10 depressive episode, 160 (18.5%) reported receipt of psychological therapies for the index depressive episode, of whom 100 (62.5% of those in receipt of psychological treatment, and 11.5% of the sample) met minimum criteria for evidence-based psychological treatment. In total, 332 individuals with ICD–10 depressive episodes (38.3% of the sample) reported receipt of and adherence to evidence-based treatment (either pharmacological or psychological). This figure rose to 41.5% \((n=282)\) among those with a severe depressive episode.

Women, the never married and those with more education were significantly more likely to receive evidence-based treatment (Table 2). Gradients were greatest for psychological treatment. The association between age and evidence-based treatment varied with modality (Fig. 1). Those aged over 60 years were least likely to receive either type of treatment. There were no statistically significant associations between unemployment, car access, or housing tenure and the likelihood of reporting receipt of evidence-based treatment. There was a statistically significant association between greater financial strain and receipt of evidence-based psychological treatment. Those not working because of ill health had high rates of evidence-based treatment.

### Table 1 Characteristics of the study participants with mild/moderate and severe ICD–10 depressive episode in 12 months prior to interview \(n=866\)\(^1\)

<table>
<thead>
<tr>
<th></th>
<th>Mild/ moderate ((n=187))</th>
<th>Severe ((n=679))</th>
<th>(P)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female gender, (n) (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, years: mean (s.d.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Screen score: mean (s.d.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Marital status, (n) (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment status, (n) (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After adjusting for depression severity, age and gender, no association between socio-economic status and evidence-based treatment reached statistical significance (Table 3). Female gender, severe depressive episode and not working because of ill health were independently associated with receipt of either antidepressant and/or psychological treatment meeting minimum evidence-based criteria, to a statistically significant degree. Educational gradients in treatment receipt were largely unaffected, but did not reach statistical significance. There was a strong association between frequency of consultation and treatment. The unadjusted odds ratio for receiving and adhering to any evidence-based treatment among those reporting five or more consultations (23.6% of the sample) \(v\) those reporting 0 or 1 consultations (40.2%) was 22.5 (95% CI 13.7–36.9, \(P<0.0001\)). Many consultations in this group were likely to have been for treatment review. Although there was a trend for socio-economic status to be associated with frequency of consultation for depression, this did not reach statistical significance.

For example, 20.9% of those with no educational qualification were in the highest attendance group, compared with 27.9% of those with A levels or higher qualifications. We found no evidence of any statistically significant interactions between frequency of attendance and socio-economic status in the association with treatment receipt and adherence (e.g. for education, likelihood ratio \(\chi^2=5.12, d.f.=4, P=0.08\)).

### DISCUSSION

#### Main findings

We estimate conservatively that 33.9% of the sample received and adhered to...
We found little evidence of socio-economic differences in rates of treatment for depression, and none that reached statistical significance after adjusting for depression severity, age and gender.

Comparison with other studies

Rates of treatment for depression in primary care

It is difficult to make comparisons with community-based surveys, and we cannot comment on people with depression who do not attend general practice. Previous

treatments for depression meeting minimum evidence-based criteria. Around one-half of those with a depressive episode in the year before interview consulted their GP about this. Three-quarters of these were prescribed an antidepressant, and in over two-thirds of individuals this met evidence-based criteria. Around 12% of the sample received and adhered to evidence-based psychological treatment. The lowest treatment rates were found in older age groups.

Table 2  Associations between receipt of treatments meeting minimum evidence-based criteria and depression severity and characteristics of study participants

<table>
<thead>
<tr>
<th></th>
<th>Antidepressant</th>
<th>P</th>
<th>Psychological treatment</th>
<th>P</th>
<th>Either treatment</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR  95% CI</td>
<td></td>
<td>OR  95% CI</td>
<td></td>
<td>OR  95% CI</td>
<td></td>
</tr>
<tr>
<td>Female (v. male)</td>
<td>1.66 (1.14–2.41)</td>
<td>0.01</td>
<td>2.19 (1.19–4.04)</td>
<td>0.01</td>
<td>1.81 (1.30–2.51)</td>
<td>0.001</td>
</tr>
<tr>
<td>Age 60–75 years¹</td>
<td>0.68 (0.40–1.14)</td>
<td>0.14</td>
<td>0.25 (0.10–0.61)</td>
<td>0.003</td>
<td>0.55 (0.35–0.85)</td>
<td>0.009</td>
</tr>
<tr>
<td>Severe depressive episode³</td>
<td>2.18 (1.60–2.96)</td>
<td>&lt; 0.001</td>
<td>1.63 (0.98–2.72)</td>
<td>0.06</td>
<td>1.93 (1.35–2.77)</td>
<td>0.001</td>
</tr>
<tr>
<td>Marital status (v. married)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>1.41 (0.93–2.14)</td>
<td>0.10</td>
<td>1.62 (0.98–2.69)</td>
<td>0.06</td>
<td>1.51 (1.01–2.24)</td>
<td>0.04</td>
</tr>
<tr>
<td>Separated, divorced or widowed</td>
<td>1.28 (0.92–1.79)</td>
<td>0.14</td>
<td>1.03 (0.57–1.89)</td>
<td>0.91</td>
<td>1.22 (0.87–1.70)</td>
<td>0.24</td>
</tr>
<tr>
<td>Employment (v. employed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>0.86 (0.43–1.72)</td>
<td>0.66</td>
<td>1.20 (0.43–3.35)</td>
<td>0.72</td>
<td>1.00 (0.47–2.13)</td>
<td>0.99</td>
</tr>
<tr>
<td>Not seeking work</td>
<td>1.05 (0.73–1.52)</td>
<td>0.80</td>
<td>0.71 (0.32–1.57)</td>
<td>0.39</td>
<td>1.00 (0.69–1.48)</td>
<td>0.97</td>
</tr>
<tr>
<td>Inactive owing to health</td>
<td>1.75 (1.09–2.81)</td>
<td>0.02</td>
<td>1.83 (0.93–3.60)</td>
<td>0.08</td>
<td>1.68 (1.07–2.64)</td>
<td>0.03</td>
</tr>
<tr>
<td>Rented accommodation³</td>
<td>1.00 (0.77–1.30)</td>
<td>0.99</td>
<td>0.97 (0.71–1.33)</td>
<td>0.83</td>
<td>1.02 (0.79–1.34)</td>
<td>0.85</td>
</tr>
<tr>
<td>No car access</td>
<td>0.89 (0.62–1.29)</td>
<td>0.54</td>
<td>0.79 (0.46–1.38)</td>
<td>0.41</td>
<td>1.01 (0.70–1.46)</td>
<td>0.95</td>
</tr>
<tr>
<td>Financial strain³</td>
<td>1.30 (1.01–1.68)</td>
<td>0.04</td>
<td>1.66 (0.97–2.85)</td>
<td>0.06</td>
<td>1.38 (1.07–1.78)</td>
<td>0.01</td>
</tr>
<tr>
<td>Education (v. A level plus)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCSE or equivalent</td>
<td>0.90 (0.67–1.22)</td>
<td>0.50</td>
<td>0.94 (0.54–1.64)</td>
<td>0.82</td>
<td>0.85 (0.61–1.18)</td>
<td>0.32</td>
</tr>
<tr>
<td>No qualifications</td>
<td>0.81 (0.54–1.23)</td>
<td>0.32</td>
<td>0.55 (0.34–0.89)</td>
<td>0.02</td>
<td>0.69 (0.47–1.02)</td>
<td>0.07</td>
</tr>
</tbody>
</table>

1. Versus participants aged 18–29 years.
2. Versus mild and moderate episodes.
3. Versus owner-occupiers.

Table 3  Associations between receipt of treatments meeting minimum evidence-based criteria and depression severity and characteristics of study participants, adjusted for the other variables

<table>
<thead>
<tr>
<th></th>
<th>Antidepressant</th>
<th>P</th>
<th>Psychological treatment</th>
<th>P</th>
<th>Either treatment</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR  95% CI</td>
<td></td>
<td>OR  95% CI</td>
<td></td>
<td>OR  95% CI</td>
<td></td>
</tr>
<tr>
<td>Female (v. male)</td>
<td>1.71 (1.12–2.60)</td>
<td>0.01</td>
<td>2.31 (1.18–4.49)</td>
<td>0.02</td>
<td>1.87 (1.27–2.75)</td>
<td>0.002</td>
</tr>
<tr>
<td>Age 60–75 years¹</td>
<td>0.91 (0.49–1.69)</td>
<td>0.76</td>
<td>0.47 (0.17–1.27)</td>
<td>0.13</td>
<td>0.78 (0.45–1.35)</td>
<td>0.36</td>
</tr>
<tr>
<td>Severe depressive episode³</td>
<td>1.91 (1.35–2.70)</td>
<td>0.001</td>
<td>1.20 (0.69–2.11)</td>
<td>0.51</td>
<td>1.61 (1.11–2.32)</td>
<td>0.01</td>
</tr>
<tr>
<td>Employment (v. employed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>0.80 (0.37–1.72)</td>
<td>0.55</td>
<td>1.06 (0.42–2.63)</td>
<td>0.91</td>
<td>0.93 (0.43–2.03)</td>
<td>0.86</td>
</tr>
<tr>
<td>Not seeking work</td>
<td>1.31 (0.87–1.99)</td>
<td>0.19</td>
<td>0.77 (0.38–1.55)</td>
<td>0.45</td>
<td>1.28 (0.83–1.97)</td>
<td>0.26</td>
</tr>
<tr>
<td>Inactive owing to health</td>
<td>1.98 (1.29–3.04)</td>
<td>0.003</td>
<td>2.28 (1.26–4.13)</td>
<td>0.008</td>
<td>1.99 (1.32–3.00)</td>
<td>0.002</td>
</tr>
<tr>
<td>Financial strain³</td>
<td>1.21 (0.95–1.53)</td>
<td>0.12</td>
<td>1.50 (0.91–2.48)</td>
<td>0.11</td>
<td>1.29 (0.97–1.70)</td>
<td>0.08</td>
</tr>
<tr>
<td>Education (v. A level plus)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GCSE or equivalent</td>
<td>0.84 (0.62–1.14)</td>
<td>0.25</td>
<td>0.83 (0.49–1.40)</td>
<td>0.48</td>
<td>0.78 (0.56–1.09)</td>
<td>0.14</td>
</tr>
<tr>
<td>No qualifications</td>
<td>0.81 (0.50–1.32)</td>
<td>0.39</td>
<td>0.61 (0.34–1.11)</td>
<td>0.10</td>
<td>0.71 (0.44–1.11)</td>
<td>0.13</td>
</tr>
</tbody>
</table>

1. Versus participants aged 18–29 years.
2. Versus mild and moderate episodes.
Studies report that 30–40% of those in the community who are depressed receive medical treatment (medication or psychological treatments; Lin & Parikh, 1999; Singleton et al., 2001), falling to 25% or less when only treatments of ‘minimal’ adequacy or better are included (Young et al., 2001).

Rates of ‘minimally adequate’ guideline-based treatment for depression may be higher in primary care settings in the UK than in the USA. In the National Comorbidity Study Replication (NCS–R; Wang et al., 2005), data from general medical settings (including primary care) indicated that 15% of those with major depressive disorder in the past year received treatment meeting these criteria, compared with 42% with a severe depressive episode in our study. Of those who reported consulting their GP for the index episode, 67% received and adhered to minimally adequate treatment, compared with 38% in the NCS–R sample.

**Evidence for ‘inverse care’?**

Apart from a qualitative study highlighting the challenges facing GPs working in socio-economically deprived areas (Chew-Graham et al., 2002), there is little evidence of ‘inverse care’ in the treatment of depression. There was little evidence in our study that socio-economic status was associated with receipt of and adherence to evidence-based treatments. Those not working because of ill health were more likely to report receiving evidence-based treatment, perhaps because depression was the certified cause of absence. Greater financial strain (a robust socio-economic indicator) was associated with more evidence-based treatment. Moreover, whereas those with the lowest educational attainment were less likely to report evidence-based psychological treatment, this association was confounded by age, gender and depression severity.

Lower socio-economic status is associated with higher primary care consultation rates in the UK (Carr-Hill et al., 1996). Although most existing studies report little evidence that demographic factors affect help-seeking when depressed, few measured individual socio-economic status precisely or attempted to adjust rates of help-seeking for the presence and severity of psychiatric disorder (Lin & Parikh, 1999). Although an early US study reported lower needs-adjusted rates of medical consultation among those with the lowest socio-economic status (Gallo et al., 1995), this association was not replicated in the NCS–R (Wang et al., 2005).

**Strengths and limitations of the study**

This is the largest UK study to test the hypothesis of an ‘inverse care law’ in the treatment of depression. Participants were recruited from among consecutive primary care attendees, and depressive episodes were confirmed using a validated, standardised clinical interview. Participating practices were recruited from across England and Wales. By recruiting individuals with a depressive episode at any time in the preceding 12 months, our sample was better suited to assessing treatment receipt and adherence than samples of current depression (Singleton et al., 2001).

We are not aware of validated screening questionnaires for identifying individuals with a recent history of depressive episodes from among primary care attendees. We based the screening questionnaire on the ‘gold standard’ of the 12-month CIDI depression interview. A primary care study in New Zealand examined GPs’ verbal use of the first two (stem) items of our waiting room questionnaire (i.e. low mood and anhedonia) as a screen for depression in the past month (Arroll et al., 2003). Among 421 attendees not taking psychotropic medication, sensitivity and specificity of this two-item screen were 97 and 67% respectively, and the positive predictive value was 18%. A high positive predictive value was needed to minimise false-positive second-stage interviews. This weighted the sample towards individuals with more severe depressive episodes.

One weakness was the high attrition rate between screening and interview. Although fewer than half of those eligible for the second stage were interviewed, screening questionnaire scores did not differ significantly between those who were and were not interviewed. Nevertheless, we cannot exclude the possibility that those who were interviewed differed in other ways, including their willingness to receive and adhere to treatment. Estimates of treatment rates should be interpreted carefully (Ayuso-Mateos et al., 2001), but it is unlikely that this would have resulted in biased estimates of association between socio-economic status and treatment receipt. The size of this association (point estimate) would remain the same unless there was an association between participation in the study, socio-economic status and treatment. In other words, our findings would only have been biased away from an association between low socio-economic status and undertreatment if individuals of low socio-economic status who did not take part were less likely to have received evidence-based treatment than the individuals of low socio-economic status who did participate. Although this was possible, it would appear unlikely, particularly since treatment rates were already very low in the latter group. We note also that our sample was relatively deprived compared with the population of England and Wales as a whole (2001 UK Census; http://www.statistics.gov.uk) on indices that included employment (study sample 51% in paid work v. 61% nationally), no educational qualification (27 v. 28% nationally) and financial strain (34 v. 14% nationally; Weich & Lewis, 1998).

A further limitation was the location of this study in primary care rather than the community. It is possible that individuals of low might not seek medical care for depression. This would appear to be borne out by the finding that 38% of those with a current depressive episode in a UK community sample had no educational qualifications, compared with 27% of our sample (Singleton et al., 2001). Although our data revealed a non-significant trend towards less frequent consultation for the index episode of depression among those without educational qualifications, there was no evidence that frequency of consultation modified the association between socio-economic status and treatment receipt and adherence. A strength of the study was that recruitment was undertaken solely on the basis of waiting to see a health professional, and not on any particular reason for consultation. Although people living on low incomes may decline or disengage from care, most continue to have contact with the healthcare system (Edlund et al., 2002; Anderson et al., 2006). We are confident that such individuals would have been invited to participate in this study, but it is not possible to generalise directly from the present findings to all those with depressive episodes in the community.

Quantifying treatment adherence is challenging and there is no agreed gold standard (Garber et al., 2004; DiMatteo & Haskard, 2006). We relied on self-report within a comprehensive, structured, face-to-face interview. We began by eliciting
evidence of a depressive episode and then used this to anchor the remainder of the interview. Although not formally validated, our approach was thorough and systematic. It is unlikely that prescribing rates were overreported. The high rate (almost 85%) of reported adherence (taking prescribed antidepressant medication ‘every day’ or ‘nearly every day’ for more than 4 weeks) may reflect overreporting. Although participants in this study may have been more likely to have adhered to treatment than those who declined to take part, a US study found that over 40% of individuals adhered to antidepressants for 6 months.

Under- or overestimates of treatment receipt and adherence are only important here if biased by socio-economic status. For this to have concealed inverse care in the treatment of depression, delivery and uptake of treatment would have to have been either systematically overreported by those with the lowest socio-economic status and/or underreported by those with higher socio-economic status. Neither was likely.

Implications for services
Depression is closely associated with socio-economic deprivation across the life span. Rates of depression are highest in areas and practices with the fewest resources. It is reassuring that those with the lowest socio-economic status in England and Wales are as likely as the more affluent to receive and adhere to evidence-based treatments, after accounting for clinical need. Nevertheless, treatment rates remain modest and the reasons for this are unclear. Older age may be a greater source of inequality in the treatment of depression than socio-economic deprivation.

Acknowledgements
We thank our colleagues at the General Practice Research Framework, particularly Louise Letley, and the many research nurses who collected the data. The study was funded by the Medical Research Council, under the Health of the Public Initiative (ref.G9900548).

REFERENCES

(Received 12 October 2006, final revision 25 January 2007, accepted 23 February 2007)

Correspondence: Professor Scott Weich, Health Sciences Research Institute, Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK. Email: s.weich@warwick.ac.uk
Characteristics and activities of acute psychiatric in-patient facilities: national survey in Italy

GIOVANNI DE GIROLAMO, ANGELO BARBATO, RENATA BRACCO, ANDREA GADDINI, ROSELLA MIGLIO, PIERLUIGI MOROSINI, BRUNO NORCIO, ANGELO PICARDI, ELISABETTA ROSSI, PAOLA RUCCI, GIOVANNI SANTONE and GIUSEPPE DELL’ACQUA

Background Legislation in 1978 led to the gradual replacement of mental hospitals in Italy with a full range of community-based services, including facilities for acute in-patient care.

Aims To survey the main characteristics of Italian public and private in-patient facilities for acute psychiatric disorders.

Method Structured interviews were conducted with each facility’s head psychiatrist in all Italian regions, with the exception of Sicily.

Results Overall, Italy (except Sicily) has a total of 4108 public in-patient beds in 319 facilities, with 0.78 beds per every 10,000 inhabitants, and 4862 beds in 54 private in-patient facilities, with 0.94 beds per 10,000 inhabitants. In 2001 the rates of psychiatric admissions and admitted patients per 10,000 inhabitants were 26.7 and 17.8 respectively. In the same year the percentage of involuntary admissions was 12.9%, for a total of 114,570 hospital days. Many in-patient facilities showed significant limitations in terms of architectural and logistic characteristics. Staffing showed a great variability among facilities.

Conclusions The overall number of acute beds per 10,000 inhabitants is one of the lowest in Europe. The survey has provided evidence of two parallel systems of in-patient care, a public one and a private one, which are not fully interchangeable.

Declaration of interest None.

In Italy, legislation in 1978 established the gradual replacement of mental hospitals with a range of community-based services, including general hospital psychiatric units and, to a lesser extent, other facilities for acute in-patient care. Law 180 stated that each general hospital unit should have no more than 15 beds, to prevent the establishment of large-scale asylum-like wards (de Girolamo & Cozza, 2000). Since then, all public mental hospitals have been closed, and public in-patient care is now provided by general hospital psychiatric units, university psychiatric clinics and (in limited areas of the country) community mental health centres operating 24 h a day. There are also 54 private in-patient facilities, which were already in place before Law 180 was enacted; fees of patients admitted to these facilities are covered by the national health service. This study reports the first phase of a national survey, the Progetto Residenze (PROGRES) Acute study, funded by the Italian Ministry of Health and jointly coordinated by the Italian National Institute of Health and the Department of Mental Health of Trieste, aimed at obtaining data on the physical characteristics, staffing arrangements, admission rules and activities of in-patient facilities.

METHOD

All 21 Italian regions agreed to participate in the study, except for Sicily. All public and private acute in-patient facilities admitting patients with a primary diagnosis of mental disorder were to be assessed.

Data collection Information about the number and location of facilities was obtained from the regional health authorities and departments of mental health. The project began in 2001 and data were collected in 2002–2003. Each facility’s head psychiatrist (previously informed by letter about the study) was administered a structured interview, previously validated and adapted from the ‘facility form’ used for a national survey of non-hospital residential facilities in Italy (de Girolamo et al., 2002; Picardi et al., 2006).

All activity data reported here refer to the year 2001 and were obtained through local computerised case registers or calculated manually for the 143 facilities (40.8%) with no electronic information system available. After data collection, thorough quality control was performed, first locally in each region, and then centrally. The admission episodes considered included both episodes of care when patients were directly admitted to a psychiatric unit, and episodes of care for patients initially admitted to another specialty ward (e.g. transfer from a medical ward after a self-harm episode). For the calculation of staffing levels, we included all staff directly involved in patient care, including psychiatrists, psychologists, nurses, nursing aids and social workers.

To assess the variability of several indicators throughout the country, we grouped the regions into four major areas, following the European Union recommendations (European Union, 2003): north-west, north-east, central and southern (including Sardinia).

Statistical analysis In-patient facility characteristics and activity data were summarised using descriptive statistics. All analyses were conducted using the Statistical Package for the Social Sciences version 12.01 for Windows.

RESULTS

In the 20 participating regions, the project surveyed 262 general hospital psychiatric units, with a total of 3431 beds; 23 university psychiatric clinics, with 399 beds; and 16 community mental health centres, located in Friuli-Venezia Giulia (n=10) and in Campania (n=6), with 98 beds. Nearly all public in-patient facilities participated in the survey. There were only four refusals (three general hospital units and one university clinic), all located in Lombardy. Overall, Italy (excluding Sicily) had 4108 public acute in-patient beds available (including the four units not included in the survey), with 0.78 beds for 10,000 inhabitants.
We also surveyed eight medical wards (limited to some areas in Tuscany) with a total of 20 beds available for psychiatric in-patients, and six additional in-patient facilities with 98 beds, which, however, did not admit involuntary patients. These latter facilities were located outside general hospitals and admitted mainly medium-term patients. As these units presented a different profile from the main three groups, we excluded them from subsequent analyses (together with the four refusals). There were 54 private in-patient facilities, with 4862 beds and 0.94 beds per 10 000 inhabitants; these units were located in ten regions, and all took part in the study. These facilities have only psychiatric beds, and cannot admit compulsory patients.

The overall number of acute, short-term psychiatric beds (public and private) in Italy (except Sicily) is therefore 8970, i.e. 1.72 per 10 000 inhabitants. Three regions (Lazio, Campania and Calabria), located in central and southern Italy, which have the fewest public acute beds, also showed the highest concentration of private beds. Table 1 summarises the main organisational characteristics of Italian in-patient facilities.

All but two of the general hospital psychiatric units were opened in the decade after the 1978 reform law was approved and nearly 60% of the 24 h community mental health centres opened in the past 6 years; in contrast, all private facilities but one were already operational in 1978.

**Physical characteristics**

Table 2 shows the main physical and logistic characteristics and the staffing of Italian acute in-patient facilities. Most public facilities (n=230, 81.3%) were located in buildings built before 1980, and only 37 (13.1%) facilities had been built since 1991. As for private facilities, all but one had been built before 1972. All general hospital units and university clinics were functionally integrated with a general hospital. Half of the general hospital units were located on the ground floor, but seven (3%) were situated in basements, and the remainder on upper floors. Approximately two-thirds of general hospital units and university clinics had an outdoor area available for in-patients, whereas this was available for all private facilities.

A considerable proportion of general hospital units (42%, n=111) and two-thirds of university clinics had no single rooms available, and only 8 of the general

---

### Table 1  In-patient facility organisation and management characteristics

<table>
<thead>
<tr>
<th>Facility characteristics</th>
<th>General hospital psychiatric units (n=262)</th>
<th>University psychiatric clinics (n=23)</th>
<th>24 h CMHCs (n=16)</th>
<th>Private in-patient facilities (n=54)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of beds: mean (s.d.)</td>
<td>13.1 (4.2)</td>
<td>17.3 (8.7)</td>
<td>6.1 (2.0)</td>
<td>90.0 (48.2)</td>
</tr>
<tr>
<td>Opening year of current structure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Up to 1977</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1978–1990</td>
<td>141 (53.8)</td>
<td>7 (30.4)</td>
<td>1 (6.2)</td>
<td>51 (94.3)</td>
</tr>
<tr>
<td>1991–2003</td>
<td>121 (46.2)</td>
<td>9 (39.2)</td>
<td>15 (93.8)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Acceps compulsory admissions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>260 (99.2)</td>
<td>13 (56.5)</td>
<td>12 (75.0)</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>2 (0.8)</td>
<td>10 (43.5)</td>
<td>4 (25.0)</td>
<td>54 (100.0)</td>
</tr>
<tr>
<td>Daytime facility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>136 (51.9)</td>
<td>11 (47.8)</td>
<td>1 (6.2)</td>
<td>46 (86.8)</td>
</tr>
<tr>
<td>Day hospital</td>
<td>115 (43.9)</td>
<td>12 (52.2)</td>
<td>2 (12.5)</td>
<td>6 (11.3)</td>
</tr>
<tr>
<td>Day centre</td>
<td>6 (2.3)</td>
<td>1 (4.3)</td>
<td>6 (37.5)</td>
<td>1 (1.9)</td>
</tr>
<tr>
<td>Both</td>
<td>4 (1.5)</td>
<td>6 (26.1)</td>
<td>7 (43.8)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>1 (0.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Out-patient clinic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>106 (40.5)</td>
<td>2 (8.7)</td>
<td>2 (12.5)</td>
<td>38 (73.1)</td>
</tr>
<tr>
<td>Out-patient centre</td>
<td>121 (46.2)</td>
<td>21 (91.3)</td>
<td>2 (12.5)</td>
<td>14 (26.9)</td>
</tr>
<tr>
<td>Crisis centre</td>
<td>16 (6.0)</td>
<td>8 (36.0)</td>
<td>3 (18.8)</td>
<td></td>
</tr>
<tr>
<td>Both</td>
<td>13 (5.0)</td>
<td>9 (39.1)</td>
<td>9 (56.3)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>6 (2.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Defined catchment area</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>261 (99.6)</td>
<td>14 (60.9)</td>
<td>16 (100.0)</td>
<td>7 (13.0)</td>
</tr>
<tr>
<td>No</td>
<td>1 (0.4)</td>
<td>9 (39.1)</td>
<td>16 (100.0)</td>
<td>47 (87.0)</td>
</tr>
<tr>
<td>Catchment area population1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Up to 100 000</td>
<td>39 (14.9)</td>
<td>1 (8.3)</td>
<td>13 (81.3)</td>
<td></td>
</tr>
<tr>
<td>100 000–250 000</td>
<td>173 (66.3)</td>
<td>7 (58.4)</td>
<td>2 (12.5)</td>
<td>2 (33.3)</td>
</tr>
<tr>
<td>250 000–500 000</td>
<td>42 (16.1)</td>
<td>3 (25.0)</td>
<td>1 (6.2)</td>
<td>2 (33.3)</td>
</tr>
<tr>
<td>Over 500 000</td>
<td>7 (2.7)</td>
<td>1 (8.3)</td>
<td>2 (33.3)</td>
<td></td>
</tr>
</tbody>
</table>

CMHC, community mental health centre.

1. Data on structures with a defined catchment area that provided information on the area inhabitants.
hospital units (3%) had at least five single rooms; also 8 (15%) private facilities had no single rooms. At least one three-bedded room was present in 126 general hospital units (48.1%), in 12 (52.2%) university clinics and in 39 (82.2%) private facilities. Availability of patient bathrooms varied: general hospital units had a mean of 0.4 (s.d.=0.2) bathrooms for each bed, and university clinics a mean of 0.3 (s.d.=0.2); the mean was slightly higher for the 24 h community mental health centres (0.6, s.d.=0.3). Private facilities had a value similar to the general hospital units (0.4, s.d.=0.1).

Most facilities (95% of general hospital units, 91% of university clinics and 83% of the community centres) had on average three or four rooms available for clinical activities and meetings with patients’ families; some had just one or two rooms. In private facilities, rooms for clinical evaluations were available everywhere, and 40% of them had more than four rooms.

**Staffing**

All public and private facilities had 24 h coverage with staff on duty at night. The 301 public facilities employed 8058 full-time equivalent professionals (excluding medical residents and volunteers), 86.5% (n=6971) of whom worked full-time. The number of staff in private facilities was substantially smaller (n=2384, of whom 1918 were working full-time); this is reflected in a much smaller equivalent full-time staff per bed in private facilities compared with any type of public facility (Table 2). There was marked variation in the total number of equivalent staff among the three public facility types, with general hospital units having 1.5 times more staff available than the university clinics, and 24 h community mental health centres having more than double the staff per bed than the general hospital units (this is probably accounted for by the fact that in the community centres the staff are also employed for out-patient care). This variation involved all staff types.

Most public facilities (n=239, 79.0%) had a full-time medical staff; in the remainder staff additionally worked part-time in other mental health services such as out-patient and residential facilities. All private facilities but two had only full-time medical and nursing staff.

Clinical supervision of various intensity and quality was available in 107 general hospital units (41.6%) and 29 private facilities (53.7%).

**Exclusion criteria for admission and length of stay**

Most facilities used exclusion criteria for admission (Table 3). Criteria varied considerably per facility type and within each facility type. In almost all private facilities admission criteria excluded the most difficult patients.

More than 90% of general hospital units (n=239) and university clinics (n=21) did not specify a maximum length of stay. When they did, it was set at 15 days. A maximum length of stay was spelled out by 23 (42.6%) private facilities; this ranged from 15 to 90 days, with the most common length being 60 days. The highest average length of stay was found in private facilities, followed by the community centres, the university clinics and the general hospital units. Even median values showed a substantial difference across facilities.
Table 3  Exclusion criteria and average length of stay

<table>
<thead>
<tr>
<th>Exclusion criteria, n (%)</th>
<th>General hospital psychiatric units (n=262)</th>
<th>University psychiatric clinics (n=23)</th>
<th>24 h CMHCs with beds (n=16)</th>
<th>Private in-patient facilities (n=54)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age &lt;18 years</td>
<td>93 (56.7)</td>
<td>6 (37.5)</td>
<td>13 (92.9)</td>
<td>24 (44.4)</td>
</tr>
<tr>
<td>Pure alcohol misuse/dependence</td>
<td>117 (71.3)</td>
<td>8 (50.0)</td>
<td>14 (100.0)</td>
<td>28 (51.9)</td>
</tr>
<tr>
<td>Alcohol misuse/dependence with comorbidity</td>
<td>4 (2.4)</td>
<td></td>
<td></td>
<td>41 (75.9)</td>
</tr>
<tr>
<td>Substance misuse</td>
<td>136 (82.9)</td>
<td>8 (50.0)</td>
<td>14 (100.0)</td>
<td>10 (18.5)</td>
</tr>
<tr>
<td>Substance misuse with comorbidity</td>
<td>6 (3.7)</td>
<td></td>
<td></td>
<td>38 (70.4)</td>
</tr>
<tr>
<td>Organic mental disorders</td>
<td>49 (29.9)</td>
<td>4 (25.0)</td>
<td>10 (71.4)</td>
<td>34 (63.0)</td>
</tr>
<tr>
<td>Severe mental retardation</td>
<td>69 (42.1)</td>
<td>6 (37.5)</td>
<td>4 (28.6)</td>
<td>22 (40.7)</td>
</tr>
<tr>
<td>High suicide risk</td>
<td>3 (1.8)</td>
<td>1 (6.3)</td>
<td>5 (35.7)</td>
<td>35 (64.8)</td>
</tr>
<tr>
<td>Severe behavioural disorders with aggressiveness</td>
<td>4 (2.4)</td>
<td>3 (18.8)</td>
<td>4 (28.6)</td>
<td>27 (50.0)</td>
</tr>
<tr>
<td>Severe physical disabilities</td>
<td>93 (56.7)</td>
<td>8 (50.0)</td>
<td>11 (78.6)</td>
<td>18 (33.3)</td>
</tr>
<tr>
<td>Severe medical disorders</td>
<td>82 (50.0)</td>
<td>7 (43.8)</td>
<td>10 (71.4)</td>
<td>20 (37.0)</td>
</tr>
<tr>
<td>Prison inmates</td>
<td>91 (55.5)</td>
<td>10 (62.5)</td>
<td>3 (21.4)</td>
<td>31 (57.4)</td>
</tr>
<tr>
<td>Other</td>
<td>152 (92.7)</td>
<td>15 (93.8)</td>
<td>13 (92.9)</td>
<td>2 (3.7)</td>
</tr>
</tbody>
</table>

Length of stay, days

<table>
<thead>
<tr>
<th>Mean (s.d.)</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>General hospital psychiatric units (n=262)</td>
<td>12.0 (3.4)</td>
</tr>
<tr>
<td>University psychiatric clinics (n=23)</td>
<td>18.5 (7.1)</td>
</tr>
<tr>
<td>24 h CMHCs with beds (n=16)</td>
<td>37.0 (55.3)</td>
</tr>
<tr>
<td>Private in-patient facilities (n=54)</td>
<td>39.7 (17.8)</td>
</tr>
</tbody>
</table>

CMHC, community mental health centre.

Table 4 shows activity data for the facilities, during the year 2001. Overall, psychiatric admission and admitted patient rates per 10,000 inhabitants in public facilities were 19.8 and 13.4 respectively, whereas in private facilities the rates were 6.9 and 4.4 respectively. There was a mean of 363.0 (s.d. = 154.7) yearly admissions for each general hospital unit, with a similar figure for university clinics (342.7, s.d. = 197.5) and a lower one for the community facilities (111.9, s.d. = 49.7). Private facilities had a higher mean number of admissions (676.3, s.d. = 409.5).

In total there were 103,260 acute admissions in the facilities for which admission data were available (n=292), with 1,227,676 hospital days. In private facilities admissions totalled 35,880 (n=53), with 1,252,049 bed-days (n=47). In the 284 public facilities providing data on individual patients, there were 70,062 admitted patients in total; in private facilities admitted patients numbered 23,097. The admission/patient ratio ranged from 1.34 in university psychiatric clinics to 1.45 in general hospital units and 1.66 in the community mental health centres, in private facilities the ratio was 1.55.

In public facilities there were 22,893 (32.7%) patients having their first-ever admission to a given facility (but possibly previously admitted to other public or private facilities); 22.3% of admitted patients were at their second admission during the year, and 8.7% had had three or more admissions to the same facility (representing ‘revolving-door’ patients). In private facilities the percentage of patients having their first-ever admission was even higher (40.9%; n=9446); the numbers of patients with two and three or more admissions were 3414 (14.8%) and 1911 (8.3%) respectively.

A mean percentage of 15.6% (s.d. = 12.5) of admitted patients in the 259 general hospital units that provided...
Table 4 Activity data of in-patient facilities (year 2001)

<table>
<thead>
<tr>
<th></th>
<th>General hospital psychiatric units</th>
<th>University psychiatric clinics</th>
<th>24 h CMHCs with beds</th>
<th>Private in-patient facilities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Rate</td>
<td>n</td>
<td>Rate</td>
</tr>
<tr>
<td>Admissions</td>
<td>260</td>
<td>18.1</td>
<td>23</td>
<td>1.5</td>
</tr>
<tr>
<td>Patients admitted</td>
<td>255</td>
<td>12.4</td>
<td>20</td>
<td>0.9</td>
</tr>
<tr>
<td>Hospital days</td>
<td>255</td>
<td>205.8</td>
<td>23</td>
<td>26.2</td>
</tr>
<tr>
<td>Patients first admitted ever</td>
<td>203</td>
<td>4.0</td>
<td>17</td>
<td>0.3</td>
</tr>
<tr>
<td>Patients with 2 admissions</td>
<td>249</td>
<td>1.7</td>
<td>20</td>
<td>0.1</td>
</tr>
<tr>
<td>in the same year</td>
<td>249</td>
<td>1.1</td>
<td>19</td>
<td>0.1</td>
</tr>
<tr>
<td>Compulsory admissions</td>
<td>254</td>
<td>2.4</td>
<td>13</td>
<td>0.1</td>
</tr>
<tr>
<td>Patients admitted compulsorily</td>
<td>242</td>
<td>1.9</td>
<td>13</td>
<td>0.1</td>
</tr>
<tr>
<td>Hospital days for compulsory patients</td>
<td>231</td>
<td>20.8</td>
<td>13</td>
<td>1.2</td>
</tr>
<tr>
<td>Patients with 2 or more compulsory admissions in the same year</td>
<td>241</td>
<td>0.2</td>
<td>13</td>
<td>0.01</td>
</tr>
</tbody>
</table>

CMHC, community mental health centre; NA, not applicable.
1. Number of facilities providing data.
2. Rate per 10 000 population.

this information came from outside the catchment area; a similar percentage held for university clinics (n=12; 18.5%, s.d.=19.1), but this figure was lower for the 24 h community centres (n=16; 3.9%, s.d.=6.7). In the few private facilities with a defined catchment area (n=7) the percentage of patients coming from outside this area was low (10.2%, s.d.=23.3).

Compulsory admissions
The percentage of compulsory admissions was 12.9% (n=12 793), with a total of 114 570 hospital days. Over the course of the year, 1.4% (n=966) of patients had had two or more compulsory admissions. Figure 2 shows admission rates and percentages of compulsory admissions in public facilities in the four geographical areas, and also the number of public and private beds per 10 000 inhabitants. Where public beds are scarce, as in the south of Italy, compulsory admissions are almost twice as frequent as in other areas. An inverse relationship is also apparent between the numbers of public beds (lowest) and private beds (highest) in the south.

DISCUSSION
This study is the first survey of acute inpatient services conducted in Italy on a nationwide scale. Our results should be interpreted keeping in mind that 24 h home treatment is available in very few areas of the country. Few units specific for alcohol and drug detoxification, or for eating disorders are operating in Italy, and there is no specialised unit for early intervention. There are only six forensic mental hospitals in Italy, all but one of which are run by the Ministry of Justice. Specialised in-patient units admitting only adolescents are rare; in most cases adolescents with severe behavioural disorders requiring in-patient care are treated in general hospital psychiatric units.

Availability of acute in-patient beds
In Italy the number of psychiatric public acute beds shows marked variation between regions; overall, public beds represent 45.8% of all acute in-patient beds. Moreover, almost all private facilities have no defined catchment area, and admit patients from different parts of the country; correlations between public and private facilities on a regional scale should therefore be considered with caution. All private in-patient facilities were built long before the reform law. Although in the north of the country the proportions of public and private beds are similar, in the central and southern regions the percentage of private beds is higher, and in the south the ratio of private to public beds is approximately 2:1. This indicates a very uneven distribution of public and private beds across the country.

It should be highlighted that private in-patient facilities are very different from the former mental hospitals, in terms of size (the 99 public mental hospitals active in
Logistic and environmental characteristics of in-patient facilities

Many facilities suffer from major logistic and architectural limitations: 3% of general hospital psychiatric units are located in basements; 42% of them have no single bedrooms; and many facilities have a considerable proportion of rooms with three or four beds (hospital rooms with more than four beds are forbidden in Italy). More than a third of general hospital units have no outdoor area for in-patients, nearly 40% have no living room other than a dining room, and some general hospital units and university clinics have no room specifically designated for clinical activities. Private facilities show a mixed picture, including both negative and positive aspects.

The unsatisfactory physical features of many general hospital psychiatric units raise serious problems. Indeed, the paramount influence of the physical in-patient setting environment on clinical and interpersonal variables – including patient and staff satisfaction – has long been recognised (Gutkowski et al., 1992; Horsburgh, 1995; Royal College of Psychiatrists, 1998; Tyson et al., 2002).

Length of stay

Private facilities presented a high average length of stay. This finding, together with other features of private facilities such as the case mix and the staffing levels, raises some doubts about their definition as acute care services. The longer stay of patients in university clinics compared with general hospital units may be due in part to a different case mix (e.g. complex cases referred for additional evaluation and treatment). The longer stay in 24-h community mental health centres is due to a difference between these services and general hospital units or university clinics – the community centres are more frequently used as flexible tools in response to a variety of critical patient needs (crisis resolution, housing needs, employment problems, family conflicts, etc.) rather than just to treat acute episodes of illness (Dell’Acqua & Mezzina, 1988; Norcio et al., 2001).

Studies of local services in Italy have consistently shown a shorter duration of in-patient stay in public facilities compared with other European countries (McCrone & Lorusso, 1999; Sytema et al., 2002). The median length of stay in general hospital units found in our study (11.4 days) was notably lower than that found in England in 2003 (18 days; Glover, 2007).

Open v. locked doors

Most general hospital units (nearly 80%) and two-thirds of university clinics had locked doors. In contrast, almost none of the community centres had a locked-door policy. Although most private facilities did not have doors always locked, they did not admit involuntary patients. In England, a 1-day survey of 118 acute psychiatric units, conducted during unannounced visit, found that only 9% of the units were locked (Ford et al., 1998). Bowers et al. (2002), in a survey of 87 acute units in the London area, found only 25% of units to be constantly locked and 45% locked only occasionally. In another survey involving 24 in-patient services in five European countries (Austria, Hungary, Romania, Slovakia and Slovenia), located in both mental and general hospitals, only 21.4% of 4191 in-patients evaluated on census day were found to be accommodated in locked units (Rittmannsberger et al., 2004). In brief, locked units admitting acute patients seem to be more common in Italy than in other European countries; since there are fewer beds, these units tend to admit a selected in-patient population with challenging behaviours, which can partly account for this finding.

Staffing

In public facilities the staff/patient ratio ranged from 1.44 to 5.17 showing that facilities for acute patients rely greatly on human resources; in contrast, ratios for private facilities were markedly smaller. The number of staff working in Italian public facilities is higher than that found in other countries, although some of the differences may be explained by the type of professionals included in the count. In the UK, the ratio of staff to resident beds in acute in-patient facilities was 1.27 (Lelliott et al., 1996), whereas studies conducted in the USA found staff/resident ratios ranging from 0.32 to 2.08 (Coleman & Paul, 2001), or higher than 3.3 (Donat, 2003); this variability suggests that there are no strict quantitative standards for staff.

The low staffing level in private facilities (0.45 staff/patient ratio) can be partly explained by a daily variation in the number of staff on site, with fewer staff present during the night shifts, because the larger size of private facilities makes possible

1978 each had 642 beds on average), physical characteristics (mental hospitals were generally old-fashioned and physically inadequate, with large numbers of patients per room, and low standards of diet, clothing and medical care) and case typology (most mental hospital patients were long-stay with admissions often lasting years, many were compulsory). It is therefore impossible to equate current private in-patient facilities with the former mental hospitals.

However, public and private facilities tend to admit a different case mix, as demonstrated by the presence of involuntary patients only in public facilities and by the exclusion criteria operated by private facilities which tend to rule out patients with challenging behaviour (e.g. violent or suicidal behaviour or comorbid substance misuse). A national census day of all in-patients (n = 7984) in acute, public and private facilities, conducted on 8 May 2003, found large differences between public and private facilities for the age and gender distribution of in-patients: more men under 35 years old had been admitted to public facilities, whereas women aged 65 years or over had more frequently been admitted to private facilities (further information available on request).

International comparison of the number of psychiatric beds is difficult (Health and Consumer Protection Directorate General, 2004; World Health Organization, 2003), because the very definition of ‘psychiatric bed’ is often unclear. World Health Organization (WHO) Mental Health Atlas data (World Health Organization, 2003) reveal that in western Europe the number of acute public in-patient beds shows substantial variation between 9.0 (Finland) and 0.3 (Greece) per 10,000 inhabitants. However, it is often difficult from WHO data to draw a clear-cut distinction between acute and long-term beds. Moreover, private and forensic beds are not included in the WHO statistics, although private beds can account for a substantial proportion of the total number of psychiatric beds, as in Italy. Although there are also six forensic mental hospitals in Italy, in 2002 the number of forensic beds was very low (0.22/10,000) compared with Sweden (1.43/10,000), The Netherlands (1.14/10,000), and Germany (0.78/10,000), and only slightly higher than in England (0.18/10,000) and Spain (0.15/10,000) (Priebe et al., 2005). Therefore, the overall rate of acute beds in Italy (1.72/10,000) is one of the lowest in western Europe.
specific organisational arrangements (e.g. one nurse present in a ward can be available in case of emergency in another ward).

Furthermore, varying case mix and facility type function may also account for observed variability in staff: nearly all general hospital psychiatric units accept compulsory admissions and provide psychiatric consultations to other hospital wards, whereas private facilities admit many patients not requiring intensive, acute care but rather needing long-term assistance and support; data from the second phase of our study, with some 3000 patients admitted to public and private facilities individually evaluated, will shed light on this issue. Official staffing requirements in Italy are currently established by regions and this issue will be given further attention.

Finally, specific professional roles (clinical psychologists, occupational therapists, etc.) are rather uncommon, suggesting the need to strengthen effective multidisciplinary teams.

Activity data
We found 19.8 admissions in public facilities per 10 000 inhabitants; if we include private admissions, we have a total rate of 26.7 psychiatric admissions per year per 10 000 inhabitants. The public admission rate is substantially lower than that found in England for the age group 16–64 years in 1999–2000 (32/10 000 inhabitants; if we include inpatient care but rather needing long-term assistance and support; data from the second phase of our study, with some 3000 patients admitted to public and private facilities individually evaluated, will shed light on this issue. Official staffing requirements in Italy are currently established by regions and this issue will be given further attention.

Involuntary admissions
Compulsory admissions represent 13% of all admissions per year in Italy, but the rate per 10 000 inhabitants is only 2.5. According to the recent percentage of compulsory admissions in relation to all admissions ranges from a low of 3.2% in Portugal to a high of 30% in Sweden, with a median value of 13.2% (Salize & Dressing, 2005). Yet, in the same European survey the rate per 10 000 inhabitants ranged from a low of 0.6 to a high of 21.8 with a median value of 7.4, which is three times the Italian rate.

In the survey of 24 different European sites by Rittmannsberger et al. (2004), 11.4% of in-patients had been admitted under compulsion. Latest UK data show a rate of about 5.5 compulsory admissions per 10 000 inhabitants for England (Department of Health, 2001). Analyses of temporal trends show substantial stability in the percentages of compulsory admissions in the past two decades (Guiana & Barbui, 2004).

Implications of the study
There is no evidence that a balanced system of mental healthcare can be provided without acute beds (Thornicroft & Tansella, 2004). Despite the importance, of acute in-patient care, its quantitative and qualitative features remain largely unexplored and many problems still await appropriate solutions (Quirk & Lelliott, 2003; Lelliott & Quirk, 2004) – and this holds true in Italy as elsewhere. Our findings highlight the need for a thorough revision of the role and the function of Italian acute care services within the framework of a comprehensive and evidence-based mental health policy.

ACKNOWLEDGEMENTS
This paper is dedicated to the memory of Ian Falcon, in gratitude for his outstanding contribution to the PROGRES projects since their beginning. The study was supported by a grant from the Ministry of Health. Special thanks are due to Dr Paul Lelliott, Dr Stephen Priebe and Dr Richard Warner for their valuable comments. Dr Giorgio Bignami provided continuous support throughout the National Mental Health Project, during which this study was conceived. Ms Elena Cuomo has provided valuable bibliographic help during the whole project.


REFERENCES


What is life like on acute psychiatric wards in the UK? A review of the research evidence.

Giovanni De Girolamo, Department of Mental Health, AUS di Bologna, Bologna and Regional Health Care Agency, Emilia-Romagna Region, Bologna; Angelo Barbato, Mario Negri Institute, Milan; Renata Bracco, Department of Mental Health, ASL Triestina, Trieste; Andrea Gaddini, Agenzia di Sanità Pubblica Regione Lazio, Rome; Rossella Miglio, Faculty of Statistics, University of Bologna, Bologna; Merilugi Morosini, National Institute of Health, Rome; Bruno Norcio, Department of Mental Health, ASL Triestina, Trieste; Angelo Picardi, National Institute of Health, Rome; Elisabetta Rossi, Department of Mental Health, ASL di Perugia, Perugia, Italy; Paola Rucci, Western Psychiatric Institute and Clinic, University of Pittsburgh, Pittsburgh, USA; Giovanni Santone, Psychiatric Clinic, United Hospitals of Ancona and Polytechnic University of Marche, Ancona; Giuseppe Dellaquila, Department of Mental Health, ASL Triestina, Trieste, Italy

Correspondence: Dr Giovanni de Girolamo, Department of Mental Health, AUSL di Bologna, Viale Pepoli 5, 40123 Bologna, Italy. Tel.: +39 051 658 4204; fax: +39 051 658 4244; email: giovanni.degirolamo@ausl.bologna.it

(First received 12 December 2005, final revision 19 December 2006, accepted 16 January 2007)
Manifest disease and motor cortex reactivity in twins discordant for schizophrenia

MARTIN SCHÜRMANNN, JUHA JÄRVELÄINEN, SARI AVIKAINEN, TYRONE D. CANNON, JOUKO LÖNNQVIST, MATTI HUTTUNEN and RIITTA HARI

Summary Schizophrenia is often associated with difficulties in distinguishing between actions of self and of others. This could reflect dysfunction of the mirror neuron system which directly matches observed and executed actions. We studied 11 people with schizophrenia and their co-twins without manifest disease, using stimulus-induced changes in the magnetoencephalographic ~20 Hz rhythm as an index of activation in the motor cortex part of the mirror neuron system. During action observation and execution, motor cortex reaction was weaker in those with schizophrenia than in their co-twins, suggesting a disease-related dysfunction of motor cognition.

Declaration of interest None.

METHOD

Participants were derived from a randomly selected subset of 335 schizophrenia-discordant twin pairs, identified (for another study) in a cohort of all 9562 pairs of same-sex twins born in Finland between 1940 and 1957 (Cannon et al, 2000). The Structured Clinical Interview for DSM-III-R Axis I Disorders (patient or non-patient edition; Spitzer et al, 1989) served for verification of the diagnoses in all participants. Interviewers were masked to zygosity and diagnostic status. A diagnosis of schizoaffective disorder, affective type in a twin with manifest disease or a psychotic disorder diagnosis in a non-affected twin led to exclusion of that twin pair (Cannon et al, 2000). Eleven twin pairs (aged 49–64 years, mean age 54.4 years, s.d. = 4.8, five monozygotic and six dizygotic) participated in the study after informed consent and ethics committee approval. All the participants with manifest disease were out-patients in stable clinical condition (further details in a data supplement to the online version of this report). For all pairs, zygosity was determined by DNA analysis (for details see Cannon et al, 2000). Neuromagnetic data were acquired during three experimental conditions: (a) rest – the participants rested in a relaxed state; (b) observation – the participants observed the experimenter manipulate a small object with her right-hand fingers; (c) action – the participants manipulated the small object with their right-hand fingers without seeing their own hand.

The left and right median nerves were stimulated alternately at the wrists (0.2 ms constant current pulses at intensities exceeding the motor threshold), once every 1.5 s. Signals from 204 planar gradiometers of a helmet-shaped whole-scalp neuromagnetometer (Vectorview, Neuromag, Helsinki, Finland) were analysed. Stimulus-related changes in the level of the ~20 Hz rhythm were quantified by first filtering signals through 14–30 Hz, then rectifying them and finally averaging them time-locked to the median nerve stimuli (approximately 100 signals averaged per condition). The strength of the rebound in each condition was then quantified (from the MEG channel with the strongest rebound suppression during action observation) as the mean level from 300 ms to 1300 ms after stimuli (Salmelin & Hari, 1994).

RESULTS

Figure 1a shows the ~20 Hz motor cortex level for one participant. The rebound, peaking at 700 ms, was abolished during object manipulation and significantly suppressed during observation, as shown previously (Schnitzler et al, 1997; Hari et al, 1998). Figure 1b and 1c illustrate the ~20 Hz reactivity in all twin pairs. For both hemispheres and for both observation and action conditions, the twins with schizophrenia showed weaker reactivity of the ~20 Hz rhythm than their non-affected co-twins (binomial test for n = 11 pairs: rest–action P = 0.033 and rest–observation NS in left hemisphere; rest–action P = 0.006 and rest–observation P = 0.006 in right hemisphere).

The rest levels of the ~20 Hz rhythm did not differ between affected and non-affected
co-twins, nor was there any statistically significant difference between the groups in the strengths of cortical responses peaking in the primary somatosensory cortex 20 ms and 35 ms after median nerve stimuli (t-test, \( P > 0.2 \)). The ~20 Hz reactivity and the dosages of antipsychotic medication were not correlated (Pearson’s \( r = 0.43, P = 0.19 \)) (further details in a data supplement to the online version of this report).

**DISCUSSION**

The ~20 Hz motor cortex rhythm in the twins with schizophrenia was systematically less reactive than in their non-affected co-twins, both during action observation and execution, with no sign of an additional mirror neuron system abnormality. Since the observed effects were not correlated with medication, we attribute them to the disease itself. The similar somatosensory cortical responses and the comparable resting levels of the rhythmic activity in non-affected and affected participants render implausible any general dysfunctioning of cortical responsiveness in the patient group. The weakened ~20 Hz reactivity, specific to clinically manifest disease in the affected twins, could be related to a deficit in motor cognition affecting both the command and the experience of action, both important for delusions of control (Frith, 2005). Further studies should test more extensively the functionality of motor and sensory mirroring in people with schizophrenia, focusing on subgroups displaying special abnormalities in the experience of action.

**ACKNOWLEDGEMENTS**

Supported by the Academy of Finland (National Centers of Excellence Programme 2006–2011), Sigrid Juselius Foundation, and the National Institute of Mental Health, USA (MH52857). We thank Ulla Mustonen for help in recruiting the participants.

**REFERENCES**


Correspondence

EDITED BY KIRIAKOS XENITIDIS and COLIN CAMPBELL

Contents  Genetic hypotheses for schizophrenia  Anticipation and the genetics of psychosis  Attention-deficit hyperactivity disorder and mood disorders in adults

Genetic hypotheses for schizophrenia

In their ‘common disease – rare alleles’ hypothesis McClellan et al (2007) come close to formulating an untenable theory. Although they refer to currently fashionable ‘candidate genes’ – e.g. dysbindin, neuregulin and DISC1 – it appears that they do not regard these as established. I agree that there is no consistency in the findings across even the largest genome scans conducted to date (sample sizes exceeding 300 and totally over 1000 sib-pairs; Crow, 2007) but disagree profoundly about the alternative.

We know that in schizophrenia: (a) incidences are more uniform across populations as one moves to the core syndrome, e.g. nuclear symptoms (Jablensky et al, 1992); (b) structural brain changes (e.g. ventricular enlargement) are consistent across populations (Chua et al, 2003) and uniform across patients relative to controls (Vita et al, 2000); (c) age at onset has a specific distribution throughout the reproductive phase of life; (d) there are gender differences (earlier onset and worse outcome in males); (e) the core syndrome comprises symptoms that are language related (i.e. specific to Homo sapiens). None of these findings would be expected if schizophrenia were a result of random mutations in a large number of genes such as McClellan et al postulate, nor would one expect variation in the form of illness within families as is generally observed.

While McClellan et al’s hypothesis promises a search for elusive rare alleles that will never reach a conclusion, Craddock et al (2007) perseverate in their claim that ‘Several genes have been implicated repeatedly as conferring risk for schizophrenia or bipolar disorder’. Comparison of the largest and most systematic linkage studies, including those of Craddock et al themselves, shows that these claims cannot be sustained (Crow, 2007).

Alternative hypotheses to the ‘rare alleles of major effect’ and the ‘polygenes of small effect’ deserve consideration. One such hypothesis (Crow, 2007) is that the variation arises in relation to characteristics that are specifically human, i.e. recent in evolution, and that it is ‘epigenetic’ in form, i.e. involves a modification of the sequence (methylation of DNA) or the associated chromosomal structure (methylation, phosphorylation or acetylation of histones) rather than a change in the DNA sequence itself. We do know that the risk for first-degree relatives is approximately 10%, whereas that for second-degree relatives is very little increased compared with the population as a whole. Although this is often held to be consistent with polygenic influence, it is also compatible with an ‘imprint’ that is applied and reapplied in meiosis (i.e. with a short time course between generations). The solution proposed is that the variation arises in relation to the change (speciation event) that defined the species, and that this is associated with the cerebral torque – the bias from right frontal to left occipital across the antero-posterior axis that is characteristic of the human brain. In contrast to McClellan et al’s rare alleles and Craddock et al’s polygenes of small effect, this hypothesis is specific and refutable – a gene has been identified that duplicated at 6 million years from the X to the Y chromosome to give rise to the ProtocadherinX/Y gene pair. This pair has been subject to accelerated evolution since the duplication event (Williams et al, 2006) and is in an unusual situation with respect to epigenetic modulation. This variation can be assessed and the hypothesis thereby tested.


T. Crow University Department of Psychiatry, Warneford Hospital, Oxford OX3 7JX, UK. Email: tim.crow@psych.ox.ac.uk doi: 10.1192/bjp.191.2.180

Authors’ reply: We are delighted that our article has stimulated discussion about strategies for gene discovery in schizophrenia. We agree that schizophrenia, like other complex traits, will be influenced by a large number of genetic and epigenetic events with a spectrum of effects. Both rare alleles of large effect and common alleles of modest effect are likely to be discovered (Craddock et al, 2007). Rare severe-effect alleles are fully compatible with familial patterns of schizophrenia because many (perhaps most) such alleles have arisen de novo in the present or recent generations. De novo mutations play havoc with predictions of conventional recurrence risk models. For example, de novo meiotic mutations (in the parental germline) increase disease concordance among monozygotic but not dizygotic twins. In contrast, de novo mitotic mutations or epigenetic events (in early embryogenesis) reduce concordance among both monozygotic and dizygotic twins.

Genetic association studies are not the most straightforward path to gene discovery for schizophrenia. Individually rare alleles cannot be identified by comparing frequencies of common alleles among unrelated patients with controls, even with enormous numbers of well-diagnosed patients, properly matched controls and
Very dense (and expensive) screening tools. To the extent that rare alleles are important to schizophrenia, study designs based on a naive 'common disease–common allele' model will yield variable and non-replicable results (King et al, 2006).

Characteristic patterns of age at onset, gender differences and brain changes associated with schizophrenia are fully compatible with causal influences of rare severe-effect events, either genetic or epigenetic. Each such event alters the expression, timing or function of one of a very large number of genes. The products of these genes converge in common pathways. Aberrations of a pathway by any of multiple mechanisms may lead to clinically similar disorders.

Crow's proposition that schizophrenia arises from the disruption of uniquely human genetic elements is very appealing. This premise, however, need not narrow the search for causes, genetic, epigenetic or environmental. Human speciation likely occurred primarily as a result of regulatory changes in genes, rather than common polymorphisms leading to changes in gene sequence (King & Wilson, 1975). The extraordinary number of repeated elements in the human genome gave rise to a vast number of new genes and regulatory mechanisms. Their architecture also created an increased risk for copying errors. Thus, one cost of the genomic complexity that enabled human brain development may be a de novo error rate that results in the maintenance of schizophrenia in the population.

Autism has recently been shown to be associated with a significantly increased frequency of rare de novo mutations (Sebat et al, 2007). These results presage the identification of many more rare mutations associated with other neurodevelopmental illnesses, as advances in technology enable the detection of ever-smaller genomic lesions. The ultimate resolution of this debate lies in gene discovery, for which we encourage the application of study designs most likely to be fruitful.


M. Blinc-Pesek University of Ljubljana, Slovenia
M. Agius Bedfordshire Centre for Mental Health Research in Association with the University of Cambridge, c/o Weller Wing, Bedford Hospital, UK.
Email: Mark.Agius@bgh.nhs.uk
doi: 10.1192/bjp.191.2.181

Attention-deficit hyperactivity disorder and mood disorders in adults
Asherson et al (2007) raise some important issues regarding adult attention-deficit hyperactivity disorder (ADHD). They state that some symptoms of bipolar disorder are similar to those of ADHD but the distinction is not difficult. However, although ADHD and classic euphoric mania (bipolar I) may be distinct, differentiation of ADHD and bipolar disorder may be difficult, especially in bipolar II, bipolar-spectrum disorder and episodes of mixed symptomatology. At times, it may be almost impossible to discriminate solely by symptoms. Irritability, excessive activity, impulsive behaviour, poor judgement and denial of problems are characteristic of both ADHD and bipolar disorder, thus making diagnosis difficult. The two also clearly occur together in some individuals: the reported overall lifetime prevalence of comorbid ADHD in people with bipolar disorder is 9.5% (Nierenberg et al, 2005); comorbidity with unipolar disorder is also frequent.

Asherson et al state that ADHD is a persistent trait whereas bipolar disorder is episodic. However, inter-episodic symptoms are common in bipolar disorder and the course of both bi- and unipolar disorder is frequently chronic; for example, up to 13% of people with bipolar disorder report continuous cycling without a well phase...
and 54% are not fully euthymic between episodes (Kupka et al., 2001).

Children of mothers with bipolar I disorder have increased rates of both unipolar disorder and ADHD, further suggesting a neurobiological overlap of these three diagnoses. Hirshfeld-Becker et al. (2006) report significantly higher rates (23.5%) of ADHD in offspring of parents with bipolar disorder compared with psychiatric comparison parents (8.4%) and non-psychiatric comparison parents (4.2%).

Drug treatments also overlap. Stimulant-type medication has been used in bipolar depression, and newer medications such as atomoxetine have similar pharmacological characteristics to some antidepressants (Lydon & El-Mallakh, 2006). Catecholaminergic antidepressants are not only potentially of benefit in ADHD but may be less likely to destabilise bipolar disorder.

There is thus a clinical and neurobiological overlap between ADHD, bipolar and unipolar disorder. Asherson et al.’s timely editorial has reminded us that ADHD in adults should not be overlooked and that further research is needed to clarify its impact on other adult psychopathology and comorbidity, particularly in mood disorders.


A. J. Kuan Department of Psychiatry, University of British Columbia, 2255 Wesbrook Mall, Vancouver BC, V6T 2A1, Canada. Email: anne.kuan@gmail.com

A. H. Young Department of Psychiatry, University of British Columbia, Canada
doi: 10.1192/bjp.191.2.181a

Authors’ reply: Kuan & Young point out that further research into the role of mood symptoms in attention-deficit hyperactivity disorder (ADHD) is essential. In a recent study of 141 adults with ADHD, 95% were found to have mood symptoms, chiefly mood instability (Kooij, 2007). We observe that in adult ADHD mood instability frequently responds to stimulants over the same time course as core ADHD symptoms, an observation reported by others. This has led to the suggestion that mood dysregulation might represent a core impairment in adult ADHD, perhaps related to the same processes that cause dysregulation of other executive processes.

Despite these observations the relationship of ADHD to mood disorders is controversial. The controversy has arisen in the context of paediatric bipolar disorder, where the distinction from ADHD is made difficult if one chooses to view irritability as a sufficient manifestation of bipolar disorder and if the requirement for episodicity is not strictly applied. However, available validation studies for the construct of paediatric bipolar disorder use elation and/or grandiosity as cardinal symptoms, rather than irritability. Narrowly defined paediatric bipolar disorder can be differentiated from ADHD, shows longitudinal stability and has plausible familial aggregation patterns (Geller & Tillman, 2005; Geller et al., 2006). Recent evidence suggests that the narrowly defined disorder can be distinguished at the behavioural and electrophysiological level from broadly construed disorder (Rich et al., 2007). Conversely, it has been argued that the intensity of irritability (Mick et al., 2005) and its temporal pattern (chronic or episodic) can distinguish paediatric bipolar disorder from ADHD (Leibenluft et al., 2006). The family study of Hirshfeld-Becker et al. (2006) is intriguing, yet the sample size is small (12 families with bipolar I disorder, 11 with bipolar II disorder), and further work is needed to clarify the rates of ADHD among relatives with narrowly defined v. broadly defined bipolar disorder.

One of the main questions to be addressed relates to how valid a diagnostic concept broadly defined bipolar disorder is, or whether mood instability/irritability in the presence of ADHD may be more adequately described by a new dimension, such as mood dysregulation (Brotman et al., 2006). Until the relevant empirical data become available, we see merit in maintaining the classic definition of mania, so that a diagnosis of bipolar disorder requires euphoria, grandiosity and episodicity, and the differential between ADHD and bipolar disorder remains explicit.


P. Asherson MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Denmark Hill, London SE5 8AF, UK. Email: P.Asherson@iop.kcl.ac.uk

A. Stringaris, W. Chen, E. Taylor Institute of Psychiatry, London, UK
doi: 10.1192/bjp.191.2.182
One hundred years ago

Etiology of General Paralysis [Contribution à l’étude de l’étiologie de la paralysie progressive]. (Cbl. fur Nerven. und Psych., March, 1906.) Mongeri, L.

The author bases his conclusions on cases occurring in his Constantinople practice. These cases are of exceptional interest from the etiological point of view in that (1) the patients comprise representatives of various nations; (2) these nations, owing to differences in religion and mode of life, vary greatly in their liability to syphilis and alcoholism.

After a preliminary description of the social and economic conditions prevailing in Constantinople the author applies the postulates thus gained to his 144 cases, and deduces the following conclusions: General paralysis is invariably preceded by syphilis. Instances where this cannot be proved may usually be accounted for either by the early exaltation of general paralysis, leading the patient to deny the existence of previous disease – the prevalence of pederastia, whereby infection may have taken place without leaving any discernible trace – or, finally, by the existence of hereditary syphilis. Syphilis alone is, however, incapable of producing general paralysis. Other accessory causes are requisite, of which by far the most important are heredity and alcoholism. A possible explanation of these facts is to be sought in the functional failure of the liver, involving loss of its poison-eliminating power. Out of thirty-eight cases examined from this point of view thirty-six had some organic or functional defect of the liver. It has been shown that toxins diminish the amount of glycogen; if the glycogen re-forms, recovery ensues; if not, the organism succumbs. On this theory the part played by alcoholism in the genesis of general paralysis is easily discernible. It is probable that heredity and intellectual overwork act in a similar manner.

General paralysis is, therefore, to be regarded as a result of an ensemble of causes. Its comparative limitation to civilized countries is explained by the fact that the necessary causes, though found singly, do not occur in combination amongst savage nations.

Bernard Hart

REFERENCES

Journal of Mental Science, January 1907, 181.
Researched by Henry Rollin, Emeritus Consultant Psychiatrist, Horton Hospital, Epsom, Surrey
doi: 10.1192/bjp.191.2.183

Corrigenda

X-chromosome abnormality and schizophrenia (letter). BJP, 190, 450. The author’s email should read: maarten.otter@filternet.nl. Authors’ reply: the fourth author’s name is R. S. Kahn. 
doi: 10.1192/bjp.191.2.183a
Book reviews

EDITED BY SIDNEY CROWN, FEMI OYEBODE and ROSALIND RAMSAY

**Journalists Under Fire: The Psychological Hazards of Covering War**


Chiding the military establishment for failing to provide adequate aftercare for wounded servicemen is a favourite media pastime and the psychological toll of conflict among service veterans has been popularised by the press. What is less well known – indeed, virtually invisible to the public gaze – is the psychological toll within the news media itself as journalists and photographers deliberately expose themselves to the risks and horrors of conflict to seek out the grotesque in pursuit of the big story and the best picture.

*Journalists Under Fire* is not a textbook of post-traumatic stress disorder or psychological trauma: rather it puts flesh on the bones of the sanitised, sterile descriptions of psychopathology in the academic literature. This book makes for uncomfortable reading: sometimes disturbing and upsetting but always compelling, Feinstein uses personal narrative to vividly and chillingly describe the psychological effects of war reporting on those journalists who bear witness to the brutality and inhumanity of conflict. The social consequences of trauma are starkly depicted: broken families, broken careers, broken lives; all too often sublimated and disguised by a fast-living, hard-drinking machismo lifestyle and culture that goes with the territory.

What is most disturbing, is not so much the incidence or nature of psychopathology among journalists, but the fact that so few of them get any sort of help or treatment. News organisations (who have a duty of care no less than the much-maligned military establishment) typically turn a blind eye and offer little in the way of support. Then there are freelancers who lack any of the benefits and protection that a concerned and responsible employer should provide.

Remarkably, this is the first published investigation looking specifically at journalists as a vulnerable group. It should give news organisations pause for thought and a stimulus to get their own house in order, before casting brickbats at the military. If they fail to act, and with no end in sight to the endless stream of war and terrorism flashed across our TV screens and news media, increasing numbers of naive young men and women will be put at risk without warning, preparation or aftercare merely to satisfy the insatiable public interest and voyeuristic appetite for war reportage.

**Forensic Psychiatry: Influences of Evil**


The Concise Oxford English Dictionary defines evil as ‘profound wickedness and depravity, especially when regarded as a supernatural force’ with the subsense of ‘something harmful or undesirable, e.g. social evils’. I give the definition because it is not in the book and ambiguity is a problem. Is that evil as in medieval, or merely undesirable? Are we dealing in hellfire and damnation, or suspension for breach of guidelines? The editors seem unconcerned with such distinctions, so dodgy business methods are thrown into the cauldron with homicide. Used in this way, as a generic term for things of which we disapprove, the concept of evil serves only to justify prejudice.

My interest in the topic began with removal from primary school after the head teachers’ ‘touch of evil’ lecture on the essential similarity between staying out late and armed robbery. Heady stuff for 9-year-olds, but it was a faith school. The head would have loved this book; it finds evil in pharmaceutical marketing and in the killing of children. It is no surprise that religion claims to identify evil wherever it resides, but it is disappointing that the inquisitorial method goes unchallenged in a book that deals also with science – or ‘science’ as the authors have it, with those quotation marks summing up their approach. Foucault dominates the references.

Foucault’s legacy is mixed. His big idea was the assault on professional power but, since attacks on doctors became a sport, he has lost the copyright. His other trademark is an impenetrable writing style. Several contributors perpetuate that legacy, without the excuse of writing in French. Parts of the book are incomprehensible or,
worse, capable of sinister interpretation. When sex offender programmes are contrasted with the cynical comments of the recipients, it looks more like old-fashioned special hospital cynicism (‘motivational interviewing’s too good for ‘em’) than radical critique. And surely therapy stands or falls by the outcome of clinical trials rather than salacious anecdotes? The gory stories titillate rather than illuminate, and it is profoundly depressing that one of the authors is a researcher nurse.

We should worry about the trend for forensic texts to include the e-word. It coincides with the growth of the fundamentalist right in the USA, and it reeks of punishment and stigma rather than treatment and rehabilitation. These are tough times for science and those who would discard it need something better than sociology and brimstone to put in its place.

Anthony Maden  The Paddock Centre, WLMHT, Crowthorne RG45 7EG, UK. Email: a.maden@imperial.ac.uk
doi: 10.1192/bjp.bp.106.029728

Risk and Resilience: Adaptations in Changing Times

This is a scholarly and thoughtful book in which Ingrid Schoon presents her work analysing longitudinal psychosocial and educational data from two UK birth cohorts, 1958 and 1970 through to the present. The introduction and the first two chapters outline the concepts and definitions that underpin the study. The work is about how children escape from disadvantage with a particular focus on academic attainment as a pathway for overcoming adversities. There is a particularly thorough discussion of the definitions of the concepts of risk and resilience and a highly readable review of the theoretical perspectives, taken over the past four decades, regarding the role of developmental influences on risk and resilience over the life course. Indeed, the first two chapters are so well done they should constitute essential reading for all behavioural scientists, mental health practitioners and those involved in forming health and education policy. Chapters three through seven constitute the empirical work charting persisting inequalities at times of marked socio-economic change in chapter three, through to personal goals and life plans by chapter seven. The reader is not assailed by numbers or statistics and each chapter gives a highly succinct summary of the findings. For example, the author concludes clearly from the data (p. 55) that ‘despite dramatic economic and social advances witnessed in the UK during the 2nd half of the 20th century, inequalities of opportunities and life chances have remained or have even become greater’. Similarly, when summarising the impact of cumulative risk effects on education outcomes, Schoon (p. 72) notes, ‘despite improving material conditions there are persistent or even increasing inequalities in academic attainment and adult psychological well-being’. It is striking in this chapter that there is much greater influence of parental social class on academic attainment in those born in the 1970 cohort compared with those born in 1958. Social contextual factors had an increasing influence on individual attainment as the 20th century progressed. There are other examples throughout the text pointing to a worrying growth in inequalities between socially disadvantaged and advantaged groups in our society as we get wealthier. The findings on protective factors are less compelling perhaps because of the increasing importance of measuring processes within the individual as a component of a multimethod approach to the impact of risk on individual attainment. Schoon acknowledges this and points to a need for future studies to combine measures of genes and physiology with psychosocial data, but also emphasises the importance of maintaining a clear scientific focus rather than tawling through data sets to see what can be found.

The final two chapters on conclusions, outlook and implications for interventions and social policy are as thoughtful and clear as the introduction, with a particularly helpful analysis focusing on the importance of available social and educational resources as well as delineating risks. Risk specificity is also clearly noted. For example, children can be resilient in the face of poverty via positive and emotionally supportive relations with parents. The latter is no protection, however, against underfunded and failing schools. The children of low-income families are less likely to survive a high-risk school system, even if they have adapted to their impoverished circumstances.

This is an excellent academic text that should be essential reference reading for mental health professionals. For researchers in the field of risk and resilience I would deem it fundamental. Psychiatrists-in-training should have access also as they would learn a great deal about the value of longitudinal, non-experimental studies that provide key information about the changing world we live in and the importance of the interplay between an individual and their environment over time.

Ian Goodyer  Section of Developmental Psychiatry, University of Cambridge. Douglas House, 18B Turpinston Road, Cambridge CB2 2AH, UK. Email: ig04@cam.ac.uk
doi: 10.1192/bjp.bp.106.031344
From the Editor’s desk

PETER TYRE

BELIEF, EVIDENCE AND THE BRITISH JOURNAL OF PSYCHIATRY

Every so often we publish a paper that almost everyone will read. The reasons why the review by Singh et al (pp. 99–105) on ethnicity and compulsory detention will attract interest is that it challenges us. The fact that Black and minority ethnic patients in 49 studies were disproportionately detained compared with White patients will surprise nobody; this statistic is a robust one that has been repeatedly found. However, they also noted that common race-based explanations ‘including that of racism within mental health services, were not supported by clear evidence’ (p. 99). At this point we cross into different territory. Can a systematic review address institutional racism, the nature and attitudes of religious faith (Leavé & King, pp. 97–98), or the effects of stress on soldiers who may have doubts on the cause for which they are fighting (Engelhardt et al, pp. 140–145)? No they cannot entirely, because all these issues involve crossing from an evidence framework to a belief one. The same evidence can be interpreted in entirely different ways by different belief proponents; the earnest belief for intelligent design has led to the creationists opposing Darwinism, and although in this conflict science says it can provide a conclusive riposte (Ayala, 2007), the alternative believers will go on believing.

There is a lot of evidence out there that needs interpreting. Black patients may be more often admitted inappropriately to psychiatric beds (Tyrer et al, 2006), and be detained nearly four times more often than White ones, but the ratio rises to > 8 in high-security psychiatric hospitals, where, for reasons that no one has explained, Black patients generally receive a diagnosis of mental illness and very rarely one of personality disorder (Leese et al, 2006). In debating this issue, and there were no fewer than four presentations on this subject at the recent Annual Meeting of the Royal College of Psychiatrists in Edinburgh, I hope that we can move the evidence and science forward. Singh et al (p. 103) found in their review that ‘few studies were hypothesis-driven or methodologically based on a testable theoretical or conceptual model’. Now that at least offers the chance of belief and evidence coming together; will those who are ready please come forward, hone your theories and test your designs. As Daniel Moynihan, a Democratic senator who had his own critics on the subject of racism, put it, ‘everyone is entitled to his own opinion, but not his own facts’.

THE BRITISH JOURNAL OF PSYCHIATRY ON SONG

Our journal loves its impact factor
That’s newly sprung in June
Our journal’s like the melody
That’s sweetly played in tune

I write this just after the June release of the impact factors of journals for 2006. Robert Burns would have scorned the use of his verse for something as mundane as the current preoccupation that learned journals have with impact factors, but this snippet nicely combines two otherwise unrelated topics. The impact factor of the British Journal of Psychiatry has now risen to 5.436, a figure more than twice its number of 12 years ago, a continuing process that reflects a great deal on the energy and commitment of past editors and many others in the publications department. This has again been achieved despite an increase of 10% in articles published. The other distinguishing feature is that the Archives of General Psychiatry (the top ranked) and the British Journal of Psychiatry are the only journals in the top 20 psychiatric journals to have a citing half-life (the median age of the articles cited in 2006) of 10 years or more. This illustrates that articles published in the Journal have both short- and long-term impact in terms of citation rates, and the ephemeral nature of some research, or what the Lancet (Anonymous, 1978) referred to some years ago as ‘elephant’s footprints in the mud’ (as they make a big impression at first and then disappear) does not apply to the British Journal of Psychiatry.

The second topic is Psychiatry in Music, and follows naturally from the late Sir Martin Roth’s aphorism that psychiatry is the most artistic of the sciences and the most scientific of the arts. We will be launching Psychiatry in Music in the Journal shortly as a companion to Psychiatry in Pictures and are looking for any examples of reflections of psychiatry in music, and whether they be erudite essays on the likely mental pathology of Robert Schumann (Slater & Meyer, 1959) or empirical observations on music therapy (Talwar et al, 2006), or an excursion into bebop (Wills, 2003) we will accommodate them all in some way. Please send them in (in the first instance, to bjp@cppsych.ac.uk). The place of the music piece is not yet determined but it will not replace the Editor’s Desk at the end of the Journal. I want to make up for the look of devastation on our school music teacher’s face when he asked me what I thought of the pieces that had been played at an afternoon’s concert recital. I named a violin sonata with enthusiasm. ‘Really’, he asked, giving me all his attention, ‘what was it you liked so much about it?’ ‘I knew that when it was finished we could all go home’, I replied. So, my apologies are half a century late and Psychiatry in Music will not be on the last page.