Psychiatry in pictures:

Psychiatry in pictures
ALLAN BEVERIDGE

Highlights of this issue:

Highlights of this issue
KIMBERLIE DEAN

EDITORIALS:

Regulatory policies on medicines for psychiatric disorders: is Europe on target?
CORRADO BARBUI and SILVIO GARATTINI

Mental health following terrorist attacks
MATTHEW G. WHALLEY and CHRIS R. BREWIN

REVIEW ARTICLES:

Psychological treatments for chronic post-traumatic stress disorder: Systematic review and meta-analysis
JONATHAN I. BISSON, ANKE EHLERS, ROSA MATTHEWS, STEPHEN PILLING, DAVID RICHARDS, and STUART TURNER

PAPERS:

Stability of early-phase primary psychotic disorders with concurrent substance use and substance-induced psychosis
CAROL L. M. CATON, DEBORAH S. HASIN, PATRICK E. SHROUT, ROBERT E. DRAKE, BOANERGES DOMINGUEZ, MICHAEL B. FIRST, SHARON SAMET, and BELLA SCHANZER
Lipids and essential fatty acids in patients presenting with self-harm
MALCOLM R. GARLAND, BRIAN HALLAHAN, MAIREAD McNAMARA, PHILIP A. CARNEY, HELEN GRIMES, JOSEPH R. HIBBELN, ANDREW HARKIN, and RONAN M. CONROY

Omega-3 fatty acid supplementation in patients with recurrent self-harm: Single-centre double-blind randomised controlled trial
BRIAN HALLAHAN, JOSEPH R. HIBBELN, JOHN M. DAVIS, and MALCOLM R. GARLAND

Treatment of social phobia: randomised trial of internet-delivered cognitive–behavioural therapy with telephone support
PER CARLBRING, MAGDALENA GUNNARSDÓTTIR, LINDA HEDENSJÖ, GERHARD ANDERSSON, LISA EKSELIUS, and TOMAS FURMARK

Risk of hip fracture in patients with a history of schizophrenia
LOUISE HOWARD, GRAHAM KIRKWOOD, and MORVEN LEESSE

Phenomenology of delirium: Assessment of 100 adult cases using standardised measures
DAVID J. MEAGHER, MARIA MORAN, BANGARU RAJU, DYMPNA GIBBONS, SINEAD DONNELLY, JEAN SAUNDERS, and PAULA T. TRZEPACZ

Reducing emotional distress in people caring for patients receiving specialist palliative care: Randomised trial
KIRI WALSH, LOUISE JONES, ADRIAN TOOKMAN, CHRISTINA MASON, JOANNE McLOUGHLIN, ROBERT BLIZARD, and MICHAEL KING

Best practice when service users do not consent to sharing information with carers: National multimethod study
MIKE SLADE, VANESSA PINFOLD, JOAN RAPAPORT, SOPHIE BELLRINGER, SUBE BANERJEE, ELIZABETH KUIPERS, and PETER HUXLEY

Cancer risk among parents and siblings of patients with schizophrenia
ITZHAK LEVAV, IRENA LIPSHITZ, ILYA NOVIKOV, INNA PUGACHOVA, ROBERT KOHN, MICHA BARCHANA, A. PONIZOVSKY, and HAIM WERNER

Neural correlates of the misattribution of speech in schizophrenia
PAUL ALLEN, EDSON AMARO, CYNTHIA H. Y. FU, STEVEN C. R. WILLIAMS, MICHAEL J. BRAMMER, LOUISE C. JOHNS, and PHILIP K. McGUIRE
SHORT REPORTS:

Police contact within 3 months of suicide and associated health service contact
KEITH R. LINSLEY, NEIL JOHNSON, and JESSICA MARTIN

Treatment adequacy for anxiety and depressive disorders in six European countries
A. FERNÁNDEZ, J. M. HARO, M. MARTINEZ-ALONSO, K. DEMYTTENAERE, T. S.
BRUGHA, J. AUTONELL, G. DE GIROLAMO, S. BERNERT, J. P. LÉPINE, and J. ALONSO

Localisation of increased prefrontal white matter in pathological liars
YALING YANG, ADRIAN Raine, KATHERINE L. NARR, TODD LENCZ, LORI
LaCASSE, PATRICK COLLETTI, and ARTHUR W. TOGA

Correspondence:

Mental health and incapacity legislation
A. Maden

Mental health and incapacity legislation
A. Buchanan

Authors’ reply:
J. Dawson and G. Szmukler

Violence and psychiatric morbidity
R. Persaud

Authors’ reply:

Assessment of manic symptoms in different cultures
M. Sanderson

Assessment of manic symptoms in different cultures
R. Reed

Author’s reply:
P. Mackin
One hundred years ago:

Henry Rollin

Book reviews:

Dementia – Mind, Meaning and the Person
Jane Garner

Hand book of Medicine in Psychiatry
Mark Tuthill and Irene Cormac

R. D. Laing: Contemporary Perspectives
Matthew Broome

Gambling as an Addictive Behaviour: Impaired Control, Harm Minimisation, Treatment and Prevention
Sanju George

The Psychiatry of Intellectual Disability
Colin Hemmings

Attachment from Infancy to Adulthood: The Major Longitudinal Studies
Katrina L. Lloyd

Prevention and Treatment of Suicidal Behaviour: From Science to Practice
Clare Nolan

Recovery Beyond Psychiatry
Glenn Roberts

Psychiatric Disorders and Pregnancy
Ian Jones
From the Editor's desk:

Peter Tyrer
Br J Psychiatry 2007 190: 188.
Highlights of this issue

BY KIMBERLIE DEAN

FATTY ACIDS AND SELF-HARM

Two papers this month examine the potential importance of fatty acids for those presenting with self-harm. Garland et al (pp. 112–117) found that patients with self-harm had low levels of both total cholesterol and essential fatty acids compared with a control group. They also report that depression and impulsivity scores were inversely correlated with essential fatty acid levels but that no differences or associations were found when platelet serotonergic measures were examined for a subgroup. Hallahan et al (pp. 118–122) conducted a double-blind randomised controlled trial of omega-3 fatty acid supplementation in a group of patients presenting with recurrent self-harm. After 12 weeks, those receiving supplementation had significantly greater improvements in depression, suicidality and daily stress scores. No impact was found on measures of impulsivity, aggression or hostility.

CANCER AND HIP FRACTURE RISK IN SCHIZOPHRENIA

Linking Israeli population registers for psychiatric disorders and cancer, Levav et al (pp. 156–161) found that both the biological parents and full siblings of individuals with schizophrenia had a reduced risk of developing cancer. This was particularly true of gender-concordant parent-offspring pairs. The authors conclude that their findings lend support to the hypothesis that a genetic factor may be responsible both for reducing cancer risk and disturbing neurodevelopment. Howard et al (pp. 129–134) report that those with schizophrenia are at increased risk of hip fracture as a result of treatment with prolactin-raising antipsychotics. The association was found to be particularly strong for men, in line with previous findings of lower bone density in men compared with women taking neuroleptics.

RANDOMISED TRIALS: INTERNET CBT AND CARER SUPPORT

Internet-based therapies for social phobia have been developed to increase access to treatment but compliance has been found to be a problem. Carlbring et al (pp. 123–128) conducted a randomised trial of internet-delivered cognitive-behavioural therapy for social phobia with the addition of weekly telephone support. Those in the treatment arm experienced a reduced level of symptoms and showed high treatment adherence, and, after 1 year, improvements were maintained. Those caring for relatives receiving palliative care are known to be at risk of psychological and physical ill health. Walsh et al (pp. 142–147) found that brief intervention by a carer advisor did not significantly reduce distress levels among such carers although qualitative benefits were reported. The authors postulate that the lack of significant benefit may be related to the brief duration of the intervention, the possibility that carers may already be well supported, or that the stress of caring for a dying relative may be particularly difficult to ameliorate.

DIAGNOSTIC STABILITY IN PSYCHOSIS AND THE PHENOMENOLOGY OF DELIRIUM

Over a 1-year follow-up period, Caton et al (pp. 105–111) found that 25% of those with a baseline DSM-IV diagnosis of substance-induced psychosis were subsequently diagnosed with a primary psychotic disorder. Compared with those whose diagnosis of substance-induced psychosis remained stable, these patients had poorer premorbid functioning, less insight and greater family mental illness. Using the Delirium Rating Scale and Cognitive Test for Delirium, Meagher et al (pp. 135–141) assessed 100 consecutive cases of delirium in a palliative care setting. The most frequent symptoms identified were sleep–wake cycle abnormalities and inattention; disorientation was found to be the least frequent cognitive deficit. Just less than half of the sample had psychotic symptoms, presenting with either perceptual disturbances or delusions but not both.

NEURAL BASIS OF AUDITORY HALLUCINATIONS

Allen et al (pp. 162–169) used functional magnetic resonance imaging to examine the brain regions involved in conscious speech appraisal. Those with a prior history of auditory verbal hallucinations (the hallucinator group) were more prone to misidentify their own speech, particularly when it had been distorted. The authors also found altered activation in the superior temporal gyrus and anterior cingulate in the hallucinatory group compared with either the control or non-hallucinator patient groups. They suggest that the occurrence of hallucinations may be related to problems in the conscious evaluation of speech origin.

SHARING INFORMATION: BEST PRACTICE

Clinicians face an ethical dilemma when a service user does not consent to the sharing of information with a carer. Slade et al (pp. 148–153) completed a synthesis of data obtained from policy review, national survey and individual interviews. The authors identified a number of key guiding principles and propose a best-practice framework. They highlight the importance of distinguishing between the sharing of general and personal information, and the importance of employing clinical judgement in balancing competing ethical pressures.

Pavel (a pseudonym) is a homeless asylum-seeker placed in Glasgow. He grew up in Ukraine and southern Russia and attended a special school for those who showed artistic promise. He had contact with the local psychiatric services in Ukraine and believed that he was subject to experimentation by them. Pavel agrees with his diagnosis of schizophrenia and finds his treatment, which includes antipsychotic medication, helpful. This picture makes reference to Carstairs State Hospital, a maximum security institution, where he was an in-patient for a period. Over the past 4 years Pavel has experienced two relapses of his illness, and arguably his paintings become less coherent and have less conventional subject matter during these periods. Pavel writes: ‘I am not interested in selling my painting. My painting is my own response to circumstances I have found myself in. I consider my paintings to be a scientific investigation into my own thinking and into my understanding of the world. For me, all philosophies and psychologies have a snowball effect: one of which I say is four dimensional. Under the old Theory of Time, these paintings I have created, for me, complete a body of work which can give others the gifts and treatments I have had’.
Regulatory policies on medicines for psychiatric disorders: is Europe on target?

CORRADO BARBUI and SILVIO GARATTINI

Summary The European Medicines Agency (EMEA) is the regulatory body that provides the institutions of the European Community with the best possible scientific advice on the quality, safety and efficacy of medicinal products. Drugs approved by the EMEA are automatically marketable in all the European member states. Since the beginning of the EMEA’s activities a number of drugs acting on the central nervous system obtained marketing authorisation. This editorial highlights some aspects of the EMEA rules that may negatively affect the evaluation of medicines for psychiatric disorders.

Declaration of interest S.G. was a member of the Committee for Proprietary Medicinal Products.

The recent revision of the European pharmaceutical legislation has given the European Medicines Agency (EMEA) new responsibilities. After more than 10 years of existence the EMEA has proved useful in ensuring member states shift towards harmonisation of pharmaceutical procedures and simplification of the process by which a central authorisation becomes valid in all the states (Garattini & Bertele’, 2001). Any opinion expressed by the EMEA on old or new products, relating to changes in therapeutic indications, approval, suspension or withdrawal of a product, has to be accepted by all members of the European Union. The system includes a centralised procedure, through the EMEA, and a decentralised procedure, whereby a new drug approved by one member state is accepted by the others after the procedure of mutual recognition. The recent revision of the European legislation (Regulation EC No. 726/2004 of the European Parliament and of the Council of 31 March 2004; Directive 2004/27/EC of the Parliament and of the Council, 31 March 2004) has extended the list of drugs that must go through the centralised procedure (Garattini et al, 2003).

Since its establishment the EMEA has issued recommendations, notes for guidance, conceptual papers and other official documents intended to guide the design and reporting of randomised controlled trials conducted for regulatory purposes. These official documents report the EMEA rules and criteria for approval of new drugs. So far, nine products acting on the central nervous system have been approved in line with these criteria, and in future years it is expected that the increasing responsibilities of the EMEA will progressively increase the number of products for psychiatric disorders submitted for approval (Garattini & Bertele’, 2003). In this still-evolving European scenario, at least three technical aspects of the EMEA rules may negatively affect the evaluation of medicines for psychiatric disorders.

PROCEDURES FOR DRUG APPROVAL

The centralised procedure is not compulsory for psychotropic drugs. In addition to the fact that the dual system of approval – centralised and decentralised – creates competition between the EMEA and the national drug agencies, with financial implications, it generates heterogeneity between countries in terms of approved indications (labels). Olanzapine, for example, has been positively assessed by the EMEA through the centralised procedure and released for marketing with the same label in all EU member states. However, a decentralised route has been followed in the case of quetiapine, marketed after 1995, and approved for the treatment of schizophrenia in the UK and for the treatment of ‘acute and chronic psychoses, including schizophrenia’, in Italy (Barbui et al, 2003). Labels have a key role in regulating the everyday prescribing and consumption of drugs: in Italy quetiapine is the only atypical antipsychotic that can be prescribed in patients without a diagnosis of schizophrenia or bipolar disorder. Off-label prescribing is not forbidden, but implies that doctors take full responsibility for the prescription and that patients give informed consent and pay the full price of the drug, as reimbursement is usually restricted to disorders stated in the label. Theoretically, approved labels should correspond to trial inclusion criteria, and it seems rather contradictory that European regulatory authorities, while strongly supporting the adoption of stringent inclusion criteria in clinical trials, with rigorous and restrictive reference to diagnostic rules, permit drugs to be licensed with generic and unspecified labels for use in clinical practice. Probably, only the abolition of the decentralised procedure will make the licensed indications of new drugs more consistent.

CONTROLLED TRIALS

At the EMEA new drugs can still be evaluated with no comparison with active alternative treatments. This means that new drugs can be proved effective and safe on their own, even though they might in fact be potentially less effective or less safe than other drugs currently in use. Although in situations where no (or only a few) active treatments are available this issue may not be relevant, in the field of psychotropic drugs, where many effective agents are available, this issue is crucial. Despite this, the demonstration of a difference against placebo, and not against an active comparator, makes a new psychotropic drug eligible for registration in Europe. If comparisons are made, the industry usually relies on demonstrating therapeutic ‘equivalence’ or ‘non-inferiority’, because this is in agreement with current EMEA requirements. This results in a high degree of uncertainty about the therapeutic role of new drugs. Even the recent revision of the European pharmaceutical legislation does not include the requirement that, when feasible, clinical studies should be conducted in comparison with reference drugs (in accordance with the Declaration of Helsinki) to establish the relative benefit of a new drug. In terms of public health needs, the concept of added value should be introduced into the legislation. This concept has two
positive consequences. First, it allows determination of whether a drug is active. If comparative trials show that a new drug is more effective than a standard one, it means that the new drug is active. Conversely, if a new drug is not more effective than a standard one, it means that the new drug is inactive or similarly active compared with the reference. In the latter scenario there is no added value. Second, the concept of added value would advance innovation in the development of drugs, because a higher threshold for the entry of new drugs would force investigators towards the development of innovative rather than ‘me too’ drugs. The current legislation, allowing investigators to demonstrate a difference against placebo, has encouraged the marketing of drugs with little degree of innovation. Investigators should be induced to design and conduct clinical trials aimed at discovering better activity, beneficial effects on different populations, and less or different toxicity.

Methodological considerations also should be taken into account. Recent data have shown that placebo-controlled trials, in comparison with active-controlled trials, tend to overemphasise the occurrence of hard outcomes, such as the rate of participants withdrawing from treatment (‘drop-outs’). In antipsychotic drug trials, for example, a systematic review showed that the proportion of participants discontinuing antipsychotics was substantially higher in placebo-controlled trials than in active-control clinical trials (Kemmler et al., 2005). In the field of psychotropic drugs, where withdrawal rates approaching or exceeding 50% are not uncommon, this may produce a problem of biased estimation of treatment effect, leading to erroneous conclusions and poor generalisability. Future revisions of the European pharmaceutical legislation should incorporate the requirement of active-control clinical trials in the evaluation of psychotropic drugs, at least in addition to placebo-controlled trials. Active-control clinical trials should be designed and powered to generate evidence of superiority (added value), providing physicians with clear indications on the therapeutic role of new medicines, with respect to older medicines already on the market.

**OUTCOMES**

A third aspect, particularly relevant to the evaluation of psychotropic drugs, is the choice of the outcome of interest. Whereas in other fields of medicine the definition of outcome measures may be a relatively straightforward task, in psychiatric disorders treatment efficacy may often be an elusive concept, typically quantified by means of rating scales. The EMEA guidance on this issue recognises that although improvement in symptoms should be documented as a difference between baseline and post-treatment score, in order to allow an estimate of clinical relevance the proportion of ‘responders’ or ‘remitters’ should be presented. Cut-off points should be defined a priori in the protocol. From a practical viewpoint this seems reasonable because it allows physicians to make judgements in terms of proportion of patients (and not means and standard deviations), absolute and relative risk differences and number needed to treat (Barbui et al., 2001). Unfortunately, this approach systematically magnifies the effect of new medicines against placebo. A situation was hypothesised of a 1-point difference in mean change in scores on the Hamilton Rating Scale for Depression between drug and placebo, and it was shown that by defining response as a minimum 12-point improvement on this scale a response rate of 50% in the drug condition and 32% in the placebo condition could be obtained (Moncrieff & Kirsch, 2005). A small difference in symptom score can thus be translated into a large and clinically relevant difference in proportions.

The EMEA rules should consider scores from rating scales and their categorisation as secondary outcome measures. Randomised controlled trials conducted for regulatory purposes should increasingly use, as primary outcomes, hard and practical measures such as suicide attempts, treatment switching, hospitalisation, school failure or truancy, job loss or even withdrawal from the trial itself (Tansella et al., 2006). The example provided by the Clinical Antipsychotic Trials of Intervention Effectiveness is paradigmatic in this regard (Lieberman et al., 2005). This study, which randomly assigned a total of 1493 patients with schizophrenia to receive olanzapine, perphenazine, quetiapine or risperidone for up to 18 months, employed as primary outcome the discontinuation of treatment for any cause. This discrete outcome was selected on the assumption that stopping or changing medication is a frequent occurrence and a major problem in the treatment of schizophrenia. The finding that 74% of patients discontinued the study medication within 18 months is a clear confirmation of the relevance of this outcome (Lieberman et al., 2005). A similar approach has been followed by the Bipolar Affective Disorder Lithium Anticonvulsant Evaluation trial, where hospital admission was defined as the primary outcome (Geddes & Goodwin, 2001). In these circumstances, the idea that hard outcome measures are suitable for practical and pragmatic clinical trials, but not for randomised controlled trials conducted for regulatory purposes, appears difficult to reconcile with the principles of evidence-based medicine.

**CONCLUSION**

In Europe, current policies on medicines for psychiatric disorders need to be further developed in order to fully comply with the EMEA mission statement of promoting ‘the protection of human health . . . and of consumers of medicinal products’ (Council of the European Communities, 1993).

**ACKNOWLEDGEMENTS**

The views presented in this editorial are those of the authors and should not be understood or quoted as being made on behalf of the EMEA and/or its scientific committees.

**REFERENCES**


Mental health following terrorist attacks

MATTHEW G. WHALLEY and CHRIS R. BREWIN

Summary  We review the current literature relating to mental health following terrorist attacks. Studies assessing symptoms of stress in the general population and those assessing the mental health of direct victims are considered. Use of mental health services following an attack is reviewed and recommendations are offered.

Declaration of interest  None.

Terrorist activity in the UK reached a peak during the 30 years from the late 1960s to the late 1990s, with hundreds of incidents associated with the political conflict in Northern Ireland. Although terrorist attacks have occurred all over the world, they have been particularly numerous in France, India, Iraq, Israel, Russia, Spain, Sri Lanka and the UK. More recently, Islamic terrorist groups have targeted London, Madrid and New York City, as well as numerous other cities in Europe, Asia and North America. Although the psychiatric impact of terrorist violence has been repeatedly noted, it is only comparatively recently that there has been systematic research into its effects on victims and on the wider community. This evidence now permits some estimate to be made of the mental health consequences of terrorism and of the challenge for psychiatric services.

Mental Health of the General Population

Studies conducted in representative samples of the general population following terrorist events can be divided into those that have investigated the prevalence of ‘substantial stress’ (the presence of a predetermined level of psychological symptoms) and those that have attempted to estimate the prevalence of diagnosable psychiatric disorders. Catchment areas studied vary from city districts, cities, and surrounding regions, to whole countries. Within the first month after the 11 September 2001 attacks, symptoms of stress were evident in individuals geographically far distant from the original incident, and nationally depressive symptoms in the USA rose for 4 weeks only to fall back to previous levels thereafter (Knudsen et al., 2003).

Schuster et al. (2001) demonstrated that the proportion of people experiencing substantial stress was negatively associated with distance from the attacks of 11 September. Therefore, for comparison purposes it is easiest to consider studies that have sampled from the city where the incident happened or from the surrounding area (details of these studies are presented as a data supplement to the online version of this editorial). Despite using a variety of different instruments, these studies reveal a close relationship between the time post-incident and the prevalence of ‘substantial stress’. Rates are extremely high in the first few days after the incident but are already in decline in the first 2 weeks and by 6–8 weeks have fallen by two-thirds. Thus for the majority of individuals significant stress symptoms are temporary and are unlikely to have lasting mental health implications (Vázquez et al., 2006). However, a significant minority will continue to have symptoms. Over 6 months after the 11 September attacks, 5.3% of New York City residents continued to meet criteria for sub-syndromal post-traumatic stress disorder (Galea et al., 2003), a condition associated with substantial functional impairment.

The data on rates of probable acute stress disorder or post-traumatic stress disorder (PTSD) show a similar pattern. Rates of PTSD in the general population attributable to single attacks may be as high as 11–13% in the first 6 weeks but decline sharply with time, with most studies indicating rates below 3% 2 months after the incident.

When considering these studies it is important to bear in mind that stress symptoms measured in the immediate aftermath of an attack are not necessarily ‘post-traumatic’. One of the hallmarks of PTSD is a sense of extreme threat that endures despite the danger having passed, something that can rarely be assumed about a terrorist attack. Thus, symptom reporting following the Madrid train bombings was particularly high in regular train passengers (Vázquez et al., 2006). Transient symptoms should in most cases therefore be regarded as a general and not necessarily inappropriate stress response, partly reflecting involvement and concern with one’s own safety as well as with the safety of the community, family and friends. Symptom reporting will also be associated with actual losses of people, possessions and employment (Galea et al., 2002; DeLisi et al., 2003).

It should also be remembered that community samples will contain a proportion of people who may have particular reasons to feel threatened by the events. The data suggest that those reporting more symptoms will include members of minority groups (Schuster et al., 2001; Galea et al., 2002; Rubin et al., 2005), people with previous experience of adversity (Galea et al., 2002, 2003) and people who have developed psychiatric disorders in response to past stresses (DeLisi et al., 2003). For these groups the typically intense levels of media coverage and general concern around terrorist attacks may increase general levels of stress by acting as a potent reminder of feared outcomes or of thematically similar experiences from the past. More research is needed into the long-term outcomes for these at-risk groups.

Children

There are fewer studies of children’s responses to terrorism. Henry et al. (2004) found no significant difference from parental reports in general levels of anxiety and depression in children from Chicago in the 100 days before and after the 11 September attacks. A number of studies have explicitly assessed children’s reactions to terrorist events. Close to 1 year after the bombing of the Alfred P. Murrah Federal Building in Oklahoma City, about 5% of elementary schoolchildren reported clinically significant levels of symptoms of PTSD (Gurwitch et al., 2002). A year later, almost 20% of middle schoolchildren living 100 miles from the city reported current bomb-related
symptoms that impaired their functioning at home or at school (Pfefferbaum et al., 2000).

Four days after 11 September 2001, 35% of a national sample of American parents reported that their child had at least one of five stress symptoms (Schuster et al., 2001). Six weeks later more than 60% of parents in the New York City metropolitan area reported that their child was upset (Schlenker et al., 2002) or had moderate post-traumatic stress reactions (Fairbrother et al., 2003). Without more normative data it is difficult to assess the significance of these reports. However, two studies have carried out diagnostic assessments in community samples of children after 11 September. One month later, 8% of Seattle children were estimated to have diagnosable levels of PTSD symptoms (Lengué et al., 2005). In New York City itself, 6 months later Hoven et al. (2005) reported that 28.6% of children had at least one probable anxiety/depressive disorder, the most common being agoraphobia (14.8%), separation anxiety (12.3%) and PTSD (10.6%).

**MENTAL HEALTH OF DIRECT VICTIMS**

Another group of studies have reported on the mental health of people with direct experience of a terrorist attack, typically using diagnostic interviews or screening tools designed to estimate the prevalence of disorder. In these studies PTSD appears to be the most common disorder attributable to the attack, followed by depression (North et al., 1999; Miguel-Tobal et al., 2005), although other sequelae include traumatic grief, panic, phobias, generalised anxiety disorder and substance misuse (prevalence rates of PTSD in direct victims of a number of major terrorist attacks are presented in the data supplement to the online version of this editorial). Direct victims of terrorist attacks are those most affected, usually by being physically present at the attack site or by having a close family member killed or injured. Despite wide variations in the number killed in the attacks and the timing of assessments, there is remarkable uniformity that within 2 years of the incident 30–40% of the people closest to the site of the attack are likely to develop a clinically diagnosable disorder. Few data are available for longer-term outcomes, but even 2½ years after the Paris attack the rate of PTSD among direct victims was 25%, and 2 years after the Pentagon attack on 11 September over 20% of employees who were present and responded to the survey were found to have PTSD. These figures emphasise that many reactions are intense and long-lasting and cannot be dismissed as normal, transient responses to traumatic events.

Studies of emergency workers have usually found considerably lower levels of psychopathology than in direct victims. Retrospective reports by body handlers describing their reactions at the time of the Oklahoma City bombing and 1 year later indicated negligible levels of PTSD and depression (Tucker et al., 2002), and North et al. (2002) reported a PTSD rate of 13% among firefighters 3 years after the bombing. Two months after the Madrid bombings Miguel-Tobal et al. (2005) found a rate of 1.2% for PTSD and 2% for depression among emergency personnel. Six months after the 11 September attack on the World Trade Center, 14.3% of those involved in the rescue effort in New York City had probable PTSD (Galea et al., 2003), but there appeared to be only a small excess of PTSD symptoms in handlers working in canine search and rescue teams who were deployed following 11 September, compared with non-deployed controls (Alvarez & Hunt, 2005).

Studies of emergency workers are hard to compare because response rates varied, and in all of them there was considerable scope for response biases to operate. Although it is not likely that these groups will respond with high levels of disorder, it is important to consider that actual levels of exposure to the attack site, and to scenes of severe injury and grotesque death, are likely to vary enormously, even among individuals attending the same incident. For example, in a study of firefighters who worked in the aftermath of the 1995 Oklahoma bombing, North et al. (2002) found that time working on the site and time spent in the ‘pit’, a particularly perilous area of the building, were associated with increased PTSD prevalence. The involvement of trauma and occupational health advisors may be of great importance in ensuring that organisations recognise the potentially toxic effects of high or prolonged levels of exposure and provide appropriate levels of protection and support.

**Children**

Again there have been fewer systematic studies of child victims. Elbedour et al. (1999) found that 50% of the daughters and 23.1% of the sons of those killed in the Hebron massacre were suffering from probable PTSD. Children were more likely to experience post-traumatic symptoms following the Oklahoma City bombing if they had been bereaved (Pfefferbaum et al., 1999). Other commentaries have drawn attention to a significant risk of psychological disorder in children who are direct victims, suffer bereavement or other losses, or have to witness repeated reminders of the attacks, including parental distress (Fairbrother et al., 2003; Hoven et al., 2005). Distress and disorder may manifest themselves in different ways depending on the child’s developmental stage, and it is likely that children’s distress is systematically underestimated by adults (Gurwitch et al., 2002; Koplewicz et al., 2004).

**RECEIPT OF MENTAL HEALTH SERVICES**

There is now a substantial evidence base indicating that PTSD can be successfully treated. Six months after the Oklahoma City bombing 69% of survivors had received some form of mental health intervention, although this might only have consisted of psychological debriefing (North et al., 1999). Similarly, 74% of survivors of the Paris bombings with PTSD received psychological treatment after the attack (Verger et al., 2004). In contrast, receipt of services in the wider population appears to be considerably lower. Three to six months after 11 September only about a quarter of those with the most severe PTSD symptoms in New York City were receiving counselling or mental health treatment (DeLisi et al., 2003). By 6–9 months after 11 September about a third of New York City residents with probable PTSD or depression had sought help from professionals, and these overwhelmingly consisted of people who had previously received mental health services (Stuber et al., 2006). Virtually nothing is known about the proportion of survivors of terrorist events with PTSD or other disorders who received appropriate, evidence-based treatment for their conditions, or how successful these interventions were at ameliorating their symptoms.

**CONCLUSIONS**

Terrorist attacks have widespread mental health effects, even on communities geographically distant from the attacks. In the
main these effects will be short-lived but there is a minority of individuals not directly involved in the incidents who will continue to experience clinical or subclinical levels of symptoms, often accompanied by functional impairment. Consistent with data on exposure and risk, 30–40% of people directly affected by terrorist action are likely to develop PTSD, and at least 20% are likely still to be experiencing symptoms 2 years later. Less is known about the mental health impact on children, but this too appears to be considerable (see the data supplement to the online version of this editorial). In contrast, there is less evidence that rescue workers and members of the emergency services are at high risk of developing disorder.

These findings have important implications for health services. Whereas some direct victims are likely to be in contact with providers of psychological services, in New York City following 11 September only around a quarter to a third of adults and children with significant post-traumatic stress symptoms received any treatment at all. Survivors with no previous contact with services were least likely to benefit from what was available. This was despite the strenuous efforts of those involved in Project Liberty, an unprecedented exercise involving over 100 mental health agencies delivering free public education and crisis counselling. If this level of unmet need is replicated elsewhere, it suggests that a targeted, active outreach programme will need to be a major feature of the response, for example using a ‘screen and treat’ approach. An important future task is to demonstrate that an outreach programme can be effective in identifying individuals with significant symptoms or functional impairment, in facilitating access to evidence-based treatment and in achieving the kind of positive health outcomes typically obtained in treatment trials for PTSD.

ACKNOWLEDGEMENTS

We thank Marylene Cloitre, Sandro Galea and Guinevere Tufnell for advice and comments.

REFERENCES


Psychological treatments for chronic post-traumatic stress disorder
Systematic review and meta-analysis

JONATHAN I. BISSON, ANKE EHLERS, ROSA MATTHEWS, STEPHEN PILLING, DAVID RICHARDS and STUART TURNER

Background The relative efficacy of different psychological treatments for chronic post-traumatic stress disorder (PTSD) is unclear.

Aims To determine the efficacy of specific psychological treatments for chronic PTSD.

Method In a systematic review of randomised controlled trials, eligible studies were assessed against methodological quality criteria and data were extracted and analysed.

Results Thirty-eight randomised controlled trials were included in the meta-analysis. Trauma-focused cognitive–behavioural therapy (TFCBT), eye movement desensitisation and reprocessing (EMDR), stress management and group cognitive–behavioural therapy improved PTSD symptoms more than waiting-list or usual care. There was inconclusive evidence regarding other therapies. There was no evidence of a difference in efficacy between TFCBT and EMDR but there was some evidence that TFCBT and EMDR were superior to stress management and other therapies, and that stress management was superior to other therapies.

Conclusions The first-line psychological treatment for PTSD should be trauma-focused (TFCBT or EMDR).

Declaration of interest None.

Chronic post-traumatic stress disorder (PTSD) is a common disorder that people may develop after exceptionally threatening and distressing events. Psychological treatments from various theoretical perspectives have been found to be effective for chronic PTSD in previous reviews (Van Etten & Taylor, 1988; Bradley et al, 2005). Some of the earlier reviews had to rely on uncontrolled trials as well as controlled ones, and on uncontrolled effect sizes. There are now sufficient numbers of randomised controlled trials of psychological treatments of chronic PTSD to allow a meta-analysis of effect sizes in such trials. We present a comprehensive systematic review and meta-analysis of randomised controlled trials assessing the efficacy of psychological treatments in reducing symptoms of chronic PTSD, and comparing the efficacy of different types of psychological treatment in reducing symptoms of this disorder.

METHOD

This review and meta-analysis derive from work undertaken in the preparation of PTSD treatment guidelines for the National Institute for Health and Clinical Excellence (NICE) in the UK (National Collaborating Centre for Mental Health, 2005). Further details of the protocol are published within the full guideline.

A systematic bibliographic search was undertaken to find randomised controlled trials of psychological treatments for PTSD from databases (EMBASE, Medline, PsycINFO and CINAHL) and the Cochrane Library, with each database being searched from inception to August 2004. Additional papers were found by hand-searching the references of retrieved articles, previous systematic reviews and meta-analyses of psychological treatments for PTSD. The search was restricted to papers with English-language abstracts. In addition, data from unpublished studies or papers in press were sought by contacting experts within the field.

Selection Studies were only considered if PTSD symptoms were the main target of treatment, all participants had had PTSD symptoms for at least 3 months following a traumatic event, at least 70% of participants had a diagnosis of PTSD, and PTSD symptoms were measured using a recognised scale. To be included studies had to be of randomised controlled design, with adult (>16 years old) participants; the studies had to report at least pre-treatment and post-treatment measures, and retain at least 50% of the original sample at the post-treatment assessment. There was no restriction regarding type of traumatic event. The minimum duration of symptoms was 1 month. Early intervention trials that only included participants with recent onset of PTSD were not included and are considered in a separate review (further details available from the author upon request). The searching and selection were done by a team of systematic reviewers led by R.M. Any disagreements with regard to inclusion or exclusion of a study were resolved by discussion with the other authors.

Validity assessment All published and unpublished papers were assessed against the following quality criteria: random sequence generation, concealment of allocation, masked assessment of outcomes, number of withdrawals, tolerability, adequate reporting of data and intention-to-treat analysis.

Data abstraction Study details including the nature of the traumatic events, participants’ characteristics and type of intervention were entered into a Microsoft Access database (version 2000), the quality criteria were applied and outcome data for included studies were entered into Review Manager version 4.2.3 for Windows. The application of quality criteria and the accuracy of outcome data were double-checked by a second reviewer.

Study characteristics An initial narrative synthesis was undertaken to describe the scope (participants, settings, intervention type, comparators, measures of effect), quality and outcomes of the studies. Three main efficacy outcomes were considered: one dichotomous outcome (retaining a diagnosis of PTSD) and two continuous outcomes (assessor-rated and self-reported severity of PTSD symptoms). Among the main
outcomes, the primary outcome was clinician-rated severity of PTSD symptoms, although this was not present for all studies.

Quantitative data synthesis
Where possible, meta-analysis was used to synthesise data, including additional meta-analyses for anxiety and depression measures where available, and numbers leaving the study early, using Review Manager. Post-treatment data (or change scores if reported instead of post-treatment data) for the psychological treatment and control condition were entered in the Review Manager tables. Dichotomous outcomes (PTSD diagnosis and leaving the study early for any reason) were analysed as a relative risk number and were calculated on an intention-to-treat basis (i.e., a ‘once randomised always analyse’ basis). This makes the conservative assumption that all participants who ceased to engage in the study had an unfavourable outcome, e.g., they left because the treatment was not acceptable and still had a diagnosis of PTSD. Continuous outcomes were analysed as standardised mean differences (SMDs) to allow for ease of comparison across studies. It was not possible to obtain intention-to-treat data for most of the trials, and we therefore used complete data for all continuous outcomes.

For consistency of presentation all data were entered into Review Manager in such a way that negative effect sizes or relative risk numbers less than 1 represented an effect that favoured the active treatment compared with the waiting-list control. Data were pooled from more than one study using a fixed-effects meta-analysis except where heterogeneity was present, in which case a random-effects model was used as described below.

Heterogeneity
To check for heterogeneity between studies, both the I²-test of heterogeneity and the χ²-test of heterogeneity (P < 0.10) as well as visual inspection of the forest plots were used. The I² statistic describes the proportion of total variation in study estimates that is due to heterogeneity (Higgins & Thompson, 2002). An I² of less than 30% was taken to indicate mild heterogeneity and a fixed-effects model was used to synthesise the results. An I² of more than 50% was taken as notable heterogeneity; in this case an attempt was made to explain the variation. If studies with heterogeneous results were found to be comparable, a random-effects model was used to summarise the results (DerSimonian & Laird, 1986). In the random-effects analysis, heterogeneity is accounted for in both the width of confidence intervals and in the estimate of the treatment effect. With decreasing heterogeneity the random-effects approach moves asymptotically towards a fixed-effects model. An I² of 30–50% was taken to indicate moderate heterogeneity. In this case, both the χ²-test of heterogeneity and a visual inspection of the forest plot were used to decide between a fixed- and random-effects model.

In order to explore heterogeneity further, sensitivity analyses were performed to consider the influence of higher-quality methodology (this was done by considering studies that used masked assessment, and those that used an intention-to-treat analysis), studies that only included females and those that only included Vietnam veterans.

Clinical effectiveness
Where psychological interventions were compared against waiting-list control groups an effect size (SMD) of —0.8 or less (e.g., a larger negative number) was considered clinically meaningful for continuous variables (a ‘large’ effect size; Cohen, 1988) and for dichotomous outcomes a relative risk of 0.65 or less (or greater than 1.54) was considered clinically meaningful. Where two active treatments were compared lower thresholds were set with an SMD of —0.5 or +0.5 for continuous variables (a ‘medium’ effect size), and for dichotomous outcomes a relative risk of 0.80 or less or 1.25 or greater was considered clinically meaningful. These thresholds came from discussions in the NICE Guideline Development Group in advance of undertaking the meta-analyses and were based on clinical experience and thresholds used in the literature (Schnurr et al., 2003). In order to be considered clinically meaningful the value had to meet the threshold criterion and the 95% confidence interval had to be greater than the threshold. If the SMD and relative risk met the threshold criterion but the 95% CI included values in the non-clinically significant range, this was interpreted as limited evidence for an effect. Similarly, if the SMD or relative risk value was below the threshold, the 95% CIs were examined to determine whether the evidence was inconclusive (in case the 95% CI included numbers greater than the threshold) or whether it could be stated that there was evidence suggesting that an effect was unlikely (where the 95% CI was entirely outside the clinically meaningful range).

Psychological treatment categories
Five separate psychological treatment categories were defined (see Appendix). These came from discussions by the NICE Guideline Development Group in advance of undertaking the meta-analyses and were based on clinical experience and categories used in the literature (Foa et al., 2000).

RESULTS
Thirty-eight studies were included in the meta-analysis. Figure 1 shows the meta-analysis profile summarising trial flow.
<table>
<thead>
<tr>
<th>Comparison</th>
<th>Clinician-rated PTSD symptoms</th>
<th>PTSD diagnosis (intent-to-treat)</th>
<th>Self-rated PTSD symptoms</th>
<th>Anxiety</th>
<th>Depression</th>
<th>Withdrawal rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFCBT v. waiting list/usual care</td>
<td>T &gt; W</td>
<td>T &gt; W</td>
<td>(T &gt; W)</td>
<td>T &gt; W</td>
<td>(W &gt; T)</td>
<td></td>
</tr>
<tr>
<td>14 studies</td>
<td>n = 649</td>
<td>n = 763</td>
<td>n = 428</td>
<td>n = 415</td>
<td>n = 625</td>
<td>n = 861</td>
</tr>
<tr>
<td>SMD = - 1.40</td>
<td>RR = 0.44</td>
<td>SMD = - 1.70</td>
<td>SMD = - 0.99</td>
<td>SMD = - 1.26</td>
<td>RR = 1.42</td>
<td></td>
</tr>
<tr>
<td>(95% CI 0.35 to 0.57)</td>
<td>(95% CI - 2.17 to - 1.24)</td>
<td>(95% CI - 1.20 to - 0.78)</td>
<td>(95% CI - 1.69 to - 0.82)</td>
<td>(95% CI 1.05 to 1.94)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMDR v. waiting list/usual care</td>
<td>E &gt; W</td>
<td>(E &gt; W)</td>
<td>(E &gt; W)</td>
<td>E &gt; W</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>5 studies</td>
<td>n = 162</td>
<td>n = 217</td>
<td>n = 156</td>
<td>n = 156</td>
<td>n = 160</td>
<td>n = 216</td>
</tr>
<tr>
<td>SMD = - 1.51</td>
<td>RR = 0.49</td>
<td>SMD = - 1.13</td>
<td>SMD = - 1.20</td>
<td>SMD = - 1.48</td>
<td>RR = 1.21</td>
<td></td>
</tr>
<tr>
<td>(95% CI - 1.87 to - 1.15)</td>
<td>(95% CI 0.28 to 0.06)</td>
<td>(95% CI - 2.13 to - 0.13)</td>
<td>(95% CI - 1.54 to - 0.85)</td>
<td>(95% CI - 1.84 to - 1.12)</td>
<td>(95% CI 0.66 to 2.22)</td>
<td></td>
</tr>
<tr>
<td>Stress management v. waiting list/usual care</td>
<td>(S &gt; W)</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>3 studies</td>
<td>n = 86</td>
<td>n = 121</td>
<td>n = 92</td>
<td>n = 109</td>
<td>n = 121</td>
<td></td>
</tr>
<tr>
<td>SMD = - 1.14</td>
<td>RR = 0.64</td>
<td>SMD = - 0.33</td>
<td>SMD = - 0.77</td>
<td>SMD = - 0.73</td>
<td>RR = 2.19</td>
<td></td>
</tr>
<tr>
<td>(95% CI - 1.62 to - 0.67)</td>
<td>(95% CI 0.47 to 0.07)</td>
<td>(95% CI - 0.47 to 1.14)</td>
<td>(95% CI - 0.23 to - 0.31)</td>
<td>(95% CI - 0.12 to - 0.33)</td>
<td>(95% CI 0.71 to 6.73)</td>
<td></td>
</tr>
<tr>
<td>Other therapies v. waiting list/usual care</td>
<td>?</td>
<td>?</td>
<td>(O &gt; W)</td>
<td>(O &gt; W)</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>2 studies</td>
<td>n = 72</td>
<td>n = 166</td>
<td>n = 166</td>
<td>n = 166</td>
<td>n = 166</td>
<td></td>
</tr>
<tr>
<td>SMD = - 0.43</td>
<td>RR = 0.79</td>
<td>SMD = - 0.61</td>
<td>SMD = - 0.48</td>
<td>SMD = - 0.25</td>
<td>RR = 3.82</td>
<td></td>
</tr>
<tr>
<td>(95% CI - 0.9 to 0.04)</td>
<td>(95% CI 0.53 to 1.18)</td>
<td>(95% CI - 0.98 to - 0.24)</td>
<td>(95% CI - 0.82 to - 0.14)</td>
<td>(95% CI 0.71 to 0.22)</td>
<td>(95% CI 1.19 to 12.29)</td>
<td></td>
</tr>
<tr>
<td>Group CBT v. waiting list/usual care</td>
<td>?</td>
<td>(GC &gt; W)</td>
<td>?</td>
<td>No data</td>
<td>No data</td>
<td>?</td>
</tr>
<tr>
<td>1 study</td>
<td>n = 97</td>
<td>n = 48</td>
<td>n = 57</td>
<td>n = 57</td>
<td>n = 271</td>
<td></td>
</tr>
<tr>
<td>SMD = - 0.72</td>
<td>RR = 0.56</td>
<td>SMD = - 0.71</td>
<td>SMD = - 0.25</td>
<td>SMD = - 0.25</td>
<td>RR = 1.00</td>
<td></td>
</tr>
<tr>
<td>(95% CI - 1.14 to - 0.31)</td>
<td>(95% CI 0.31 to 1.01)</td>
<td>(95% CI - 1.2 to - 0.22)</td>
<td>(95% CI - 0.71 to 0.22)</td>
<td>(95% CI 0.64 to 1.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMDR v. TFCBT</td>
<td>?</td>
<td>(E = T)</td>
<td>(E = T)</td>
<td>?</td>
<td>(E = T)</td>
<td></td>
</tr>
<tr>
<td>6 studies</td>
<td>n = 187</td>
<td>n = 267</td>
<td>n = 206</td>
<td>n = 186</td>
<td>n = 206</td>
<td>n = 287</td>
</tr>
<tr>
<td>SMD = 0.02</td>
<td>RR = 1.14</td>
<td>SMD = - 0.17</td>
<td>SMD = - 0.14</td>
<td>SMD = - 0.32</td>
<td>RR = 0.87</td>
<td></td>
</tr>
<tr>
<td>(95% CI - 0.5 to 0.55)</td>
<td>(95% CI 0.70 to 1.85)</td>
<td>(95% CI 0.45 to 0.11)</td>
<td>(95% CI 0.48 to 0.20)</td>
<td>(95% CI 0.9 to 0.26)</td>
<td>(95% CI 0.58 to 1.30)</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Table 1 (Continued)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>Clinician-rated PTSD symptoms (intent-to-treat)</th>
<th>PTSD diagnosis</th>
<th>Self-rated PTSD symptoms</th>
<th>Anxiety</th>
<th>Depression</th>
<th>Withdrawal rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFCBT v. stress management</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>6 studies</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>SMD = -0.27</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>(95% CI = -0.71 to 0.16)</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>TFCBT v. other therapies</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>3 studies</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>SMD = -0.81</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>(95% CI = -1.19 to -0.42)</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>EMDR v. stress management</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>3 studies</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>SMD = -0.35</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>(95% CI = -0.90 to 0.19)</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>EMDR v. other therapies</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>No data</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>Stress management v. other therapies</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>1 study</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>SMD = -1.22</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>(95% CI = -2.09 to -0.35)</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>Group TFCBT v. group CBT (non-trauma-focused)</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>1 study</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>SMD = -0.12</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
<tr>
<td>(95% CI = -0.34 to 0.1)</td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
<td><img src="https://via.placeholder.com/150" alt="Image" /></td>
</tr>
</tbody>
</table>

1. Key to comparison: X > Y, evidence that X has clinically important advantages over Y; X = Y, limited evidence that X has clinically important advantages over Y; ?, evidence is inconclusive so it is not possible to determine whether there is a clinically important difference; X < Y, there is evidence suggesting that there is unlikely to be a clinically important difference.

CBT, cognitive–behavioural therapy; EMDR, eye movement desensitisation and reprocessing; GC, group CBT, non-trauma-focused; GT, group TFCBT; O, other therapies; PTSD, post-traumatic stress disorder; RR, relative risk; S, stress management; SMD, standardised mean difference; TFCBT, trauma-focused CBT; W, waiting list/usual care.
Study characteristics

Two additional randomised controlled trials met inclusion criteria but differed in mode of delivery (Lange et al. 2003; Neuner et al. 2004), and one further trial compared two versions of TFCBT (exposure and cognitive therapy) with each other (Tarrier et al., 1999a,b). These studies could not be included in the meta-analysis.

Quantitative data synthesis
Table 1 provides details of the quantitative data synthesis. It highlights that TFCBT and EMDR were better than waiting-list/control on most outcome measures. Stress management was better on some outcomes, and ‘other therapies’ appeared to be the least effective. Unfortunately none of the studies reported adverse effects and therefore it was not possible to analyse these. However, most studies did report withdrawal rates and these are included in Table 1.

Sensitivity analyses
Masked assessment
The EMDR studies using masked assessment showed evidence favouring EMDR over waiting-list on reducing the severity of PTSD symptoms (clinician-rated measures) (three studies, n=120; SMD=−1.54, 1.54, 95% CI −1.95 to −1.12) similar to that in all EMDR studies (see Table 1). The TFCBT studies using masked assessment showed evidence favouring TFCBT over waiting-list on reducing the severity of PTSD symptoms (clinician-rated measures) (seven studies, n=308; SMD=−1.70, 95% CI −2.47 to −0.93) similar to that in all TFCBT studies.

Vietnam veteran studies
One EMDR study considered only Vietnam veterans. This showed less evidence favouring EMDR over waiting-list on reducing the severity of PTSD symptoms (clinician-rated measures) (one study, n=25; SMD=−0.97, 95% CI −1.81 to −0.13) than the other EMDR studies (see Table 1). One TFCBT study considered only Vietnam veterans using the primary outcome measure; this showed less evidence favouring TFCBT over waiting-list on reducing the severity of PTSD symptoms (clinician-rated measures) (one study, n=24; SMD=−0.22, 95% CI −1.03 to 0.58) than the other TFCBT studies.

Female studies
The EMDR studies including only female participants showed evidence favouring EMDR over waiting-list on reducing the severity of PTSD symptoms (clinician-rated measures) (two studies, n=57; SMD=−1.67, 95% CI −2.30 to −1.04) similar to that in all EMDR studies. The TFCBT studies including only female participants showed more evidence favouring TFCBT over waiting-list on reducing the severity of PTSD symptoms (clinician-rated measures) (six studies, n=358; SMD=−2.06, 95% CI −2.70 to −1.42) than all TFCBT studies.

Intention-to-treat analysis
None of the EMDR studies reported using an intention-to-treat analysis so this could not be assessed. The TFCBT studies using an intention-to-treat analysis showed more evidence favouring TFCBT over waiting-list on reducing the severity of PTSD symptoms (clinician-rated measures) (six studies, n=332; SMD=−1.82, 95% CI −2.76 to −0.89) than all TFCBT studies.

DISCUSSION
We identified 38 randomised controlled trials of psychological treatments for PTSD. Trauma-focused cognitive–behavioural therapy showed clinically important benefits over waiting-list or usual care on all measures of PTSD symptoms. In addition, there was limited evidence that it also has clinically important effects on depression and anxiety. The effectiveness of eye movement desensitisation and reprocessing was also generally supported by the meta-analysis, but the evidence base was not as strong as that for TFCBT, both in terms of the number of trials available and the certainty with which clinical benefit was established. Furthermore, there was limited evidence that TFCBT and EMDR were superior to supportive/non-directive treatments, hence it is highly unlikely that their effectiveness is due to non-specific factors such as attention. There was limited evidence for stress management and group cognitive–behavioural therapy, but ‘other therapy’ (supportive/non-directive therapy, psychodynamic therapies and hypnotherapy) that focused on current or past aspects of the patient’s life other than the trauma or on general support did not show clinically important effects on PTSD symptoms, depression or anxiety. However, this might be due to the limited number of studies available and does not mean that these treatments were shown to be ineffective.

The treatments most supported by the review (individually delivered TFCBT and EMDR) are both trauma-focused psychological treatments that specifically address the patient’s troubling memories of the traumatic event and the personal meanings of the event and its consequences. Direct comparisons of these two approaches did not reveal any significant advantages of
one over the other, with respect to either treatment outcome or speed of therapeutic change (Taylor et al, 2003).

Heterogeneity

There is clearly considerable clinical diversity within the studies considered. The separation of different active interventions into groups partially addresses their impact on clinical diversity, but not all trials within the same group used identical interventions. The differences were most marked in the ‘other therapy’ group, which had in common the absence of cognitive-behavioural techniques and trauma-focused work. There was also diversity in the TFCBT group, which included both exposure-only and trauma-focused cognitive therapy interventions.

Another source of heterogeneity was the quality of the studies. Sensitivity analyses of higher-quality and lower-quality studies were performed to explore this further. There was some limited evidence that higher-quality studies (those including masked assessment of outcome or intention-to-treat analysis) showed better outcomes than the lower-quality studies. This finding contradicts previous research (Moher et al, 1998) that has found an association between poorer methodology and more favourable results for the intervention. It may reflect the fact that the better studies tended to be more recent and associated with refinement of techniques. They also included most of the female-only studies. The fact that female-only studies showed a better response to TFCBT than mixed studies and male-only studies is difficult to interpret. It may be that the female-only studies used more effective interventions, that the trauma of rape is more amenable than other traumas to effective TFCBT, or that for some undetermined reason women are more responsive to TFCBT than men. Interestingly, a similar superiority in female response has been found for pharmacological treatment of PTSD (National Collaborating Centre for Mental Health, 2005). The finding that studies including only Vietnam veterans produced worse responses to TFCBT and EMDR might have contributed to the female studies finding and also suggests that Vietnam veterans are a particularly difficult population to treat.

As with all psychological treatment trials, there are issues with the control group. The development of a psychological treatment placebo is difficult, if not impossible, as is masking of participants and therapists. In several of the waiting-list or usual care conditions it was apparent that some (usually poorly defined) treatment was going on. The main effect of this is likely to have made it more difficult for the active intervention to show itself to be superior to the control condition.

Tolerability

Unfortunately none of the studies reported adverse effects. It remains unclear whether no adverse effects occurred, or whether they were not described. This is a key shortcoming in the trials identified. Most studies reported withdrawals by group. There are likely to be several different factors that determine withdrawal rates, including the tolerability of the intervention. There was limited evidence that TFCBT and other therapies fared worse than waiting-list or usual care on this outcome measure, but there was no significant difference in withdrawal rates in direct comparisons between any of the active treatments. The higher-quality TFCBT studies showed no difference in withdrawal rates when compared with waiting-list or usual care. Some people find it difficult to fully engage in psychological treatment because it requires a significant commitment of time and emotion. For some people with PTSD it may initially be difficult and overwhelming to disclose details of their traumatic events. It is also well recognized that some patients may be subject to initial adverse effects such as increased re-experiencing following exposure treatment (Pirman et al, 1991; Foa et al, 2002; Hackmann et al, 2004). Withdrawal rates of up to 30% in some studies suggest that the active treatments were not always acceptable to those receiving them. It is possible that in these cases devoting several sessions to establishing a trusting therapeutic relationship and emotional stabilisation, before addressing the traumatic event, might lead to greater acceptability.

Clinical implications

Our results suggest that trauma-focused psychological treatments (TFCBT and EMDR) are effective for chronic PTSD. Indeed, the effect sizes compare favourably with those found for cognitive–behavioural therapy in depressive and anxiety disorders (National Collaborating Centre for Mental Health, 2004; National Collaborating Centre for Primary Care, 2004). These treatments are normally delivered on an individual out-patient basis over 8–12 sessions. A course of trauma-focused psychological treatment should be offered to everyone with chronic PTSD. The results also suggest that not all chronic PTSD will benefit from these treatments; other approaches should then be considered, including extending the number of sessions, trying an alternative form of trauma-focused psychological treatment and the augmentation of trauma-focused psychological treatment with a course of pharmacological treatment. A recent meta-analysis has suggested that pharmacological interventions are unlikely to be as clinically effective as trauma-focused psychological interventions and should therefore be used as a second-line treatment (National Collaborating Centre for Mental Health, 2005).

Future research

Further well-designed trials of psychological treatments are required, including further comparison studies of one type of psychological treatment against another. There is a need for large-scale studies (phase 4) to find out whether the results will survive in real practice. Future trials should consider adverse events and tolerability of treatment in more detail. Our results suggest that several of the currently available treatments might benefit from modifications that would make them more acceptable to people with chronic PTSD.
A controlled evaluation of cognitive behavioural posttraumatic stress disorder.

With cognitive restructuring in treatment of PTSD.

Clinical Psychology

Multidimensional meta-analysis of psychotherapy for PTSD.


*Power, K., McGoldrick, T., Brown, K., et al. (2002) A controlled comparison of eye movement desensitization and reprocessing versus exposure plus cognitive restructuring versus waiting list in the treatment of...


*Studies that were part of the meta-analysis.
Stability of early-phase primary psychotic disorders with concurrent substance use and substance-induced psychosis

CAROL L. M. CATON, DEBORAH S. HASIN, PATRICK E. SHROUT, ROBERT E. DRAKE, BHAOANERGES DOMINGUEZ, MICHAEL B. FIRST, SHARON SAMET and BELLA SCHANZER

Background  The stability of the diagnostic distinction between a substance-induced psychosis and a primary psychotic disorder co-occurring with substance use is not established.

Aims  To describe DSM–IV diagnostic changes over 1 year and determine the predictive validity of baseline indicators of the substance-induced psychosis v. primary psychosis distinction.

Method  We conducted a 1-year follow-up study of 319 psychiatric emergency department admissions with diagnoses of early-phase psychosis and substance use comorbidity.

Results  Of those with a baseline DSM–IV diagnosis of substance-induced psychosis, 25% had a diagnosis of primary psychosis at follow-up. These patients had poorer premorbid functioning, less insight into psychosis and greater family mental illness than patients with a stable diagnosis of substance-induced psychosis. Reclassifying change cases to primary psychoses on follow-up, key baseline predictors of the primary/substance-induced distinction at 1 year also included greater family history of mental illness in the primary psychosis group.

Conclusions  Further study of substance-induced psychoses should employ neuroscientific and behavioural approaches. Study findings can guide more accurate diagnoses at first treatment.

Declaration of interest  None. Funding detailed in Acknowledgements.

Comorbid substance use is frequently observed among patients presenting for treatment with symptoms of psychosis (Serper et al, 1999; Weaver et al, 2003; Arseneault et al, 2004; Green et al, 2005). Among patients presenting with a first episode of psychosis, lifetime comorbidity with substance use disorder has been observed in a third to nearly a half of admissions (Van Mastrigt et al, 2004; Barnes et al, 2006; Mauri et al, 2006). Diagnostic certainty in early-phase psychotic disorder is often difficult to achieve (Drake et al, 2003) and is challenged further when psychosis co-occurs with the use of alcohol or drugs (Greach et al, 2005). Diagnostic change over time has been observed in longitudinal studies of primary psychotic disorders (McGorry, 1994; Schwartz et al, 2000). Despite the clinical significance of a differential diagnosis between a primary and a substance-induced psychotic disorder, surprisingly little is known about longitudinal diagnostic stability and change in psychotic disorders co-occurring with alcohol or drug use. A change in diagnosis from a substance-induced psychosis to a primary psychosis can reflect the evolution of an illness, the availability of new information about onset or course, or unreliable diagnostic assessments (Schwartz et al, 2000). Psychomimetic drug use may precipitate a schizophrenia-like illness (Andreason et al, 1988; Boutros & Bowers, 1996; Zammit et al, 2002) or may evolve into a chronic psychotic disorder over time (McLellan et al, 1979). Yet systematic evidence for such a diagnostic shift is lacking. The distinction between a substance-induced psychosis and a primary psychotic disorder is important because these two disorders require fundamentally different approaches to treatment.

In the study reported here we used follow-up data from participants in an earlier study to address the stability of DSM–IV primary and substance-induced psychotic disorders; predictors of change in diagnosis during the follow-up; and the 1-year predictive validity of the key variables that distinguished the primary and substance-induced psychosis groups at baseline.

METHOD

Study aims  Our study consisted of a 1-year follow-up assessment of a sample of 386 patients with early-phase psychosis and substance use (Caton et al, 2005). We reported previously that among this patient group at baseline assessment (Caton et al, 2005), patients with substance-induced psychosis had greater personal and parental substance use disorders and more often experienced visual hallucinations, whereas patients with primary psychosis had greater overall psychopathology on the Positive and Negative Syndrome Scale (PANSS; Kay et al, 1992).

To study diagnostic stability and change over the first year of follow-up, we compared diagnostic assessments made at baseline with diagnostic assessments made at the 6-month and 12-month follow-up points. We focused on the primary distinction between psychosis and substance-induced psychosis. We observed substance-induced psychotic episodes in participants with baseline primary psychotic disorder whose diagnostic designation by definition remained stable. However, the main focus of our research was on cases with a change from a baseline diagnosis of substance-induced psychosis to a follow-up diagnosis of primary psychosis. To study the predictive validity of key variables distinguishing the two diagnostic groups at baseline, we used baseline assessments of demographic, family and clinical variables, and the follow-up diagnosis at 1 year.

Design and setting  Research methods in this longitudinal cohort study have been described in detail elsewhere (Caton et al, 2005). Briefly, participants were recruited from five psychiatric emergency departments in upper Manhattan.

Participants  The study sought to identify people experiencing psychosis in an early phase. We followed the precedent established in prior research on early psychosis (Schwartz et al, 2000) by excluding those whose first admission to hospital for psychosis occurred...
more than 6 months prior to the index admission. We did not include individuals who had experienced an extended duration of continuous psychotic symptoms in the absence of prior treatment, out of concern that psychotic symptoms might already be chronic. Participants were English- or Spanish-speaking, aged 17–45 years, had at least one psychotic symptom assessed during administration of the research protocol and had used alcohol or drugs within the preceding 30 days. All patients who met these criteria were eligible for the study, regardless of psychosis diagnosis.

Of the 386 participants meeting DSM-IV criteria for primary or substance-induced psychotic disorder at baseline, follow-up data were obtained on 319 (83%). Of the 67 who were not interviewed post-baseline, 31 were lost to follow-up, 16 left the region and could not be interviewed, 11 were incarcerated and could not be interviewed, 8 refused to continue their participation in the study, and 1 died. Compared with the interviewed group, those not interviewed had greater homelessness, more unemployment and poorer family support. There was no difference in gender, age, race, level of education, jail or prison history, or baseline diagnosis of primary or substance-induced psychosis. Characteristics of the interviewed group are shown in Table 1.

Data collection
Participants were initially interviewed at baseline after voluntary informed consent was obtained. They were contacted monthly to obtain information on clinical status and service use, and were re-interviewed in depth at 6 months and 12 months. Follow-up interviews were typically conducted in the community by trained assessors with master’s degrees in psychology or social work. The research protocol was approved by the institutional review boards of the New York State Psychiatric Institute/Columbia University Medical Center and the other institutions from which participants were recruited.

Assessments
Research diagnostic assessments at baseline and follow-up
Research diagnoses were made using the Psychiatric Research Interview for Substance and Mental Disorders (PRISM; Hasin et al, 1996, 2006), which was developed to assess psychiatric and substance use

comorbidity using DSM-IV criteria (American Psychiatric Association, 1994). A detailed description of the PRISM interview, DSM-IV guidelines to distinguish between a primary psychotic disorder and substance-induced psychosis, and the implementation of these guidelines is given by Caton et al (2005). Test–retest reliability for psychotic symptoms in the PRISM is good to excellent (κ = 0.63–0.76; Hasin et al, 1996), and the PRISM differentiation between primary and substance-induced psychotic disorders was good to excellent (κ = 0.75–0.86; Hasin et al, 2006). Validation of PRISM diagnoses using psychiatrists’ re-evaluations of Spanish-speaking patients showed very good to excellent agreement (κ = 0.74–0.85 for current psychosis; Torrens et al, 2004).

The PRISM follow-up interview was administered in community settings, hospital or in the project offices. Additional data sources for the PRISM diagnosis included diagnostic assessments of clinical staff, hospital charts (baseline only), family/collateral reports of substance use and onset/offset of psychosis, and urine toxicological screens at baseline and follow-up. Symptoms and substance use were considered present when indicated by any data source. When a source indicated that psychotic symptoms antedated heavy substance use, or persisted during at least 4 weeks of abstinence, the PRISM assigned a primary diagnosis.

We compared PRISM primary and substance-induced psychosis at baseline with the 1-year follow-up diagnosis. Diagnostic stability was defined as having the same category (primary or substance-induced psychosis) at baseline and follow-up, and diagnostic change was defined as a shift from baseline substance-induced psychosis to primary psychosis at either the 6-month or 12-month follow-up points. The strict decision rules of the PRISM/DSM-IV procedure minimise the probability of over-diagnosis of a primary psychotic disorder (e.g. a diagnosis of substance-induced psychosis is the ‘default’ in DSM-IV criteria when there is insufficient evidence to support a primary psychotic diagnosis). Sufficient evidence includes psychotic symptoms preceding the onset of substance use, persistence of symptoms for a substantial period after cessation of use, or substantially excessive symptoms given the type, duration and amount of substance used. A diagnosis of primary psychotic disorder is treated as a lifetime designation in this study, although DSM–IV specifies that substance-induced episodes can occur during the 12-month interval in people with a primary psychotic disorder at baseline. New substance-induced psychosis was
diagnosed at follow-up only if the baseline episode remitted for at least 2 months. These cases, unlike those with a diagnostic change from substance-induced to primary disorder, do not represent a true change in diagnostic distinction. An illness classified as either primary psychosis or substance-induced psychosis could have been in remission at either the 6-month or 12-month follow-up with no change in diagnostic category.

**Baseline assessment of the sample**

To explore the predictive validity of baseline characteristics distinguishing primary from substance-induced psychosis at baseline, we used demographic data and information on living arrangement, education, employment, criminal justice contacts, out-of-home placement, current family support and participants’ reports of family history from the Community Care Schedule (Caton, 1997). Family history of mental illness was indicated by a participant’s report of a parent having undergone psychiatric treatment. Parental substance misuse was based on the participant’s report of a parent’s problems with drugs or alcohol (treated or untreated).

Psychiatric symptoms were assessed with the PANSS (Kay et al, 1992). The PANSS total score on overall psychopathology was used for the analyses reported here. The PRISM provided information on visual and auditory hallucinations. Psychosocial, educational and occupational functioning in childhood, adolescence and adulthood were rated with the Premorbid Adjustment Scale (PAS, Cannon-Spoor et al, 1982). The PAS overall score was used in the analyses reported here. The Scale to Assess Unawareness of Mental Disorders (SUMD; Amador et al, 1993) indicated insight into having a mental illness or a reaction to heavy drug use. The instrument yields two scores: ‘unawareness of symptom’ score (lack of awareness of the existence of a psychotic symptom) and ‘misattribution of symptom’ score (lack of understanding that a psychotic symptom is a manifestation of a mental illness or is related to alcohol or drug use).

**Analysis**

Participants’ diagnoses were classified as ‘primary’ or ‘substance-induced’ based on PRISM assessment at three points in time: baseline, 6 months and 12 months. In studying diagnostic stability and change, the distinction between the primary and substance-induced psychosis is the only diagnostic dimension herein reported (e.g. a change from schizophrenia to schizoaffective disorder would not be considered a change for this analysis). When baseline and follow-up diagnoses were compared, three diagnostic categories were created: stable primary psychosis, stable substance-induced psychosis and change from substance-induced psychosis to primary psychosis. Subsequent substance-induced psychotic episodes in participants with a prior diagnosis of primary psychosis did not warrant a change in diagnosis.

These three diagnostic groups were compared on the demographic, family and clinical domains outlined above. We were especially interested in the differences between the ‘change’ group and the stable primary psychosis and substance-induced psychosis groups, and for each domain we examined the binary distinctions between the change group and each of the stable groups. We used logistic regression analyses (Kleinbaum et al, 1998) with the binary diagnostic distinctions as the outcomes and the variables in the domains as explanatory variables. Because of the large number of possible comparisons in these analyses, we adopted the following procedure for containing type I error. Within each domain, we examined model-based likelihood ratio chi-squared test (LRT) omnibus tests to determine if there was evidence that the variables in the domain were related to either the change vs. primary psychosis comparison or the change vs. substance-induced psychosis comparison. If the omnibus test was significant, we examined tests of the individual variables within the domain. Each of these individual variables was also tested using the likelihood ratio test from the logistic regression. This allowed a unified treatment of continuous and categorical variables within the domain. The omnibus tests for the family and clinical domains were adjusted for demographic variables. We show both the adjusted and unadjusted LRT tests for the individual variables.

In a final analysis, the change group (n=34) and the stable primary psychosis group (n=186) were combined to create a group of people who all had a 1-year primary psychosis diagnosis (n=220). We compared this group with the stable substance-induced psychosis group at 1 year (n=99) using the set of baseline demographic, family and clinical characteristics that we had used previously (Caton et al, 2005) to examine the diagnostic distinction at baseline. We entered all these variables at once in a multivariate logistic regression (Kleinbaum et al, 1998) which estimated the unique effect of each variable. Statistical significance was determined using the P<0.05 level and two-tailed tests of significance.

**RESULTS**

**Diagnostic stability and change**

At follow-up, 285 participants (89%) retained their baseline diagnostic category. We identified 10 participants with a baseline diagnosis of primary psychotic disorder that remitted during the follow-up interval who experienced a new substance-induced psychotic episode at some point in the follow-up interval (e.g. onset of psychotic symptoms followed drug ingestion and later remitted within a 4-week drug-free period). This group shared many baseline characteristics with the stable primary psychosis group, including similar scores on positive symptoms (mean total PANSS score 66.5 in contrast to 66.7 for cases of primary psychosis without subsequent substance-induced episodes). However, 80% had a diagnosis of substance dependence in contrast to 45% of those with primary psychotic disorder and no substance-induced psychotic episode. The low number of people in this group obviates meaningful comparisons on baseline predictors. Since their diagnostic classification remained primary psychosis (i.e. the new substance-induced episode did not invalidate the baseline primary classification), these 10 cases were included in the primary psychosis group.

Thirty-four participants (11%) had a change in diagnosis from substance-induced psychosis at baseline to primary psychosis at follow-up (the ‘change’ group). Nearly three-quarters of these (74%; n=25) changed in the first 6 months post-baseline as a result of persistent psychotic symptoms in the absence of substance use. Significant numbers of those in the change group (71%) and the stable substance-induced psychosis group (61%) also carried a diagnosis of misuse of or dependence on any drug (including alcohol) at follow-up, in contrast to 33% in the stable primary psychosis group. The most common primary psychosis diagnoses in the change group were schizophrenia or schizoaffective disorder (n=15; 44%), psychotic mood disorder (n=9; 26%) and psychotic disorder not otherwise specified (n=8; 24%).
Change group v. the stable primary psychosis group

There was no significant difference in demographic characteristics (omnibus LRT=6.2, d.f.=7, NS) (Table 1) or family history (omnibus LRT=5.45, d.f.=2, NS) (Table 2) when the change group and the primary psychosis group were compared. When the clinical domain was considered (Table 3), the difference between the change group and the stable primary psychosis group was significant (omnibus LRT=13.23 d.f.=4, P<0.05). Bivariate tests suggest that the difference between the two groups was owing to the lower baseline PANSS score – indicating less psychopathological disorder – in the change group compared with the stable primary psychosis group. Adjusted and unadjusted bivariate comparisons on the total PANSS score for the primary psychosis group and the change group were significant (P<0.05). There was no significant difference between the two diagnostic groups in bivariate tests of the premorbid adjustment scores or the unawareness of psychosis and misattribution scores.

Substance dependence and associated clinical characteristics (Table 4) differed significantly between the stable primary psychosis group and the change group (omnibus LRT=20.6, d.f.=3, P<0.01). Bivariate comparisons between the stable primary psychosis group and the change group suggest that the difference is chiefly a result of differences in substance misuse or dependence, and to a lesser degree to differences in suicidal ideation. Most (83%) of the change group had a baseline diagnosis of substance dependence, compared with 47% of the stable primary psychosis group. Bivariate tests showed significant group differences (P<0.01) for the unadjusted comparison and a comparison adjusted for demographic variables. Nearly half (47%) of the change group had suicidal ideation at baseline, compared with 28% of the stable primary psychosis group. The bivariate comparison was significant (P<0.05) for the unadjusted comparison, a finding that did not persist when a comparison was adjusted for demographic variables. The two diagnostic groups showed no significant difference in baseline visual hallucinations.

Change group v. the stable substance-induced psychosis group

The change group did not differ significantly from the stable substance-induced psychosis group on demographic characteristics: omnibus LRT=2.49, d.f.=7, NS (see Table 1). However, the family history variables differed between these two groups: omnibus LRT=9.95, d.f.=2,

### Table 2  Family history characteristics of the three diagnostic groups

<table>
<thead>
<tr>
<th></th>
<th>Primary psychosis group (n=186)</th>
<th>Substance-induced psychosis group (n=99)</th>
<th>Change group (n=34)</th>
<th>Statistical test&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
<td>Primary v. change</td>
</tr>
<tr>
<td>Parental mental illness</td>
<td></td>
<td></td>
<td></td>
<td>Unadjusted χ²</td>
</tr>
<tr>
<td>Yes</td>
<td>29 (16)</td>
<td>7 (7)</td>
<td>10 (29)</td>
<td>3.38</td>
</tr>
<tr>
<td>No</td>
<td>157 (84)</td>
<td>92 (93)</td>
<td>24 (71)</td>
<td></td>
</tr>
<tr>
<td>Parental substance use</td>
<td></td>
<td></td>
<td></td>
<td>Unadjusted χ²</td>
</tr>
<tr>
<td>Yes</td>
<td>57 (31)</td>
<td>40 (40)</td>
<td>16 (47)</td>
<td>3.35</td>
</tr>
<tr>
<td>No</td>
<td>129 (69)</td>
<td>59 (60)</td>
<td>18 (53)</td>
<td></td>
</tr>
</tbody>
</table>

1. Likelihood ratio chi-squared test, d.f.=1.
2. Adjusted for age, gender, race, marital status and education level.
*P<0.05, **P<0.01.

### Table 3  Clinical characteristics of the three diagnostic groups

<table>
<thead>
<tr>
<th></th>
<th>Primary psychosis group (n=186)</th>
<th>Substance-induced psychosis group (n=99)</th>
<th>Change group (n=34)</th>
<th>Statistical test&lt;sup&gt;1&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (s.d.)</td>
<td>Mean (s.d.)</td>
<td>Mean (s.d.)</td>
<td>Primary v. change</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Unadjusted χ²</td>
</tr>
<tr>
<td>Premorbid adjustment scale score</td>
<td>0.32 (0.14)</td>
<td>0.31 (0.12)</td>
<td>0.37 (0.15)</td>
<td>3.43</td>
</tr>
<tr>
<td>PANSS</td>
<td>66.72 (21.25)</td>
<td>54.65 (15.45)</td>
<td>57.71 (12.75)</td>
<td>6.35&lt;sup&gt;*&lt;/sup&gt;</td>
</tr>
<tr>
<td>Unawareness score</td>
<td>2.80 (1.57)</td>
<td>1.75 (1.70)</td>
<td>2.59 (1.70)</td>
<td>0.50</td>
</tr>
<tr>
<td>Misattribution score</td>
<td>2.97 (1.81)</td>
<td>2.21 (2.05)</td>
<td>2.75 (1.92)</td>
<td>0.45 (1)</td>
</tr>
</tbody>
</table>

PANSS, Positive and Negative Syndrome Scale
1. Likelihood ratio chi-squared test, d.f.=1.
2. Adjusted for age, gender, race, marital status and education level.
*P<0.05.
Table 4  
Substance use disorder and associated clinical characteristics of the three diagnostic groups

<table>
<thead>
<tr>
<th></th>
<th>Primary psychosis group (n=186)</th>
<th>Substance-induced psychosis group (n=99)</th>
<th>Change group (n=34)</th>
<th>Statistical test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>Primary v. change</td>
</tr>
<tr>
<td></td>
<td>[Unadjusted $\chi^2$]</td>
<td>[Adjusted $\chi^2$]</td>
<td></td>
<td>Unadjusted $\chi^2$</td>
</tr>
<tr>
<td>Any drug use or dependence</td>
<td>Yes 87 (47)</td>
<td>Yes 85 (86)</td>
<td>Yes 28 (82)</td>
<td>15.76**</td>
</tr>
<tr>
<td></td>
<td>No 99 (53)</td>
<td>No 14 (14)</td>
<td>No 6 (18)</td>
<td>0.24</td>
</tr>
<tr>
<td>Visual hallucinations</td>
<td>Yes 25 (13)</td>
<td>Yes 26 (26)</td>
<td>Yes 6 (18)</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>No 161 (87)</td>
<td>No 73 (74)</td>
<td>No 28 (82)</td>
<td>1.08</td>
</tr>
<tr>
<td>Suicidal ideation, past 12 months</td>
<td>Yes 52 (28)</td>
<td>Yes 30 (30)</td>
<td>Yes 16 (47)</td>
<td>4.64*</td>
</tr>
<tr>
<td></td>
<td>No 134 (72)</td>
<td>No 69 (70)</td>
<td>No 18 (53)</td>
<td>3.06</td>
</tr>
</tbody>
</table>

1. Likelihood ratio chi-squared test, d.f.=1.
2. Adjusted for age, gender, race, marital status and education level.
*P < 0.05, **P < 0.01.

Predictive validity of key baseline variables

To test the predictive validity of baseline differences between primary psychotic disorders and substance-induced psychoses in determining psychosis diagnosis at the 1-year assessment, the change group (n=34) was combined with the stable primary psychosis group (n=186) to create a new primary psychosis group (n=220) based on the 1-year diagnosis. The stable substance-induced psychosis group retained its sample size of 99 participants based on the 1-year diagnosis. Table 5 shows the results of a logistic regression for the test of the predictive validity of baseline demographic, family and clinical variables in determining the primary vs. substance-induced psychosis distinction at 1 year. When 1-year psychosis diagnosis was the outcome, three variables that had been found to distinguish the primary and substance-induced psychosis groups at baseline remained the same. The primary psychosis group had greater overall psychopathology assessed with the PANSS, whereas the substance-induced psychosis group had greater substance misuse/dependence and greater visual hallucinations. Although parental substance misuse no longer remained significant at the $P < 0.05$ level, the odds ratio of 1.5 remained within the 95% confidence interval. Importantly, we found that

Table 5  
Logistic regression results for test of predictive validity of baseline variables in determining the distinction between primary psychosis and substance-induced psychosis at the 1-year follow-up (change group added to primary group)

<table>
<thead>
<tr>
<th>Variables</th>
<th>b</th>
<th>(s.e.)</th>
<th>OR</th>
<th>(95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.00</td>
<td>0.02</td>
<td>1.00</td>
<td>(0.96–1.04)</td>
</tr>
<tr>
<td>Female</td>
<td>–0.01</td>
<td>0.33</td>
<td>0.99</td>
<td>(0.52–1.87)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>–0.34</td>
<td>0.33</td>
<td>0.72</td>
<td>(0.37–1.38)</td>
</tr>
<tr>
<td>White/other</td>
<td>–0.81</td>
<td>0.49</td>
<td>0.44</td>
<td>(0.17–1.16)</td>
</tr>
<tr>
<td>Married/cohabiting</td>
<td>0.45</td>
<td>0.37</td>
<td>1.56</td>
<td>(0.75–3.24)</td>
</tr>
<tr>
<td>High-school diploma</td>
<td>–0.11</td>
<td>0.38</td>
<td>0.90</td>
<td>(0.43–1.89)</td>
</tr>
<tr>
<td>Some college</td>
<td>0.16</td>
<td>0.36</td>
<td>1.17</td>
<td>(0.58–2.36)</td>
</tr>
<tr>
<td>Parental substance use</td>
<td>0.42</td>
<td>0.30</td>
<td>1.52</td>
<td>(0.84–2.74)</td>
</tr>
<tr>
<td>Parental mental illness</td>
<td>–0.98</td>
<td>0.49</td>
<td>0.38</td>
<td>(0.15–0.98)</td>
</tr>
<tr>
<td>Total PANSS score</td>
<td>–0.04</td>
<td>0.01</td>
<td>0.96</td>
<td>(0.95–0.98)</td>
</tr>
<tr>
<td>Any drug use/dependence</td>
<td>1.87</td>
<td>0.35</td>
<td>6.48</td>
<td>(3.25–12.91)</td>
</tr>
<tr>
<td>Visual hallucinations</td>
<td>1.11</td>
<td>0.37</td>
<td>3.04</td>
<td>(1.49–6.22)</td>
</tr>
</tbody>
</table>

1. The reference groups for categorical demographic variables were: male, African-American, single, no high-school diploma.
parental mental illness was greater in the primary psychosis group.

**DISCUSSION**

The primary psychosis vs. substance-induced psychosis distinction was remarkably stable over the 1-year follow-up period. Subsequent substance-induced psychotic episodes that occurred in 10 participants with a prior diagnosis of primary psychosis did not warrant a change in diagnosis by PRISM/DSM–IV convention, but clinicians should follow such patients closely to ensure that treatment prescriptions are appropriate, given these patients’ excessive use of alcohol and drugs.

We observed a change in diagnostic category from substance-induced psychosis at baseline to primary psychotic disorder at the 1-year follow-up in 34 study participants, representing about 25% of those diagnosed with substance-induced psychosis at baseline. Greater instability in substance-induced psychosis diagnoses compared with primary psychosis diagnoses had been observed previously (Whitty et al., 2005). The frequency of our research diagnostic assessments over the course of follow-up revealed that the greatest number of diagnostic changes occurred in the first 6-month period of follow-up. The change group shared some of the characteristics of both the stable primary psychosis and the stable substance-induced psychosis groups, but important differences were observed. In contrast to the stable primary psychosis group, the change group had markedly greater rates of substance use disorder, a characteristic shared with the stable substance-induced psychosis group and a small group of participants with primary psychosis who experienced substance-induced psychotic episodes in the follow-up period. Heavy substance misuse overlying presentation of psychotic symptoms in these patients undoubtedly added greater complexity to the diagnostic process. Other factors possibly influencing the diagnostic process include language and cultural differences, unreliable histories, presence of Axis II disorders and cognitive problems.

The significantly lower level of baseline psychotic symptoms in the change group compared with the stable primary psychosis group indicates that at intake these patients’ psychotic disorder was less severe compared with those whose primary psychosis was fully manifest. The greater suicidal ideation in the change group compared with the stable primary psychosis group despite less severe psychotic symptoms underscores their need for thorough assessment and appropriate crisis treatment.

The change group differed from the stable substance-induced psychosis group at the initial presentation on three important dimensions: they had more parental mental illness, poorer premorbid adjustment and less insight into psychosis. The first two of these factors suggest a greater inherent vulnerability to psychosis in the change group compared with their counterparts in the stable substance-induced psychosis group. Clinicians should therefore attend to these indicators and follow such patients longitudinally to monitor the course of psychotic symptoms to ensure diagnostic accuracy and the most appropriate treatment prescriptions, which may ultimately include antipsychotic medication.

Reasons for a change from substance-induced psychotic disorder at baseline to primary psychotic disorder at the 1-year follow-up include several possibilities. The first is that there really was no change in diagnostic status over the follow-up year. Some of the cases diagnosed as substance-induced psychosis at baseline might have actually been primary psychotic disorders that were misdiagnosed owing to the cross-sectional nature of the baseline assessment, and did not have the benefit of observation over time. Moreover, features of the DSM–IV diagnostic criteria for psychotic disorders as implemented in the PRISM may lead to unstable diagnoses, for example if psychotic symptoms co-occur with substance use and an adequate substance-free period does not occur. In such cases the default DSM–IV diagnosis is substance-induced psychosis, consistent with the intent of this diagnostic system to avoid overdiagnosing as primary psychiatric disorders syndromes that are largely the effects of intoxication or withdrawal. Upon follow-up it might be possible to determine whether psychotic symptoms persisted in a subsequent substance-free period, leading to a more accurate diagnostic determination. Thus, ‘change’ cases could be an artefact of the diagnostic criteria rather than indicating true evolution of the disorder. However, a second reasonable possibility is that a substance-induced episode might be a marker for an emerging primary psychosis that was not yet manifest at the first admission. Such individuals might be especially vulnerable to the psychotomimetic properties of substances in the prodromal period prior to the development of a full psychotic disorder. Third, the first episode of a substance-induced psychosis might be part of a process of moving toward an autonomous psychotic disorder in those chronically misusing drugs. Chronic, heavy drug use may alter the brain chemistry in individuals who would not otherwise develop a primary psychosis (Boutros & Bowers, 1996). A clearer delineation of the relationship of substance use and psychosis requires further study employing neuroscientific as well as behavioural approaches. Findings from this investigation should be viewed as preliminary, owing to the small sample size and the unique demographic and social characteristics of the study population.

Of the four key predictors that distinguished the primary psychosis group from the substance-induced psychosis group at baseline (Caton et al., 2005), three – diagnosis of drug misuse/death, total PANSS score and visual hallucinations – remained as key predictors of the diagnostic distinction at 1 year. These findings support conclusions from a cross-sectional investigation reported previously (Rosenthal & Miner, 1997). Parental substance misuse was no longer significant at the 0.05 level, although its odds ratio of 1.5 remained within the 95% confidence interval. In addition, parental mental illness was found to be greater in the primary psychosis group – a finding that emerged at the 1-year follow-up that was not observed at baseline.

**Clinical implications**

The predictive validity of these baseline variables underscores their utility in assisting psychiatric emergency clinicians to make more accurate diagnoses and more appropriate treatment prescriptions when patients with early-phase psychotic disorders and substance use comorbidity initially present for treatment. Longitudinal follow-up of patients initially presenting with psychosis and substance use comorbidity is warranted by the occurrence of heavy substance misuse overlying presentation of psychotic symptoms, adding greater complexity to the diagnostic process, and the greater instability of substance-induced psychosis diagnoses.

**Limitations**

Our findings are based on an ethnically mixed sample of substance-using patients...
recruited from New York City psychiatric emergency departments, and might not be generalisable to other populations selected from different types of service settings (Kirkbride et al., 2006), although further research could clarify this issue. In addition, our findings are based solely on behavioural data. A clearer delineation of the relationship of substance use and psychosis requires further study employing neuroscientific as well as behavioural approaches. Continued longitudinal follow-up beyond 1-year will clarify the long-term outcome of disorders characterised by psychosis and substance use comorbidity.

ACKNOWLEDGEMENTS

Support for this investigation was provided by research grants RO1 DA10539 and 2R01-DA10539-06 from the US National Institute on Drug Abuse and K05AA014223 from the National Institute on Alcohol Abuse (DSH). The authors are indebted to Ingrid Ramirez, Eustace Hsu and Milagros Ventura for their assistance with data processing and manuscript preparation.

REFERENCES


British Journal of Psychiatry


Correspondence: Professor Carol L. M. Caton, Department of Psychiatry, New York State Psychiatric Institute, Columbia University, 1051 Riverside Drive, New York, 10032; DEBORAH S. HASIN, PhD, Mailman School of Public Health, New York State Psychiatric Institute, Columbia University, New York; PATRICK E. SHROOT, PhD, Department of Psychology, New York University, New York; ROBERT E. DRAKE, MD, PhD, Department of Psychiatry, Dartmouth Medical School, Lebanon; BOANERGES DOMINGUEZ, MS, Mailman School of Public Health, Columbia University, New York; MICHAEL B. FIRST, MD, SHARON SAMET, MSW, New York State Psychiatric Institute, New York; BELLA SCHANZNER, MD, Department of Psychiatry, Columbia University, New York, New York, USA

(First received 28 July 2005, final revision 14 August 2006, accepted 29 September 2006)
Lipids and essential fatty acids in patients presenting with self-harm†

MALCOLM R. GARLAND, BRIAN HALLAHAN, MAIREAD MCNAMARA, PHILIP A. CARNEY, HELEN GRIMES, JOSEPH R. HIBBELN, ANDREW HARKIN and RONAN M. CONROY

Background Low cholesterol has been reliably demonstrated in people who self-harm.

Aims To determine whether people who self-harm also have low levels of essential fatty acids (EFAs) and to examine associations between the EFAs and serotoninergic function.

Method Depression, impulsivity and suicidal intent were measured in patients with self-harm (n=40) and matched controls, together with plasma lipids and EFAs. Platelet serotoninergic studies were carried out in a subgroup (n=27).

Results Patients with self-harm had significantly more pathology on all psychometric measures, lower mean total cholesterol levels (4.18 (s.d. =0.93) v. 4.87 (s.d. =0.83) mmol/l, P=0.003) and lower mean total EFA levels (89.5 (15.6) v. 103.7 (17.1) μg/ml, P=0.0001) than controls after adjustment for confounding variables. Total n-3 and n-6 EFA levels were also significantly lower. Impulsivity and depression scores were significantly inversely correlated with both n-6 EFAs and n-3 EFAs, but were not associated with total or low-density lipoprotein cholesterol levels. Platelet serotoninergic measures did not differ between groups, and were not related to psychobiological measures.

Conclusions Lower plasma EFA levels combined with low cholesterol concentrations were associated with self-harm as well as impulsivity and affect. This was not related to platelet serotoninergic measures.

Declaration of interest None

Self-harm is a high-risk pathological behaviour which clearly has multiple determinants. Mounting epidemiological, basic scientific and clinical intervention data indicate that low levels of circulating lipids, omega-3 essential fatty acids (n-3 EFAs) and cholesterol are risk factors for impulsive and depressive behaviours. An association of low cholesterol with a recent act of self-harm has been frequently demonstrated (Garland et al, 2000; Lester, 2002). The n-3 EFAs are selectively concentrated in the brain but are obtained exclusively from diet, seafood and fish being the primary source. Countries with greater per capita rates of seafood consumption have lower rates of major depression, bipolar depression, post-partum depression and mortality from homicide (Hallaahan & Garland, 2005). Greater intake of fish was associated with a lower risk of suicide among 260,000 Japanese men followed for 17 years (Hirayama, 1990) and a lower risk of suicidal ideation among 1767 Finns (Tanskanen et al, 2001). Since self-harm occurs with disproportionate frequency among people with these pathologies, it seems likely that it might be more frequent among people with low plasma n-3 EFAs.

Patients with self-harm exhibit a convergence of many constructs related to serotoninergic dysfunction (e.g. depression, impulsivity and violence; Mann, 2003). Low concentrations of serotonin metabolites (5-hydroxyindoleacetic acid (5-HIAA)) in the cerebrospinal fluid of those who attempt suicide is a well-replicated finding and is frequently linked to impulsive and self-destructive behaviours (see Roggenbach et al, 2002). The primary aim of this study was to examine lipid and EFA levels in patients following an act of self-harm. Secondary aims were to examine associations between lipids/EFAs and two common psychopathological constructs in self-harm: impulsivity and depression. We also sought to explore, in a subgroup, any mediating influence of the serotonergic system.

METHOD

Participants

Consecutive patients (n=40) presenting to the accident and emergency unit of Galway University Hospital were recruited over an 18-month period and tested within 24 h of admission. Exclusion criteria were: consumption of fish more than once a week; age <16 or >65 years; requiring resuscitation with fluids or ventilatory support; serious injury or medical complication following self-harm; current psychiatric diagnoses of addiction, psychotic disorders, or eating disorders; presence of any illness, treatment or diet known to affect plasma cholesterol; history of cardiovascular or lipid disorder; recent weight loss. Patients with self-harm involving psychotropic substances or previously taking prescribed psychotropic medication (apart from low-dose benzodiazepines) were also excluded as these substances would interfere with measures of platelet serotonin (Owens & Nemeroﬀ, 1994). Controls, matched for age and gender, were recruited from the medical day ward. The same exclusion criteria (including the requirement for fish consumption no more than once weekly) were applied, with the addition that they had no previous or current psychiatric history. There were no refusals among patients or controls.

Assessments

In addition to demographic data, the following were documented: body mass index (BMI); weekly alcohol intake (with cut-off points defined by the Royal College of Psychiatrists’ guidelines (1986) for men (≥21 units/week) and women (≥14 units/week). In addition, all participants were diagnosed according to ICD-10 criteria (World Health Organization, 1993) with the relevant Axis I diagnosis. Menstrual cycle status was recorded as being follicular, peri-ovulatory or luteal.

The psychometric measures used were as follows. The Suicide Intent Scale (SIS; Beck et al, 1974) is a 15-item scale which is rated from 0 to 2 for each item, giving a total score range of 0–30. The Beck Depression Inventory (BDI; Beck et al, 1961) is a 21-item self-rated instrument.

See pp. 118–122, this issue.
The Barratt Impulsivity Scale–II (BIS; Barratt, 1994) is a 34-item instrument in which all items are answered on a 4-point scale. There are three sub-scales: attentional impulsivity (e.g. 'I concentrate easily', 'I get easily bored when solving thought problems'), motor impulsivity (e.g. 'I do things without thinking', 'I am self-controlled') and non-planning impulsivity (e.g. 'I plan tasks carefully', 'I finish what I start').

**Laboratory methods**

Fasting antecubital venous blood was drawn from all participants in the morning. All samples were immediately frozen and stored at −80°C until use. Plasma for EFA analysis was transported on dry ice by overnight courier to the USA (to J.R.H.). Owing to insufficient blood, serotonergic data are missing for 13 patients (and controls). Cholesterol measures and EFA data were available for all participants.

**Cholesterol**

Total plasma cholesterol was measured on a Beckman Synchron CX7 Analyser by an enzymatic timed end-point method (Allain et al, 1974).

**Plasma EFAs**

Fatty acids were extracted from plasma by adding 200 μl to 2 ml trichloromethane (CHCl₃), 1 ml BHT-MeOH and a known quantity of 23:0 methyl ester as an internal standard. After brief vortexing 1 ml of 0.2 mol/l disodium hydrogen phosphate (Na₂HPO₄) was added. The samples were capped under nitrogen and vortexed again. After centrifugation CHCl₃ was removed and dried under nitrogen. Samples were methylated under boron trifluoride/methanol for 60 min. Samples were kept cold and under nitrogen throughout the analysis to prevent oxidation. Gas chromatography was performed on a HP 5890 series II with a flame ionisation detector, an autosampler and a FFAP capillary column. Peaks were identified using authentic standards. Fatty acids were quantified by comparison with peak areas of the 23:0 internal standard. When samples were subjected to thawing and refreezing, within- and between-run coefficients of variance were less than 0.3% and 5% respectively.

**Platelet studies**

Many studies support the association between platelet serotonergic measures and psychopathology, the platelet having physiological properties in common with the neuron (Owens & Nemeroff, 1994). This measure has the advantage of being less invasive than other methods of assessing serotonergic function (e.g. lumbar puncture for 5-hydroxyindole acetic acid (5-HIAA) or neuroendocrine probes).

Pellets from platelet-rich plasma were frozen at −80°C for later use in the (³H)-paroxetine binding assay as described previously (Kelly et al, 1999). The data for each participant were reduced initially by using iterative curve-fitting routines (GraphPad version 2.0; http://www.graphpad.com/prism/Prism.htm) to yield dissociation equilibrium constants (Kₐ) and maximal binding capacity (Bₘ₉) (saturation analyses – one-site binding hyperbola). The Bₘ₉ reflects the number of functioning serotonin transporter molecules on the platelet membrane; the Kₐ is an inverse measure of the affinity of the ligand (paroxetine) for the transporter.

**Statistical analysis**

Data were analysed using Stata 9.2/SE Release. Linear regression was used to estimate mean values and their confidence intervals for patients and controls, with adjustment for confounding variables. However, because some of the fatty acid measures were not normally distributed, statistical comparisons between patients and controls on these measures were made using ordered logistic regression with conversion of the fatty acid measures to deciles. Robust variance estimation (Huber–White estimation) was used in all regression models to compensate for the non-independence of observations within case–control pairs. Partial correlations were calculated between fatty acid measures and psychometric scores, adjusting for the effects of social class, alcohol and smoking. These correlations were calculated with both fatty acid and psychometric measures transformed to decile scores to guard against the potential influence of outlier values.

**RESULTS**

Table 1 shows the demographic characteristics of the patients and controls. In general, controls were from a higher social class than patients and a higher proportion were employed. Patients had a higher prevalence of smoking and excess alcohol intake, were less likely to exercise regularly but had a similar mean BMI.

Mean total n-6 EFA levels were not associated with social class (P=0.653), smoking (P=0.192), alcohol consumption (P=0.471), or regular exercise (P=0.540). Mean total n-3 EFA levels were not associated with social class (P=0.448), smoking (P=0.740), alcohol consumption (P=0.610), or regular exercise (P=0.808). Mean total cholesterol levels were not associated with social class (P=0.616), smoking (P=0.876), or regular exercise (P=0.486). They were, however, associated with alcohol consumption (P=0.037), cholesterol levels being 0.43 IU/l lower in those with excess alcohol consumption. Mean high-density lipoprotein (HDL) cholesterol levels were not associated with social class (P=0.070), smoking (P=0.906), or regular exercise (P=0.653). They were, however, associated with alcohol consumption (P=0.043), mean levels being 0.2 IU/l higher in those with excess alcohol consumption. Mean low-density lipoprotein (LDL) cholesterol levels were not associated with social class (P=0.423), smoking (P=0.640), or regular exercise (P=0.425). They were, however, associated with alcohol consumption (P<0.001), mean levels being 0.73 IU/l lower in those with excess alcohol consumption.

To examine the effect of alcohol on the estimation of differences in lipid levels between patients and controls, we performed regressions unadjusted and adjusted for alcohol consumption, followed by Wald post hoc comparisons of the coefficients associated with case–control status between the models. For total cholesterol, the coefficient associated with case–control status was unchanged when alcohol consumption was added to the model (Wald P=0.766). Although HDL levels did not differ significantly between patients and controls, the coefficient was also unchanged when alcohol was added to the model (Wald P=0.750). Likewise, the changes in the coefficients for LDL cholesterol (Wald P=0.328) and triglycerides (P=0.805) were non-significant.

Two of the patients met ICD–10 criteria for depressive disorder and a further four met criteria for adjustment disorder. Scores on the BDI were significantly higher in patients than controls. The BIS scores were also higher in patients than controls, both overall and on the three sub-scales.
Table 2 shows the lipid measures. Patients had significantly lower total cholesterol concentrations than controls and lower LDL cholesterol concentrations. However, when LDL levels were corrected for total cholesterol using the cholesterol/LDL ratio, the difference between patients and controls was not statistically significant. The other lipid indices, HDL cholesterol and triglycerides, were similar in patients and controls. Adjusting lipid comparisons for alcohol, smoking and social class did not make any substantive change to these findings. Subsequent addition of age and gender to the models was not associated with any of the variables. No consistent relationship was demonstrated between lipids and psychopathology (data not shown).

There were no significant differences in platelet serotonergic indices between patients and controls, although B_max for patients was lower, at 371 (s.d.=265) vs. 458 (s.d.=229) fmol/mg, for controls. The K_a was 0.313 (s.d.=0.44) and 0.304 (s.d.=0.49) nM for patients and controls respectively.

Table 3 shows mean fatty acid levels in patients and controls, adjusted for social class, smoking and alcohol consumption. The levels of statistical significance are based on ordered logistic regression using deciles of fatty acids. Two comparisons between patients and controls are shown: the first compares absolute levels of fatty acids, corrected for social class, alcohol and smoking; the second also adjusts for total fatty acid level.

Total fatty acids, total saturates and total monounsaturates did not differ significantly between patients and controls after controlling for alcohol, social class and smoking status. However, there were significant differences in levels of EFAs, total n-6 fatty acids and linoleic acid, with higher levels in controls. Total n-3 fatty acids were also higher in controls than in patients, as were eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) levels. There were no significant differences between patients and controls in arachidonic acid or alpha-linoleic acid levels.

When we controlled for total fatty acid level, in addition to social class, smoking and alcohol, the differences reported above remained statistically significant. In addition, patients had a significantly higher level of monounsaturates.

Table 4 shows the partial correlations between fatty acids and scores on the BIS and the BDI, adjusted for social class, alcohol consumption and smoking. There were significant negative correlations between both impulsivity and depression scores and total n-6 and n-3 fatty acids.

We also examined the relationship of total cholesterol and LDL cholesterol to impulsivity and depression scores, using multiple regression to adjust for differences in mean score levels between patients and controls, as well as for smoking, alcohol and social class. Neither total cholesterol nor LDL cholesterol was associated with variation in impulsivity or depression scores, when adjusted in this manner. Among the patients there was no relationship between lipid parameters and degree of suicidal intent.

**DISCUSSION**

Compared with controls, matched for age, gender and crudely matched for (low) weekly fish consumption, patients with self-harm had significantly lower levels of total cholesterol and LDL concentrations, as reported previously (Garland et al, 2000). They also had lower circulating concentrations of total n-3 and n-6 fatty acids as well as the principal central nervous system n-3 EFAs, DHA and EPA. Each of these findings was significant after adjustment for alcohol consumption, smoking,
social class and other demographics. Finally, in a regression model, low total n-3 and n-6 EFA levels were significantly associated with greater depression and impulsivity scores. There was no such independent relationship for cholesterol/LDL ratio and psychometric scores. For the subgroup analysed, there was no difference in platelet serotonergic measures between patients and controls, and no association with psychobiological parameters.

Weaknesses in this study included the small population, although it was sufficiently powered to detect the primary findings. None the less, such a small study may result in some false-negative conclusions and therefore any negative findings should be treated as absence of evidence, not evidence of absence. Likewise, for the positive findings small sample size limited the statistical power to control for possible confounding factors. The disparity in demographic variables between patients and controls was marked and potentially important; however, we were fortunate that the only confounder in the psychobiological data was an association between alcohol consumption and total, LDL and HDL cholesterol levels. Accordingly minimal adjustment was necessary prior to analysis. The loss of samples for the serotonergic studies was unfortunate and these data need to be considered with caution owing to the possibility of a false-negative result (type II error).

With such caveats in mind, our findings suggest there is a greater risk of self-harm and psychopathology predisposing to self-harm in individuals with low plasma

### Table 2
Mean (s.d.) lipid variables in patients with self-harm and controls

<table>
<thead>
<tr>
<th>Variable</th>
<th>Patients</th>
<th>Controls</th>
<th>$P^1$</th>
<th>$P^2$</th>
<th>$P^3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total cholesterol, mmol/l</td>
<td>4.18 (0.93)</td>
<td>4.87 (0.83)</td>
<td>&lt;0.001</td>
<td>0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Triglycerides, mmol/l</td>
<td>1.39 (0.71)</td>
<td>1.13 (0.60)</td>
<td>0.088</td>
<td>0.565</td>
<td>0.789</td>
</tr>
<tr>
<td>HDL cholesterol, mmol/l</td>
<td>1.67 (1.00)</td>
<td>1.56 (0.35)</td>
<td>0.823</td>
<td>0.241</td>
<td>0.542</td>
</tr>
<tr>
<td>LDL cholesterol, mmol/l</td>
<td>2.03 (0.85)</td>
<td>2.79 (0.79)</td>
<td>&lt;0.001</td>
<td>0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cholesterol/LDL ratio</td>
<td>2.94 (0.96)</td>
<td>3.28 (1.00)</td>
<td>0.128</td>
<td>0.296</td>
<td>0.075</td>
</tr>
</tbody>
</table>

HDL, high-density lipoprotein; LDL, low-density lipoprotein.
1. Original logistic regression unadjusted for confounding factors.
2. Adjusted for alcohol and smoking.
3. Adjusted for alcohol, smoking and social class.

### Table 3
Plasma concentrations of fatty acids in patients with self-harm and controls

<table>
<thead>
<tr>
<th>Plasma fatty acids, µg/ml</th>
<th>Patients</th>
<th>Controls</th>
<th>$P^1$</th>
<th>$P^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (95% CI)</td>
<td>Mean (95% CI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total fatty acids</td>
<td>243.9 (225.9–261.9)</td>
<td>267.2 (247.0–287.4)</td>
<td>0.215</td>
<td>0.233</td>
</tr>
<tr>
<td>Total saturates</td>
<td>79.5 (73.1–86.0)</td>
<td>87.6 (79.5–95.7)</td>
<td>0.221</td>
<td>0.206</td>
</tr>
<tr>
<td>Total monounsaturates</td>
<td>76.6 (68.9–84.3)</td>
<td>74.0 (66.0–82.0)</td>
<td>0.422</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total essential fatty acids</td>
<td>87.8 (82.3–93.3)</td>
<td>105.6 (99.4–111.7)</td>
<td>0.002</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Total n-6 fatty acids</td>
<td>79.5 (74.3–84.6)</td>
<td>93.8 (88.2–99.4)</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>Linoleic acid (18:2 n-6)</td>
<td>59.7 (55.4–63.9)</td>
<td>71.4 (67.0–75.7)</td>
<td>0.006</td>
<td>0.019</td>
</tr>
<tr>
<td>Arachidonic acid (20:4 n-6)</td>
<td>14.4 (13.1–15.6)</td>
<td>16.1 (14.6–17.6)</td>
<td>0.113</td>
<td>0.295</td>
</tr>
<tr>
<td>Total n-3 fatty acids</td>
<td>8.3 (7.4–9.1)</td>
<td>11.8 (10.3–13.3)</td>
<td>&lt;0.001</td>
<td>0.011</td>
</tr>
<tr>
<td>α-Linolenic acid (18:3 n-3)</td>
<td>1.7 (1.5–1.9)</td>
<td>1.9 (1.7–2.2)</td>
<td>0.298</td>
<td>0.676</td>
</tr>
<tr>
<td>Eicosapentaenoic acid (20:5 n-3)</td>
<td>1.7 (1.3–2.1)</td>
<td>3.2 (2.4–4.0)</td>
<td>0.002</td>
<td>0.015</td>
</tr>
<tr>
<td>Docosahexaenoic acid (22:6 n-3)</td>
<td>3.7 (3.2–4.2)</td>
<td>5.1 (4.4–5.8)</td>
<td>0.003</td>
<td>0.016</td>
</tr>
</tbody>
</table>

1. Ordered logistic regression, adjusted for alcohol, smoking and social class.
2. Ordered logistic regression, adjusted for alcohol, smoking and social class and total fatty acid level.

### Table 4
Partial correlations between deciles of fatty acid measures (absolute and relative) and deciles of the Barratt Impulsivity Scale and Beck Depression Inventory

<table>
<thead>
<tr>
<th></th>
<th>Partial correlation</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barratt Impulsivity Scale (decile)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n-6, µg/ml (1 decile)</td>
<td>$-0.309$</td>
<td>0.007</td>
</tr>
<tr>
<td>n-3, µg/ml (1 decile)</td>
<td>$-0.238$</td>
<td>0.038</td>
</tr>
<tr>
<td>Beck Depression Inventory (decile)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n-6, µg/ml (1 decile)</td>
<td>$-0.295$</td>
<td>0.010</td>
</tr>
<tr>
<td>n-3, µg/ml (1 decile)</td>
<td>$-0.301$</td>
<td>0.009</td>
</tr>
</tbody>
</table>

1. Correlations are adjusted for social class, alcohol intake and smoking.
cholesterol (both total and LDL cholesterol) and low n-3 and n-6 EFAs. Such predisposition is likely to be longstanding, as one of the two psychopathological constructs we measured, impulsivity, is a trait construct. Although cholesterol is an important component of the neuronal membrane (Engelberg, 1992), we feel it is acting as a marker or ‘bystander’ for EFA levels and is not the primary source of the deficit in the participants with self-harm. When at the lower end of the spectrum, low EFA levels (in dietary supplementation studies) are associated with low cholesterol levels (Harris et al., 1997); the brain synthesises its own cholesterol de novo (Pardridge & Mietus, 1980) and there is no relationship between peripheral and central (nervous system) levels, contrasting with the tight correlation between EFAs in plasma and cerebrospinal fluid (Yao et al., 2002). In this study, any relationship between impulsivity and depression scores and cholesterol/LDL ratio disappeared when adjusted for EFA levels. Finally, attempts to correlate cholesterol with neurotransmitter function have generally been negative, whereas there are strong data for the EFAs (although we found no such correlation in our own small sample).

Causative factors for low lipid and EFA levels

Although our estimate of EFA intake was crude, the observed differences in EFA levels between groups are unlikely to be accounted for purely by differences in dietary intake. Apart from socio-demographic factors that were controlled for in regression analysis, factors accounting for the observed lower levels of cholesterol and EFAs warrant consideration.

Perinatal factors

Levitsky & Strupp (1995) demonstrated the enduring effects of malnutrition in the perinatal period on learning in rats. Looff et al. (1991) have shown similar effects in bottle-fed children, and breast milk is a richer source of EFAs. Data on perinatal nutrition were not recorded in our study.

Stress

Several mechanisms may bring about stress-related depletion of EFAs, although this has not been demonstrated in human models. Chronic restraint stress in rats was associated with increased lipid peroxidation, with resultant neuronal phospholipid depletion (Gulyaeva et al., 1989). Reductions in membrane-protective antioxidants, such as superoxide dismutase, in a rat immobilisation stress model (Sosnovskii et al., 1993) have been demonstrated as another potential pathway.

Stress and diet

Periods of emotional stress, as would be a prelude to self-harm, can lead to changes in diet, such as a switch from lipid- and protein-rich foods to carbohydrate-rich foods (Cohen et al., 2002). This could alter levels of both cholesterol and EFAs. In our study, dietary history was not recorded, although there was no difference in BMI between groups.

Genetic factors

Allelic variation in one of many genes encoding enzymes in the anabolic and catabolic pathways of EFA and lipid metabolism could account for differences in peripheral levels of EFAs and lipids. However, apart from the report of polymorphism in the phospholipase A₂ gene in schizophrenia (Peet et al., 1998), no other abnormalities have been reported in psychiatric populations.

Low cholesterol/EFA and behaviour

As proposed by Engelberg (1992) there is a sound theoretical basis for linking low cholesterol levels with impaired serotonergic neurotransmission and a number of in vitro studies support this (Heron et al., 1980; Scanlon et al., 2001). However, in accordance with this study, clinical investigations of the cholesterol-serotonin relationship (using a variety of measures of serotonergic activity) have, with the exception of that of Terao et al. (2000), been negative (e.g., Alvarez et al., 1999).

More studies support a direct role for EFAs in neurotransmission and neuronal function in general (see Hallahan & Garland, 2005). For example, concentrations of serotonin, dopamine and their respective metabolites nearly doubled after only 18 days in piglets that were fed formulas supplemented with arachidonic acid (an n-6 EFA) and DHA (an n-3 EFA) (de la Presa Owens & Innis, 1999). In a rat model (Delion et al., 1996) of diet-induced n-3 EFA deficiency, increased 5-HT₂ platelet receptor density (indicative of reduced serotonin neurotransmission) was reported. The only human studies of which we are aware (for a review see Hibbelsn, 1999) reported significant positive correlations between 5-HIAA in the cerebrospinal fluid and a variety of n-3 and n-6 EFAs in 45 healthy volunteers and an effect of n-3 EFA supplementation on adenosine diphosphate (ADP)-induced serotonin amplification in platelets of patients with schizophrenia (Yao et al., 2002). However, the results are difficult to interpret. Our sample may have been too small to detect significant relationships between EFA and platelet serotonin.

Cholesterol/EFAs and psychiatric disorder

There appears to be a J-shaped relationship between cholesterol and mortality, cohorts with low cholesterol being at increased risk from violent death, including suicide (reviewed in Garland et al., 2000). Importantly, more recent data suggest that the therapeutic lowering of cholesterol (by diet or statins) does not increase non-cardiovascular mortality (including suicide) (Muldoon et al., 2001) or psychological morbidity (Stewart et al., 2000). However, unlike the fibrates, neither of these cholesterol-lowering methods lowers EFAs. An earlier and widely cited meta-analysis of cholesterol-lowering trials (Muldoon et al., 1990) that reported the increase in suicide analysed mainly trials using fibrate. Although there are fewer data for epidemiological links between EFAs and mental health, they are less equivocal. Strong negative correlations were found between national per capita fish consumption and prevalence of post-partum depression (Hibbelsn, 2002) and depression (Hibbelsn, 1999). The 100-fold increase in the prevalence of depression in North America over as many years directly correlates with the increase in consumption of saturated fats and n-6 EFAs as part of the ‘modern’ Western diet, at the expense of n-3 EFAs (Hibbelsn, 1999).

Lipids and EFAs in clinical populations

There is now clear evidence that self-harm is associated with lowered cholesterol (Lester, 2002). This reflects similar deficits in EFAs reported in the only previous study of self-harm of which we are aware (Huan et al., 2004), in which significant diminution of n-3 EFAs (but not n-6 EFAs) was found in 100 patients but not in 100
matched controls, after controlling for dietary intake of fish. Two studies support an association between low levels of n-3 EFAs and impulsive/violent behaviour (Virkkunen et al., 1987; Stevens et al., 1995). Moreover, there is mounting evidence that EFA deficits are present in syndromal psychiatric disorders (Hallahan & Garland, 2003). Our accompanying paper (Hallahan et al., 2007, this issue) details a randomised controlled trial of n-3 EFA supplementation in a similar population with self-harm.

REFERENCES


Malcolm R. Garland, MD, MRCPsych, MRCPI, Brian Hallahan, MD, MRCPsych, Department of Adult Psychiatry, Galway University Hospital and the Clinical Sciences Institute, National University of Ireland, Galway; MAIREAD McNAMARA, PhD, Department of Experimental Pharmacology, National University of Ireland, Galway; PHILIP A. CARNEY, FRCPsych, Department of Adult Psychiatry, Galway University Hospital and the Clinical Sciences Institute, National University of Ireland, Galway; HELEN GRIMES, PhD, Department of Clinical Biochemistry, Galway University Hospital; JOSEPH R. HIBBELN, MD, Laboratory of Membrane Biochemistry and Biophysics, National Institute on Alcohol Abuse and Alcoholism, Rockville, Maryland, USA; ANDREW HARKIN, PhD, Department of Experimental Pharmacology, National University of Ireland, Galway; RONAN M. CONROY, DSc, Department of Epidemiology and Biostatistics, The Royal College of Surgeons in Ireland, Dublin, Ireland.

Correspondence: Dr Malcolm R. Garland, St Ita’s Hospital, Portrane, County Dublin, Ireland. Email: mgarland@ireland.com

(First received 15 November 2005, final revision 29 August 2006, accepted 3 October 2006)
Omega-3 fatty acid supplementation in patients with recurrent self-harm

Single-centre double-blind randomised controlled trial†

BRIAN HALLAHAN, JOSEPH R. HIBBELN, JOHN M. DAVIS and MALCOLM R. GARLAND

Background Trials have demonstrated benefits of long-chain omega-3 essential fatty acid (n-3 EFA) supplementation in a variety of psychiatric disorders.

Aims To assess the efficacy of n-3 EFAs in improving psychological well-being in patients with recurrent self-harm.

Method Patients (n=49) presenting after an act of repeated self-harm were randomised to receive 1.2 g eicosapentaenoic acid plus 0.9 g docosahexaenoic acid (n=22) or placebo (n=27) for 12 weeks in addition to standard psychiatric care. Six psychological domains were measured at baseline and end point.

Results At 12 weeks, the n-3 EFA group had significantly greater improvements in scores for depression, suicidality and daily stresses. Scores for impulsivity, aggression and hostility did not differ.

Conclusions Supplementation achieved substantial reductions in surrogate markers of suicidal behaviour and improvements in well-being. Larger studies are warranted to determine if insufficient dietary intake of n-3 EFAs is a reversible risk factor for self-harm.

Declaration of interest Pronova (now Epax) AS, Lysaker, Norway, provided the active preparation and placebo but were not otherwise involved in the study. Funding detailed in Acknowledgements.

†See pp. 112–117, this issue.

Therapeutic trials of the n-3 essential fatty acids (EFAs) eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA) have yielded promising data for a broad spectrum of psychiatric disorders (Hallahlan & Garland, 2003). Moreover, rapid advances in the basic science have confirmed the central role of the n-3 EFAs in a variety of pathophysiological processes, particularly those involving monoaminergic neurotransmission (Sublette et al, 2004). Patients with recurrent non-fatal self-harm exhibit affective disturbances and impulsive and violent behaviours that have been repeatedly linked to deficits in monoaminergic (particularly serotonergic) neurotransmission (Mann, 2003). Self-harm is a major cause of morbidity worldwide and presents a major therapeutic challenge. This single-centre 12-week randomised active placebo-controlled trial is the first of n-3 EFA supplementation in patients with self-harm.

METHODS

Setting
All participants were recruited from the accident and emergency (A&E) Department of Beaumont Hospital, an academic teaching hospital in Dublin, Ireland. Approximately 300 cases of self-harm are treated annually. The study was approved by the hospital ethics committee.

Participants
Consecutive patients (aged 16–64 years) who presented acutely with self-harm by whatever means and had a lifetime history of at least one other episode were candidates for study inclusion and were approached in the A&E department. Exclusion criteria were: current history of addiction, substance misuse, psychosis or eating disorder; currently receiving psychotherapy; known history of dyslipidemia; any treatment, diet or illness known to interfere with lipid or n-3 EFA metabolism; weight loss >10% over the previous 3 months; taking supplements containing n-3 EFAs or consuming fish more than once per week; changes to, or introduction of, psychotropic medication during the previous 6 weeks; unwillingness to participate in the study or living outside the greater Dublin area. During the course of the study patients continued to receive standard psychiatric care at either primary or specialised level and had changes to their psychotropic medication as prescribed by their treating agency. At the time of screening, all patients were free of intoxicants, had not suffered significant harm as a result of the self-harm act and had mental capacity. In addition to the outcome measures at baseline, participants completed the Structured Clinical Interview for DSM–III–R, Axis II version (SCID–II, First et al, 1997).

Sample size
Power analysis was carried out using nQuery Advisor, version 4 (Statistical Solutions, MA, USA). Using published data on completed suicide (Zahl & Hawton, 2004) and recurrence of self-harm (Kapur et al, 2004), the number needed to detect a 20% difference (P<0.05) between groups would be approximately 1500 and 100 respectively, which was not feasible. However, to detect a 20% difference (P<0.05) in outcome measures using psychometric instruments (s.d. ranges 2–9), a total population of between 35 and 50 was estimated.

Outcome measures
Changes in the following psychological domains at study end point v. baseline were the outcome measures.

Depression
The 21-item Beck Depression Inventory (BDI; Beck et al, 1961) and the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960) were used to assess depression.

Suicidality and aggression
The Overt Aggression Scale, Modified (OAS–M, Coccaro et al, 1991) was used. This is a clinician-administered semi-structured interview that was adapted from the OAS (Yudofsky et al, 1986) for use with out-patients. It measures type, severity and frequency of aggressive behaviour. The three sub-scales measure aggression, irritability and suicidality on 4-point Likert scales. The instrument has satisfactory psychometric properties and is sensitive to change (Coccaro et al, 1991).
**Impulsivity**

This was assessed with Immediate and Delayed Memory Tasks (IMT/DMT; Dougherty et al., 2002), which is a computer-based variant of the Continuous Performance Test (CPT) and is designed to objectively measure attention, memory and impulsivity. The IMT/DMT has satisfactory psychometric properties and is sensitive to change (Dougherty et al., 2002).

**Daily stresses**

The Perceived Stress Scale (PSS; Cohen et al., 1983), a 14-item measure of the degree to which situations in one’s life are appraised as stressful, was used to measure stress. Items are rated on a 5-point frequency scale ranging from 0, ‘never’, to 4, ‘very often’, for how often in the past month the person has experienced stress-related feelings and thoughts. A total perceived stress score is obtained by reversing the scoring on the positive items and summing responses across the 14 items. Higher scores indicate more stress. The instrument is widely used and has satisfactory psychometric properties.

The Daily Hassles and Uplifts Scale (DHUS; Kanner et al., 1981) is a well-validated instrument which consists of a list of 18 items (such as children, parents, spouse, work, time for self and health) that can be sources of strain, stress and hassle. Items were clustered into five domains of sources of stress: family (parents, children, spouse, in-laws and other relatives); self (appearance, time for self and health); roles (work-related stress, relationships with co-workers and household jobs); social-environmental (housing, environment, security and socio-political situation); and financial. Sources of uplift were measured with a similar list of items, with a few modifications. Respondents were asked to indicate whether each item had been a source of pleasure or enjoyment for them during the past week. Items were clustered into the same five domains as above. For the purposes of analysis, total hassles scores were subtracted from total uplifts scores.

**Intervention, randomisation and masking**

Participants were prescribed four identical capsules of either active agent or placebo (both provided by Epax (formerly Pronova) AS, Lysaker, Norway), to be taken in the morning. Active capsules (EPAX 5500) contained 305 mg EPA and 227 mg DHA (total EPA plus DHA of 532 mg/capsule). The total dose equalled 2128 mg/day of EPA plus DHA. Placebo contained 99% corn oil and a 1% EPA/DHA mixture. This ensured a degree of equality in the incidence of ‘fishy breath’, the most frequent side-effect of taking active treatment. Adherence was encouraged by weekly telephone contact and ascertained from pill count. Participants who did not return or who missed appointments were deemed to be non-adherent. An independent colleague dispensed either active or placebo capsules according to a computer-generated list. The code was only revealed to the researchers once data collection was complete.

**Statistical methods**

Analysis was performed using the Statistical Package for the Social Sciences (SPSS), version 12.0.1 for Windows) and the last-observation-carried-forward (LOCF) method. Significance was set at 0.05 and all tests were two-tailed. Baseline data between groups were compared using the $\chi^2$-test for categorical variables and independent $t$-tests for psychometric tests. A repeated analysis of variance (ANOVA) with baseline rating scores entered as covariates was used to compare changes in psychopathology over time in the two groups. For psychometric data that were dichotomised, the $\chi^2$-test was used. Logistic regression was used to assess the independence of sub-scale scores.

**RESULTS**

**Participants**

A total of 49 patients entered the study, with 27 randomised to placebo and 22 to n-3 EFA (Fig. 1). No patient withdrew consent during the study period. Baseline socio-demographic characteristics were similar in both groups, except for marital status: more patients were married in the n-3 EFA than the placebo group (Table 1). With the exception of the BDI (mean placebo group score 32.22, mean n-3 EFA group score 38.41, F=2.7, P=0.04), the mean baseline scores for all psychometric instruments were similar for both groups. The mean scores for all psychometric instruments were well in excess of published normative data.

**Effects of n-3 EFA supplementation on psychopathology**

**Depression**

The baseline BDI scores were used as covariates in the subsequent statistical analysis (for all psychometric variables, baseline scores were entered as covariates). Table 2 shows the significant improvements in BDI scores at 12 weeks (P<0.004) in the n-3 EFA group. Moreover, more patients in the n-3 EFA group attained more than 50% (P<0.001) and 70% (P=0.001) reduction (response and remission respectively) in symptoms (Table 3). Similar data were observed for the HRSD. Figure 2 shows the improvements in depression scores at all measurements.

**Suicidal ideation and self-harm**

Taking a categorical value of zero to indicate no suicidal ideation and a value of 1 or more to indicate its presence (OAS–M suicidality sub-scale), more participants reported no suicidal ideation in the n-3 EFA group (n=14) than the placebo group.
392 assessed for eligibility
343 excluded
325 not meeting exclusion
145 first episode
239 schizophrenia
31 antipsychotics
42 frequent fish eaters
9 dietary supplements
4 organic brain injury
11 medical sequelae of self-harm
23 age <16 or >65
21 present substance misuse
18 refused to participate

49 randomised

27 placebo
6 discontinued treatment early
2 left district
2 lost to follow-up
2 admitted to psychiatric hospital
1 discontinued treatment late (refused to continue)
20 completed study

22 active
0 discontinued treatment early
3 discontinued treatment late
2 lost to follow-up
1 left district
19 completed study

Fig. 1 Trial CONSORT diagram.

(n=8, P=0.018), although the difference in the mean total OAS–M scores did not quite reach significance (P=0.057). Logistic regression indicated that neither the irritability and aggression sub-scales (both of which demonstrated no difference between the two groups) nor the depression scores had any effect on the suicidality sub-scale. There were no completed acts of suicide during the study period. Fourteen patients presented with self-harm episodes, 7 from the placebo and 7 from the n-3 EFA group (P=0.65); none was life threatening.

**Impulsivity/aggression**
No differences were observed at study end point in outcome measures (IMT/DMT, OAS–M total score) (Table 2).

**Perception of daily stresses**
Participants in the n-3 EFA group demonstrated a greater drop in mean PSS score at 12 weeks compared with those on placebo (P=0.021, Table 2). Significantly more patients attained a 50% drop ("response") in the n-3 EFA group at 12 weeks compared with placebo (P=0.006, Table 3). Significance was not reached for a 70% reduction (remission) – as only 5 patients (3 from the n-3 EFA group and 2 from the placebo group) attained this improvement in scores. Participants in the n-3 EFA group demonstrated a greater mean improvement in DHUS score at study end point compared with placebo (P=0.027). Any changes were independent of depression scores (logistic regression). Owing to the method used to calculate one mean DHUS score, response and remission data were not calculated. Figure 2 shows improvements in daily stress scores at all measurements.

**Safety, tolerability and adherence**
Four patients complained of mild gastric discomfort during the study and 3 of these were in the n-3 EFA group. A further 9 patients (2 placebo and 7 n-3 EFA) described a 'fish-like taste' from the capsules or 'fish-like breath' after consuming the capsules. No patients discontinued the study because of adverse events. There

<table>
<thead>
<tr>
<th>Variable</th>
<th>n-3 EFA group</th>
<th>Placebo group</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Co-varied mean</td>
<td>Change</td>
<td>95% CI</td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td>17.5</td>
<td>17.52</td>
<td>12.1–22.9</td>
</tr>
<tr>
<td>Hamilton Rating Scale for Depression</td>
<td>12.2</td>
<td>13.93</td>
<td>8.5–16.0</td>
</tr>
<tr>
<td>Perceived Stress Scale</td>
<td>19.9</td>
<td>9.65</td>
<td>16.4–23.4</td>
</tr>
<tr>
<td>Daily Hassles and Uplifts Scale</td>
<td>10.1</td>
<td>18.91</td>
<td>22.9–2.6</td>
</tr>
<tr>
<td>Overt Aggression Scale</td>
<td>10.3</td>
<td>10.57</td>
<td>6.7–13.8</td>
</tr>
<tr>
<td>Immediate Memory Task</td>
<td>36.2</td>
<td>1.41</td>
<td>32.8–39.6</td>
</tr>
<tr>
<td>Delayed Memory Task</td>
<td>37.0</td>
<td>–1.13</td>
<td>32.5–40.4</td>
</tr>
</tbody>
</table>

n-3 EFA, omega-3 essential fatty acids.
1. Positive scores for change indicate a benefit in mental health. The means have been adjusted by co-varying with the baseline mean for each group.

Table 2 Outcome variables for continuous data at study end point (12 weeks)
Improvement in mean psychometric score during the course of the study. Self-rated instruments were administered every 4 weeks and observer-rated instruments every 6 weeks. All patients were on antidepressants, with two-thirds taking prescribed psychotropic medication. Almost two-thirds of patients were taking antidepressants, with 8 in the placebo group (P = NS).

Psychotropic medication
Almost two-thirds of patients were taking prescribed psychotropic medication (excluding anti-psychotics at baseline (Table 1). All were on antidepressants, with two-thirds taking prescribed benzodiazepines.

Table 3  Response and remission data

<table>
<thead>
<tr>
<th>Variable</th>
<th>n-3 EFA group</th>
<th>Placebo group</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Absent</td>
<td>Present</td>
<td>Absent</td>
</tr>
<tr>
<td>Overt Aggression Scale – Suicidality &gt; 1</td>
<td>14</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Beck Depression Inventory</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response</td>
<td>9</td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td>Remission</td>
<td>9</td>
<td>13</td>
<td>23</td>
</tr>
<tr>
<td>Hamilton Rating Scale for Depression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response</td>
<td>10</td>
<td>11</td>
<td>14</td>
</tr>
<tr>
<td>Remission</td>
<td>13</td>
<td>9</td>
<td>23</td>
</tr>
<tr>
<td>Perceived Stress Scale</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response</td>
<td>13</td>
<td>9</td>
<td>25</td>
</tr>
<tr>
<td>Remission</td>
<td>19</td>
<td>3</td>
<td>25</td>
</tr>
</tbody>
</table>

1. Response, $\geq$ 50% reduction; remission $\geq$ 70% reduction.

During the trial period 5 participants in the placebo group were started on an antidepressant and 2 had the dose of their antidepressant increased, compared with 2 and 1 respectively in the n-3 EFA group (P = NS).

DISCUSSION
Since the mechanisms of action of the n-3 EFAs are poorly understood and they have been reported to improve a wide range of psychopathologies, from type I symptoms in schizophrenia (Peet & Horrobin, 2002) to aggression in healthy people (Hamazaki et al, 2001) we deliberately chose a population (recurrent self-harm) that would manifest a broad spectrum of psychological constructs and behaviours that could be measured for change reasonably quickly over time. Accordingly, our population, although not syndromally homogeneous, had scores on measures of depression, suicidality, aggression, daily stresses and impulsivity that indicated much greater in psychopathology than a healthy population. The high rates of Axis II disorder in the participants and levels of psychotropic prescribing also reflect this.

A principal driving force behind the improved functioning of the n-3 EFA group was the improvements in affect (Fig. 2). It is noteworthy that scores were continuing to diverge from those of the placebo group at study end point, suggesting that continued improvement might have occurred with a longer period of supplementation. This is supported by data from clinical trials in which supplementation was for 16 weeks (Hamazaki et al, 1996) or 20 weeks (Gesch et al, 2002). Although decreases in suicidality and daily stress perception were less dramatic than improvements in depression, they were independent of changes in depression score, indicating different factors...
mechanisms of n-3 EFA action. However, the lack of change in aggression or impulsivity measures suggests a degree of specificity, although there are robust data from randomised controlled trials that n-3 EFAs reduce aggression (Hamazaki et al, 1996; Gesch et al, 2002). Indeed, Zanarini & Frankenburg (2003) reported a significant change in aggression (OAS–M score) after supplementation for only 4 weeks in people (n=30) with borderline personality disorder. For a discussion on the possible mechanism of action of n-3 EFAs at the molecular level, we refer to our accompanying paper (Garland et al, 2007, this issue).

It is important to stress that treatment as usual was carried out by the patient’s treating psychiatrist/general practitioner, and the benefits seen were additional to normal care. The dosage of n-3 EFAs used is comparable to most other supplementation studies in psychiatric populations. The treatment was well tolerated, with only minimal side-effects, and for this population adherence was excellent (Spirito, 1996). The inclusion of a small amount of fish oil in the placebo preparation (with attendant ‘fish breath’ effects) made it unlikely that patients would have known which group they were in; this was confirmed at the end of the study (data not shown). Although 14 patients reported episodes of self-harm during the study, it was known a priori that the study was insufficiently powered to detect significant differences between groups.

Our accompanying paper details compositional deficits in a variety of n-3 EFAs in a different population of patients with self-harm (Garland et al, 2007, this issue). Taken together, the data suggest that reversal of an n-3 EFA deficit improves psychological status in such patients. The source of such deficits and the mechanism of improvement with supplementation remain a matter of speculation. Patients undergoing a period of life stress may be expected to consume different diets (Oliver & Wardle, 1999) and this may apply to fish consumption; alternatively deficits may be engendered by genomic or proteomic differences in metabolic pathways. We will examine this in future studies using blood samples taken from our patients.

Clearly there is now a mandate to proceed with a large-scale multicentre study with the primary end points of recurrence of self-harm and completed suicide, the principal targets of intervention.

ACKNOWLEDGEMENTS

BRIAN HALLAHAN, MD, MRCPsych, Department of Psychiatry, Beaumont Hospital and The Royal College of Surgeons in Ireland; JOSPEH R. HIBBELN, MD, Laboratory of Membrane Biochemistry and Biophysics, National Institute on Alcohol Abuse and Alcoholism, Rockville, Maryland, USA; JOHN M. DAVIS, MD, Institute of Psychiatry, University of Illinois at Chicago, Illinois, USA; MALCOLM R. GARLAND, MD, MRCPsych, MRCPI, Department of Psychiatry, Beaumont Hospital and the Royal College of Surgeons in Ireland, Ireland.

Correspondence: Dr Malcolm R. Garland, St Ita’s Hospital, Portrane, County Dublin, Ireland. Email: mgarland@ireland.com

(First received 31 January 2006, final revision 29 August 2006, accepted 4 October 2006)

REFERENCES


Treatment of social phobia: randomised trial of internet-delivered cognitive–behavioural therapy with telephone support

PER CARLBRING, MAGDALENA GUNNARSDOTTIR, LINDA HEDENSJO, GERHARD ANDERSSON, LISA EKSELIUS and TOMAS FURMARK

Background Although effective therapies for social phobia exist, many individuals refrain from seeking treatment owing to the embarrassment associated with help-seeking. Internet-based cognitive–behavioural self-help can be an alternative, but adherence is a problem.

Aims To evaluate a 9-week programme based treatment supplemented by short, weekly telephone calls. As evidenced by several trials there are effective psychosocial treatments for social phobia (Rodebaugh et al, 2004). However, far from all of those with this condition seek treatment (Baldwin & Buist, 2004). Apart from a shortage of skilled therapists, long waiting lists and high costs, a major problem is that those with social phobia may not seek therapy because of the fear of embarrassment associated with help-seeking (King & Poulos, 1998; Newman et al, 2003). Consequently, a challenge is to increase the accessibility of evidence-based psychological treatment. A recent approach is internet-based self-help with minimal therapist contact by email (Andersson et al, 2006). Although the results are promising, there is a problem with treatment adherence. In this study we attempted to increase adherence by adding short, weekly telephone calls to the internet-based self-help programme.

Method In a randomised controlled trial the effects of internet-based cognitive–behavioural therapy in the treatment group (n=29) were compared with a waiting-list control group (n=28).

Results Compared with the control group the treated participants experienced greater reductions on measures of general and social anxiety, avoidance and depression. Adherence to treatment was high, with 93% finishing the complete treatment package. One year later all improvements were maintained.

Conclusions This study provides evidence to support the use of internet-based treatment supplemented by short, weekly telephone calls.

Declaration of interest None. Funding detailed in Acknowledgements.

As evidenced by several trials there are effective psychosocial treatments for social phobia (Rodebaugh et al, 2004). However, far from all of those with this condition seek treatment (Baldwin & Buist, 2004). Apart from a shortage of skilled therapists, long waiting lists and high costs, a major problem is that those with social phobia may not seek therapy because of the fear of embarrassment associated with help-seeking (King & Poulos, 1998; Newman et al, 2003). Consequently, a challenge is to increase the accessibility of evidence-based psychological treatment. A recent approach is internet-based self-help with minimal therapist contact by email (Andersson et al, 2006). Although the results are promising, there is a problem with treatment adherence. In this study we attempted to increase adherence by adding short, weekly telephone calls to the internet-based self-help programme.

METHOD

Participants were selected by means of a computerised screening interview consisting of the highly specific and sensitive Social Phobia Screening Questionnaire (SPSQ; Furmark et al, 1999), the self-rated version of the Montgomery–Åsberg Depression Rating Scale (MADRS–S; Svanborg & Åsberg, 1994) and ten additional questions regarding current and past treatment. To be included in the study, participants had to meet the following ten criteria:

(a) fulfil the DSM–IV (American Psychiatric Association, 1994) criteria for social phobia according to the SPSQ;
(b) have a total score of below 31 on the MADRS–S depression scale and a score of less than 4 on the suicide item of this scale;
(c) agree to undergo no other psychological treatment for the duration of the study, and have no history of earlier cognitive–behavioural therapy;
(d) if taking prescribed drugs for anxiety or depression, the dosage had to be constant throughout the study;
(e) have access to a computer with internet connection;
(f) be at least 18 years old;
(g) live in Sweden;
(h) be able to speak to the therapists on the telephone on a weekly basis.

Outcome measures The following social anxiety scales constituted the primary outcome measures: the Liebowitz Social Anxiety Scale self-report version (LSAS–SR; Liebowitz, 1987), the Social Phobia Scale (SPS) and Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998) and the Social Phobia Screening Questionnaire (SPSQ; Furmark et al, 1999). In addition, the following secondary measures were used to assess general anxiety, depression and quality of life: the Beck Anxiety Inventory (BAI; Beck et al, 1988),

Outcomes were assessed at baseline, 3 months after randomisation and post-treatment (9 weeks post-randomisation).

In summary, a randomised controlled trial was conducted to evaluate a 9-week programme based treatment supplemented by weekly telephone calls. The results showed that the treatment group experienced greater reductions on measures of general and social anxiety, avoidance and depression compared to the waiting-list control group. The study highlights the potential of internet-based self-help as a viable alternative to traditional therapy, which can be further improved by the addition of telephone support to enhance adherence.

Funding detailed in Acknowledgements.
the MADRS–S (Svanborg & Åberg, 1994), and the Quality of Life Inventory (QoLI; Frisch et al., 1992). The outcome measures were administered after all inclusion criteria were met, i.e. after the SCID interview. Internet administration of questionnaires has generally resulted in adequate psychometric properties (see Carlbring et al., 2007).

**Intervention**

Whereas those in the control group remain on a waiting list and receive no treatment, those in the treatment group receive internet-administered self-help including minimal therapist contact via email supplemented with short weekly telephone calls as outlined below.

**Treatment**

The treatment was based on established cognitive–behavioural methods as described in self-help books (e.g. Rapee, 1998; Antony & Swinson, 2000). The text, consisting of 186 pages, was taken from an existing manual (Furmark et al., 2006), divided into nine modules and adapted for the internet. Each module included information, exercises and an interactive quiz, and ended with three to eight essay questions. Participants were asked to explain in their own words the most important sections of the module they had just completed, provide thought records, and describe their experience with and outcome of their exposure exercises. The questions were intended to promote learning and to enable the online therapists to assess whether the participants had assimilated the material. For each module participants were required to post at least one message in an online discussion group about a predetermined topic.

Feedback on homework assignments was usually given within 24 h after participants had sent their answers by email. On the basis of these emails, an assessment was made of whether the participant was ready to continue; if so, the password to the next module was sent. If not, the participant received instructions on what needed to be completed before proceeding to the next module.

**Telephone calls**

One weekly telephone call was made by the therapists to each participant in the treatment group. The purpose was to provide positive feedback and to answer any questions the participant might have regarding the modules. All conversations were timed, and each of the nine calls lasted an average of 10.5 min (s.d. = 3.6).

**Therapists**

The therapists were two students completing their last semester of the Master’s degree programme to become clinical psychologists. The mean total time per week spent on each participant was approximately 22 min, including telephone calls, administration, and reading and responding to emails. Hence, the total human contact time per participant including screening was over 2.5 h.

**Statistical analysis**

Significance testing of group differences in demographic data and pre-treatment measures was conducted with χ² and t-tests. Participants’ scores before and after treatment were analysed using two-way analysis of variance with repeated measures. These were followed by t-tests with Bonferroni-corrected P values, set at 0.0125. This limit was obtained by dividing the traditional alpha level with the maximum number of individual group comparisons (i.e. 4). Effect sizes (Cohen’s d) were calculated both within and between groups, and all calculations were based on the pooled standard deviation.

**RESULTS**

The flow of participants through the trial is shown in Fig. 1 and the characteristics of the sample are given in Table 1.

**Attrition**

Two participants, one in each condition, were excluded from the analysis since they started other treatment during the period.

---

**Table I** Demographic description of the participants

<table>
<thead>
<tr>
<th></th>
<th>Treatment group (n = 29)</th>
<th>Control group (n = 28)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>17 (59)</td>
<td>20 (71)</td>
</tr>
<tr>
<td>Male</td>
<td>12 (41)</td>
<td>8 (29)</td>
</tr>
<tr>
<td>Age, years</td>
<td>Mean (s.d.)</td>
<td>Range</td>
</tr>
<tr>
<td></td>
<td>32.4 (9.1)</td>
<td>19–52</td>
</tr>
<tr>
<td></td>
<td>32.9 (9.2)</td>
<td>22–51</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married/living together</td>
<td>11 (38)</td>
<td>13 (46)</td>
</tr>
<tr>
<td>Married/living together with children</td>
<td>5 (17)</td>
<td>4 (14)</td>
</tr>
<tr>
<td>Living apart with children</td>
<td>1 (3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Living apart without children</td>
<td>2 (7)</td>
<td>2 (7)</td>
</tr>
<tr>
<td>Single with children</td>
<td>1 (3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Single without children</td>
<td>9 (31)</td>
<td>9 (32)</td>
</tr>
<tr>
<td>Education, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Elementary education</td>
<td>2 (7)</td>
<td>2 (7)</td>
</tr>
<tr>
<td>High school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not completed</td>
<td>0 (0)</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Completed</td>
<td>8 (28)</td>
<td>5 (18)</td>
</tr>
<tr>
<td>Community college</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completed</td>
<td>1 (3)</td>
<td>1 (4)</td>
</tr>
<tr>
<td>College/university</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not completed</td>
<td>10 (34)</td>
<td>10 (34)</td>
</tr>
<tr>
<td>Completed</td>
<td>8 (28)</td>
<td>9 (32)</td>
</tr>
<tr>
<td>Self-rated computer experience, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Far above average</td>
<td>10 (34)</td>
<td>9 (32)</td>
</tr>
<tr>
<td>Above average</td>
<td>7 (24)</td>
<td>6 (21)</td>
</tr>
<tr>
<td>Average</td>
<td>10 (34)</td>
<td>9 (32)</td>
</tr>
<tr>
<td>Below average</td>
<td>2 (7)</td>
<td>4 (14)</td>
</tr>
</tbody>
</table>

I. No significant differences existed between the groups according to χ² tests.
Pre-treatment differences

The two groups did not differ significantly on any of the measures at the pre-treatment assessment ($t_{4}=0.08–0.76$, $P=0.94–0.45$).

Primary outcome measures

Significant group by time interactions were obtained for all primary measures (Table 2). For all scales post hoc t-tests with Bonferroni-corrected $P$ values indicated that the intervention group had improved significantly between the pre-treatment and post-treatment assessments ($t_{5}=6.3–7.2$, $P<0.001$), whereas the control group had not ($t_{2}=0.4–1.2$, $P>0.23$). Moreover, following treatment the treated group had lower social anxiety levels on all scales compared with controls ($t_{5}=3.6–5.1$, $P<0.001$).

Secondary outcome measures

Significant group by time interactions were obtained for scores on the MADRS–S and BAI (Table 3). The QoLI only showed a trend ($P=0.08$). For all secondary scales, post-hoc t-tests with Bonferroni-corrected $P$ values indicated that the intervention group had improved significantly between the pre-treatment and post-treatment assessments ($t_{5}=3.2–4.6$, $P<0.004$), whereas the control group had not ($t_{2}=0.6–1.1$, $P>0.29$). Moreover, following treatment, the intervention group had lower distress levels on two of the scales (MADRS–S and BAI) compared with the control group ($t_{5}=2.8–3.4$, $P<0.007$). However, we found no significant post-treatment difference in QoLI score between the groups ($t_{5}=1.6$, $P=0.12$).

Effect sizes

The mean within-group effect size was high at $d=0.95$. The between-group effect size varied markedly across the different measures: the highest value was found on the SPSQ ($d=1.31$) whereas the lowest was found for QoLI ($d=0.39$). The mean between-group effect size across all measures was $d=1.00$.

One-year follow-up

Of the 29 people in the treatment group, 28 returned the 1-year follow-up questionnaires; hence intention-to-treat analysis was again used. Paired $t$-tests showed that there were significant differences between pre-treatment and follow-up scores ($t_{24}=3.4–8.1$, $P<0.003$), but almost no
Table 2 Social phobia: main and interaction effects and pooled effect sizes for each group

<table>
<thead>
<tr>
<th></th>
<th>Pre-treatment Mean (s.d.)</th>
<th>Post-treatment Mean (s.d.)</th>
<th>Main effect</th>
<th>Interaction effect (time x group)</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time</td>
<td></td>
<td>Group</td>
<td></td>
<td>Within group</td>
</tr>
<tr>
<td>Liebowitz Social Anxiety Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fear/anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment group</td>
<td>36.0 (11.7)</td>
<td>24.2 (12.0)</td>
<td>16.8***</td>
<td>3.1</td>
<td>32.2***</td>
</tr>
<tr>
<td>Control group</td>
<td>34.2 (10.6)</td>
<td>36.1 (12.3)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avoidance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment group</td>
<td>35.2 (12.9)</td>
<td>21.6 (12.8)</td>
<td>34.2***</td>
<td>3.0</td>
<td>29.2***</td>
</tr>
<tr>
<td>Control group</td>
<td>33.8 (11.0)</td>
<td>33.3 (11.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Phobia Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment group</td>
<td>36.2 (15.2)</td>
<td>20.0 (15.0)</td>
<td>28.3***</td>
<td>6.1*</td>
<td>27.8***</td>
</tr>
<tr>
<td>Control group</td>
<td>37.8 (16.5)</td>
<td>37.7 (16.4)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Interaction Anxiety Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment</td>
<td>41.3 (13.2)</td>
<td>27.1 (11.1)</td>
<td>23.1***</td>
<td>8.8**</td>
<td>28.3***</td>
</tr>
<tr>
<td>Control group</td>
<td>42.9 (12.1)</td>
<td>43.6 (14.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social Phobia Screening Questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment group</td>
<td>29.7 (7.8)</td>
<td>20.3 (9.1)</td>
<td>27.7***</td>
<td>9.9***</td>
<td>41.2***</td>
</tr>
<tr>
<td>Control group</td>
<td>31.4 (9.4)</td>
<td>32.3 (8.9)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*P < 0.05, **P < 0.0125, ***P < 0.001.
1. Cohen’s d.
2. Higher scores indicate a better quality of life.

Table 3 Depression, anxiety level and quality of life: main and interaction effects and pooled effect sizes for each group

<table>
<thead>
<tr>
<th></th>
<th>Pre-treatment Mean (s.d.)</th>
<th>Post-treatment Mean (s.d.)</th>
<th>Main effect</th>
<th>Interaction effect (time x group)</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time</td>
<td></td>
<td>Group</td>
<td></td>
<td>Within group</td>
</tr>
<tr>
<td>Montgomery–Åsberg Depression Rating Scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment group</td>
<td>13.4 (8.4)</td>
<td>8.5 (5.9)</td>
<td>5.7*</td>
<td>3.4</td>
<td>12.4***</td>
</tr>
<tr>
<td>Control group</td>
<td>13.5 (6.0)</td>
<td>14.5 (7.2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beck Anxiety Inventory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment group</td>
<td>14.5 (8.1)</td>
<td>8.2 (7.9)</td>
<td>15.4***</td>
<td>2.8</td>
<td>10.2**</td>
</tr>
<tr>
<td>Control group</td>
<td>15.1 (8.8)</td>
<td>14.5 (9.0)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of Life Inventory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summary score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment group</td>
<td>0.7 (2.0)</td>
<td>1.4 (1.8)</td>
<td>9.9**</td>
<td>1.16</td>
<td>3.2</td>
</tr>
<tr>
<td>Control group</td>
<td>0.5 (1.7)</td>
<td>0.7 (1.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*P < 0.05, **P < 0.0125, ***P < 0.001.
1. Cohen’s d.
2. Higher scores indicate a better quality of life.

robust post-treatment $\nu$. follow-up changes ($t_{25}=0.2–1.9, P<0.86–0.07$). The only significant post-treatment $\nu$. follow-up change was in QoLI ($t_{25}=2.55; P=.017$). The following 1-year results were observed:

- LSAS–SR, sub-scale fear, mean 22.31 (s.d.=11.43); LSAS sub-scale avoidance, mean 20.55 (s.d.=11.95); SPS, mean 20.28 (s.d.=14.60); SIAS, mean 25.69 (s.d.=10.63); SPSQ total score, mean 18.00 (s.d.=10.12); MADRS–S, mean 7.93 (s.d.=7.75); BAI, mean 7.62 (s.d.=8.93); QoLI, mean 1.94 (s.d.=1.67). In sum, the average effect size at follow-up was $d=1.02$. There was no correlation between
number of postings on the online discussion and change scores at post-treatment or follow-up assessment (all $r < 0.17$ and all $P > 0.40$).

**DISCUSSION**

Participants receiving the intervention improved significantly on all the measures used, whereas those on the waiting list control group did not. Specifically, the treated participants achieved significant improvement on measures of social anxiety, fear, avoidance, depression and general anxiety. The differences in quality of life were marked but not significant, and this may represent a type II error. Apart from power issues, a reason could be that the treatment period was too short to have an impact on this measure, which has broad items such as 'number of children'. In fact, at the 1-year follow-up the quality of life measure had caught up and was significant. Unfortunately, since people on the waiting list received treatment before the follow-up data were collected, there is no between-group comparison at the 1-year follow-up.

**Effect size**

The treatment had a substantial within-group effect size (Cohen’s $d = 0.95$), which should be compared with the within-group effect size reported in a meta-analysis by Taylor (1996) for placebo ($d = 0.48$), exposure alone ($d = 0.81$), cognitive therapy ($d = 0.63$), social skills training ($d = 0.65$), and cognitive and exposure therapy combined ($d = 1.06$).

**Adherence**

Although self-administered treatments for various problems have shown promising results in many studies (Carlbring & Andersson, 2006), a crucial problem is the low adherence to treatment. This study added weekly telephone calls, which resulted in a considerably higher proportion of participants finishing the entire treatment package within the 9-week time frame, compared with a previous study (Andersson et al., 2006) conducted without telephone support (93% vs. 62%). However, direct comparisons are needed to draw firm conclusions regarding the relative value of whether therapist interaction over the telephone improves retention and outcomes.

**Limitations**

One of the advantages of internet-based therapy is the possibility of treating people who would not otherwise seek treatment. Asking participants to come to a clinical selection interview might induce a self-selection bias for people with less severe problems. Our study was designed to target anyone with social phobia, whether they were able to travel or not; we therefore decided to administer the clinical interview over the telephone, which might have compromised diagnostic reliability. Because the research staff never met the participants in person, there was a risk of including those with extreme suicidal tendencies. To minimise this risk we excluded people who, according to their MADRS-S, responses were suicidal. In theory this might have led to a sample of people who were less depressed than participants in other studies. However, the results on the measures are comparable with those reported elsewhere for the target population (Orsillo, 2001). Nevertheless, it is still uncertain how the treatment would affect a more severely depressed group.

Another caveat with this study is that the educational level of the participants was high. One in three Swedish adults aged 25–64 years has some form of post-secondary education (Statistics Sweden, 2003). That is considerably lower than in our study sample, which raises the question of how well the treatment would work with individuals with lower levels of education. Also, as the sample was selected from individuals who had expressed an interest in an internet-administered self-help programme, it is possible that selection biases yield a more effective result for this treatment compared with standard live treatment. Finally, a major weakness is the sole reliance on self-report measures. A clinical global impression and a behavioural test including psychophysiological measures would have strengthened the results.

**Future research**

As we did not include a comparison treatment, specificity of the findings cannot be assured. Consequently, future studies should investigate the issue of specificity of internet-based self-help interventions, the role of community online support and the non-specifics of therapist contact that are likely to be present in both telephone and internet consultations. Furthermore, larger studies are needed to allow an examination of individual characteristics and treatment response. Additionally, comparisons with standardised face-to-face therapy are imperative (compare with Carlbring et al., 2003). Dismantling studies are strongly encouraged in order to evaluate the cost-benefit of brief or more intensive combined treatments (e.g. internet plus live therapist sessions in severe cases).

**ACKNOWLEDGEMENTS**

This study was funded in part by a grant from the Swedish Research Council and the Söderström Königska Foundation.

**REFERENCES**


Risk of hip fracture in patients with a history of schizophrenia

LOUISE HOWARD, GRAHAM KIRKWOOD and MORVEN LEESE

Background There is evidence of an association between decreased bone mineral density, schizophrenia, and prolactin-raising antipsychotic medication. However, it is not known whether this is clinically significant.

Aims To investigate whether patients with a history of schizophrenia are at increased risk of hip fracture.

Method In a case–control study, we compared cases of ‘hip fracture’ on the General Practice Research Database (n=16 341) with matched controls (n=29 889).

Results Hip fracture was associated with schizophrenia (OR=1.73; 95% CI 1.32–2.10), and prolactin-raising antipsychotics (OR=2.6; 95% CI 2.43–2.78), in the univariate analysis. In the multivariate analysis, prolactin-raising antipsychotics were independently associated with hip fracture but schizophrenia was not. A significant interaction between gender and antipsychotics was found in the association with hip fracture (P=0.042; OR=2.12 (95% CI 1.73–2.59) for men, OR=1.93 (95% CI 1.78–2.10) for women.

Conclusions The association between prolactin-raising antipsychotic medication and hip fracture may have serious implications for public health. Mental health service patients may require preventive measures including dietary and lifestyle advice.

Declaration of interest None.

Funding detailed in Acknowledgements.

A growing number of recent reports have suggested that there is an association between decreased bone mineral density and schizophrenia (Halbreich et al, 1995; Abraham et al, 2003), particularly in patients treated with psychotropic medication (Halbreich & Palter, 1996; O’Keane & Meaney, 2005). However, these studies have not examined whether this is clinically significant, leading to an increased risk of fractures. The prevalence of functional psychoses is 4 per 1000 (Jenkins et al, 1997), with approximately half of that figure relating to schizophrenia. The identification of schizophrenia as a risk factor for osteoporosis would therefore have important public health implications.

Hip fracture is the most important fracture in terms of patient morbidity and mortality and for utilisation of health service resources (Cumming et al, 1997). As fracture is common in the older population, a small increase in the risk of fracture associated with psychotic disorders could have a considerable public health effect. We therefore chose to investigate whether there was an association between schizophrenia and hip fractures, using a case–control study design with data from a UK primary care data-set, the General Practice Research Database (GPRD). Our hypothesis was that patients with a history of schizophrenia would have a significantly increased risk of antipsychotic-induced osteoporotic hip fractures compared with a control group matched for age and general practice.

METHOD

Data source

The GPRD was set up in the UK in 1987 and contains the computerised medical records of approximately 5% of the UK population in primary care (Walley & Mantgani, 1997). Data recorded include prescription details, clinical events, preventive care provided, specialist referrals, hospital admissions and their major outcomes. The data collected are audited regularly and the participating general practices subject to a number of quality checks, including internal validation by cross-checking within practices and by comparisons with national statistics (Walley & Mantgani, 1997). Only practices that comply with this quality control (i.e. are ‘up-to-research standard’) are retained within the database. The data are representative of the general population (Walley & Mantgani, 1997), although there is a bias towards larger group practices. Diagnoses of schizophrenia recorded on the GPRD have been validated in several studies (Nazareth et al, 1993; Howard et al, 2002) and hip fracture recording is reported to be particularly complete (van Staa et al, 2001).

Case–control analysis

Stata version 8.2 for Windows was used for statistical analysis. All patients registered on the GPRD between 1 August 1987 and 22 November 1999 with a recorded diagnosis of fractured neck of femur or ‘hip fracture’ were identified, and designated as cases. Two controls per case were identified, matching on age, gender, general practice and duration of available GPRD data. For statistical efficiency a 1:1 ratio of cases to controls is ideal when the number of cases can be chosen. In our study, since the number of cases was limited, the number of controls was increased to two per case in order to achieve adequate power. Each case was assigned a date of diagnosis, defined as the date of the first hip fracture, and matching control individuals were assigned an identical ‘pseudo’ date of diagnosis. Only records that were up to research standard were used.

Variables included the patients’ medical and psychiatric history, medication history and demographic details, as well as lifestyle factors (alcohol consumption, smoking and body mass index). All recorded diagnoses of schizophrenia were extracted and recorded for each case and control on an ever/never basis. Ever having had a prescription for a neuroleptic drug prior to or on the day of the first fracture was extracted and recorded. Where comorbid disorders were examined all disorders under the main relevant ICD–9 heading were included e.g. ‘intestinal diseases’ includes all diseases under this heading in ICD–9 (World Health Organization, 1978). To
enable all the data to be used, a category 'missing' was created for each variable where needed. Imputation would have added no extra information, since all the variables available to make imputations were already included in the analysis.

Variables that had been previously identified in the research literature as being significantly associated with hip fracture were examined in univariate analyses and those that were significantly associated in this analysis were selected for multivariate analysis. Conditional logistic regression was used to calculate an unadjusted odds ratio for each selected variable. The selected variables were added one by one into bivariate models to identify potential confounders with the schizophrenia diagnosis variable and to see how much, if at all, they reduced or increased the odds ratio obtained with the schizophrenia diagnosis variable alone. A multivariable model was fitted using conditional logistic regression in a forward stepwise process using likelihood ratio tests. Variables that had become non-significant were removed and again the fit of the model was tested. By including interaction terms, gender and age at the time of the first fracture were tested to assess for modification of the effect of the schizophrenia diagnosis and of ever having had a prescription for neuroleptic medication on the incidence of hip fracture.

Ethical approval was granted by the Scientific Advisory and Ethical Group at the Medicines Control Agency, who are responsible for ethical issues for all projects using the GPRD.

**RESULTS**

There were 16,341 cases of hip fracture and 29,889 controls. There were two controls per case for 13,548 of the cases and one control per case for the remaining 2793. The mean age of the cases and controls at the time of first fracture was 79 years (s.d. = 12), and 12,854 (79%) of the cases were female. Variables identified as having an association with hip fracture in the univariate analysis included cerebrovascular disease (OR = 1.89, 95% CI 1.76–2.03), blood disorders (OR = 1.85, 95% CI 1.73–1.97), intestinal disorders (OR = 1.75, 95% CI 1.69–1.86), eye disorders (OR = 1.11, 95% CI 1.06–1.17), ear disorders (OR = 1.02, 95% CI 0.97–1.08), urinary disorders (OR = 1.50, 95% CI 1.42–1.57), laxatives (OR = 1.97, 95% CI 1.89–2.07), opioid analgesics (OR = 1.94, 95% CI 1.81–2.07), non-opioid analgesics (OR = 1.85, 95% CI 1.77–1.93), inhaled corticosteroids (OR = 1.26, 95% CI 1.17–1.36) and injected corticosteroids (OR = 1.07, 95% CI 0.96–1.19). More details of variables of particular interest to this study, including psychotropic medications, are presented in Table 1. Medication with heparin was considered for inclusion in the analysis, but this was ruled out as only 8 cases and 5 controls were exposed to this drug. Fitting bivariate models for each of these variables along with the schizophrenia variable identified only one having a substantial effect on the odds ratio attached to schizophrenia diagnosis, namely having ever had a prescription for

<table>
<thead>
<tr>
<th>Exposure</th>
<th>Cases (n = 16,341)</th>
<th>Controls (n = 29,889)</th>
<th>Univariate analysis</th>
<th>Bivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>OR (95% CI)</td>
<td>OR for each variable adjusting for schizophrenia</td>
</tr>
<tr>
<td>Schizophrenia diagnosis</td>
<td>100 (0.61)</td>
<td>110 (0.37)</td>
<td>1.73 (1.32–2.28)</td>
<td>2.60</td>
</tr>
<tr>
<td>Any neuroleptic</td>
<td>2246 (13.74)</td>
<td>1779 (5.95)</td>
<td>2.60 (2.43–2.78)</td>
<td>2.60 (1.00)</td>
</tr>
<tr>
<td>Any SSR1</td>
<td>955 (5.84)</td>
<td>892 (2.98)</td>
<td>2.01 (1.82–2.21)</td>
<td>2.01</td>
</tr>
<tr>
<td>Any anticonvulsant</td>
<td>869 (5.32)</td>
<td>805 (2.69)</td>
<td>2.01 (1.82–2.22)</td>
<td>2.01</td>
</tr>
<tr>
<td>Any tricyclic</td>
<td>2905 (17.78)</td>
<td>3259 (10.90)</td>
<td>1.78 (1.69–1.89)</td>
<td>1.78 (1.69)</td>
</tr>
<tr>
<td>Any hypnotic</td>
<td>4195 (25.67)</td>
<td>5149 (17.23)</td>
<td>1.68 (1.60–1.77)</td>
<td>1.68</td>
</tr>
<tr>
<td>Alcohol intake over recommended limit4</td>
<td>243 (1.49)</td>
<td>288 (0.96)</td>
<td>1.60 (1.33–1.92)</td>
<td>1.60 (1.73)</td>
</tr>
</tbody>
</table>

1. Any of chlorpromazine, thioridazine, droperidol, flupentixol, fluphenazine, haloperidol, pimozide, promazine, sulpiride, trifluoperazine, zuclopenthixol acetate and dehydrochloride as well as depot versions of flupentixol, fluphenazine, haloperidol, zuclopenthixol and pipotiazine.
2. Selective serotonin reuptake inhibitor: any of citalopram, fluoxetine, fluvoxamine, paroxetine and sertraline.
3. Any of carbamazepine, phenytoin, sodium valproate, lamotrigine and gabapentin.
4. Alcohol level compared with intake within recommended limits, defined as no more than 21 units a week for men and 14 units a week for women.
5. Smoking status compared with non-smoker.
6. Compared with normal body mass index defined as 18.5–24.9 kg/m².
There was also a significant interaction between age at the time of the first fracture (categorised) and having ever had a prescription for neuroleptic medication ($\chi^2=28.27$, d.f.=3, $P<0.001$, likelihood ratio test). The interaction was most marked for those aged under 85 years ($n=31,071$; OR for neuroleptic medication 2.29, 95% CI 2.06–2.53) compared with those aged 85 years and over ($n=15,113$; OR for neuroleptic medication 1.66, 95% CI 1.49–1.85). Since there was an interaction for both age and gender, the model was also fitted separately with the data-set stratified by both these variables into women under 85 years old ($n=23,289$; OR for neuroleptic medication 2.31, 95% CI 2.06–2.58); women 85 years or older ($n=13,027$; OR=1.65, 95% CI 1.47–1.86); men under 85 years old ($n=77,888$; OR=2.42, 95% CI 1.89–3.09); and men 85 years old or older ($n=20,989$; OR=2.12, 95% CI 1.51–2.99). Odds ratios were calculated for each of the individual neuroleptic medications which went to make up the combined neuroleptic variable, and are presented in Table 4. Also included here are figures for prochlorperazine, which was not included as part of the combined neuroleptic variable as it was felt that this might have been used largely as an anti-emetic. Atypical antipsychotic medications (which were only prescribed in the last few years of this data-set), were also considered separately, as there were too few prescribed to include in a combined antipsychotic variable, and odds ratios for these are also presented.

**DISCUSSION**

We found an independent significant association between hip fracture and prolactin-raising antipsychotic medications used prior to the index fracture. The relationship between schizophrenia and hip fracture was confounded by neuroleptics, and the relationship therefore appears to be owing to the effect of neuroleptics rather than diagnosis *per se*. There is increasing evidence that long-term antipsychotic-induced hyperprolactinaemia is associated with bone mineral density loss which appears to be mediated by secondary hypogonadism (Bilici et al, 2002; O’Keane & Meaney, 2005). However, ours is the first study, to our knowledge, to demonstrate an association between antipsychotic medication and a clinically relevant outcome of osteoporosis, i.e. fracture

---

**Table 2 Adjusted odds ratios for hip fractures in women: General Practice Research Database 1987–1999 (n=36,330)**

<table>
<thead>
<tr>
<th>Exposure</th>
<th>OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schizophrenia diagnosis</td>
<td>1.01</td>
<td>0.72–1.40</td>
</tr>
<tr>
<td>Any neuroleptic 1</td>
<td>1.93</td>
<td>1.78–2.10</td>
</tr>
<tr>
<td>Body mass index 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>2.17</td>
<td>1.78–2.64</td>
</tr>
<tr>
<td>Overweight</td>
<td>0.55</td>
<td>0.50–0.62</td>
</tr>
<tr>
<td>Obese</td>
<td>0.35</td>
<td>0.30–0.41</td>
</tr>
<tr>
<td>Any SSR1</td>
<td>1.24</td>
<td>1.10–1.40</td>
</tr>
<tr>
<td>Any laxative</td>
<td>1.36</td>
<td>1.29–1.44</td>
</tr>
<tr>
<td>Any anticonvulsant</td>
<td>1.39</td>
<td>1.23–1.57</td>
</tr>
<tr>
<td>Opioid analgesics</td>
<td>1.38</td>
<td>1.27–1.50</td>
</tr>
<tr>
<td>Cerebrovascular disease</td>
<td>1.48</td>
<td>1.36–1.62</td>
</tr>
<tr>
<td>Blood diseases</td>
<td>1.51</td>
<td>1.40–1.63</td>
</tr>
<tr>
<td>Non-opioid analgesics</td>
<td>1.47</td>
<td>1.39–1.55</td>
</tr>
<tr>
<td>Any tricyclic</td>
<td>1.16</td>
<td>1.09–1.25</td>
</tr>
<tr>
<td>Any hypnotic</td>
<td>1.16</td>
<td>1.10–1.23</td>
</tr>
<tr>
<td>Diseases of urinary system</td>
<td>1.17</td>
<td>1.10–1.24</td>
</tr>
<tr>
<td>Smoking status 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Current smoker</td>
<td>1.28</td>
<td>1.14–1.43</td>
</tr>
<tr>
<td>Ex-smoker</td>
<td>1.00</td>
<td>0.85–1.19</td>
</tr>
<tr>
<td>Hormone replacement therapy</td>
<td>0.62</td>
<td>0.51–0.75</td>
</tr>
<tr>
<td>Inhaled corticosteroids</td>
<td>1.11</td>
<td>1.01–1.23</td>
</tr>
<tr>
<td>Hypertensive disease</td>
<td>0.87</td>
<td>0.82–0.93</td>
</tr>
<tr>
<td>Ear diseases</td>
<td>0.91</td>
<td>0.85–0.96</td>
</tr>
</tbody>
</table>

SSRI, selective serotonin reuptake inhibitor.
1. Neuroleptic medication prescribed before or on day of fracture.
2. Compared with normal body mass index defined as 18.5–24.9 kg/m$^2$.
3. Smoking status compared with non-smoker.
of the neck of femur or hip fracture. A history of neuroleptic use was found to be the most significant predictor of an increased risk of hip fracture, with the exception of being underweight.

A number of recent studies have indicated that low bone mineral density is highly prevalent in people with a chronic psychiatric disorder treated with prolactin-raising antipsychotic medication (Bilici et al., 2002; Liu-Seifert et al., 2004). Meaney et al. (2004) also reported that higher doses of potent typical and atypical antipsychotic medications that block dopamine D2 receptors were associated with increased rates of both hyperprolactinaemia and bone mineral density loss. However, possibly because of a lack of research into the secondary consequences of antipsychotic-induced hyperprolactinaemia, the World Health Organization does not include antipsychotic drugs in its list of prescribed drugs associated with the development of osteoporosis.

Although postmenopausal women are generally most at risk of osteoporosis, possibly owing to low serum oestradiol concentrations (Cummings et al., 1998), we found a higher risk of fracture associated with neuroleptic medication in men. This result is in agreement with several studies of psychiatric patients, which found significantly lower bone mineral density in men than women associated with neuroleptic use (Halbreich et al., 1995; Hummer et al., 2005). These gender differences may be owing to the age differences in onset of schizophrenia (Hafner et al., 1998): men have an age at onset approximately 5 years younger than that in women, and illness-related factors including medication will therefore have had a longer-lasting impact on male patients. An alternative explanation suggested by Hummer & Huber (2004) is that women with schizophrenia take better care of themselves with regard to adequate nutrition and exercise than men and therefore have less osteoporosis. Unfortunately the data are not available here to test either hypothesis.

There are a number of other known risk factors for osteoporosis among patients with schizophrenia which may be acting as confounders here, including inadequate exercise and exposure to sunshine, poor nutrition, cigarette smoking and polydipsia (Naidoo et al., 2003). Of these, only smoking could be controlled for in this analysis and was found to be a significant factor, but the association was not as strong as with neuroleptic medication. Other mechanisms may also be relevant in causing hip fracture: for example, neuroleptic medications are known to cause sedation, orotostatic hypotension and extrapyramidal side-effects, which may predispose some patients to these treatments to falls (Misra et al., 2004).

We used the GPRD, a large UK primary care data-set, which provided one of the largest data-sets of hip fracture. Like other studies, we found an increased risk of hip fracture to be associated with smoking (Cummings et al., 1997), low body mass index (Farahmand et al., 2000), alcohol intake (for men) (Yuan et al., 2001) and anticonvulsants (Kinjo et al., 2005), and obesity to be protective against hip fracture (Farahmand et al., 2000), giving a high level of face validity to this study.

### Limitations of the study

Although using a large, nationally representative database provides important data from a large sample, detailed clinical information is less available than in smaller studies. Diagnostic categories found on the GPRD are not operationalised and are therefore unlikely to be exactly the same as those found in research or psychiatric practice, and information such as bone mineral density is not available. Lifestyle variables (body mass index, smoking and alcohol intake) are recorded optionally and were missing in a significant number of cases and controls. We created a ‘missing’ category to enable us to use these fields in the analysis. Residual confounding is therefore possible. Some prescribing of neuroleptics occurs in secondary care only and information on secondary care prescriptions was not available. For this reason details of dosages of antipsychotic medication over time were not reliable, and we categorised neuroleptic exposure as a dichotomous variable (ever/never); some patients might not have received large doses of antipsychotics, and if this were the case the relationship between antipsychotic medication might be owing to mechanisms such as falls, rather than secondary to hyperprolactinaemia. In addition, we could not
examine the effect of atypical antipsychotic medications because too few patients had been prescribed them on the GPRD during the exposure period, nor could we examine the effect of individual neuroleptic drugs in a multivariable analysis as there was insufficient statistical power. Larger, more up-to-date data-sets could address these issues in the future.

Although case-control studies are prone to bias, one advantage of using a data-set such as the GPRD is that the two major biases in this type of study – selection and recall bias – should be minimised by the prospective collection of data by general practitioners. Reverse causation is normally a possible explanation in case-control studies; this possibility should have been excluded here by looking at exposure prior to the occurrence of the first fracture. These findings should also be generalisable.

This study found significant evidence of an association between a diagnosis of schizophrenia and hip fracture, which appeared to be partly explained by neuroleptic medication. This adds to the growing body of evidence of an association between neuroleptic medication and bone mineral density loss. Patients with psychiatric disorders are less likely to have their medical illness diagnosed (Koranyi, 1979; Koran et al, 1989; Redelmeier et al, 1998) and medically managed (Redelmeier et al, 1998), and there is some evidence to suggest that they are less likely to have osteoporosis screened for or treated compared with age-matched control patients (Bishop et al, 2004). If this is the case this has serious public health implications, because the patients who have taken long-term neuroleptic medications are precisely those patients who are probably not being screened for osteoporosis. The evidence base for routinely screening patients prescribed neuroleptics is not available at present and clinicians urgently need more data on who is at highest risk and when (we do not yet know whether psychiatric patients are most at risk of developing osteoporosis after antipsychotic medication is initiated, or after dose-dependent long-term exposure). Randomised controlled trials of interventions to prevent fractures in these patients would enable more effective prophylaxis to be provided by mental health services and by primary care. However, if our findings are replicated, preventive measures should become part of the treatment of patients taking long-term prolactin-raising antipsychotic drugs and may include advice to patients about the importance of a balanced diet containing sufficient amounts of calcium and vitamin D, regular weight-bearing exercise, avoidance of tobacco, caffeine and alcohol, and sufficient exposure to sunlight.

ACKNOWLEDGEMENTS

We thank Professor Richard Hubbard and Christopher Smith, University of Nottingham for their valuable advice and assistance. The study was funded by the Sir Halley Stuart Trust.

REFERENCES


Table 4 Odds ratios for individual antipsychotic medications prescribed before or on day of fracture: General Practice Research Database 1987–1999

<table>
<thead>
<tr>
<th>Drug</th>
<th>Cases, n</th>
<th>Controls, n</th>
<th>Unadjusted odds ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chlorpromazine</td>
<td>270</td>
<td>204</td>
<td>2.48 (2.06–2.99)</td>
</tr>
<tr>
<td>Droperidol</td>
<td>20</td>
<td>5</td>
<td>7.65 (2.86–20.41)</td>
</tr>
<tr>
<td>Flupentixol</td>
<td>233</td>
<td>302</td>
<td>1.43 (1.20–1.71)</td>
</tr>
<tr>
<td>Fluphenazine</td>
<td>29</td>
<td>16</td>
<td>3.34 (1.81–6.16)</td>
</tr>
<tr>
<td>Haloperidol</td>
<td>307</td>
<td>191</td>
<td>2.97 (2.47–3.57)</td>
</tr>
<tr>
<td>Pimozide</td>
<td>19</td>
<td>19</td>
<td>1.86 (0.98–3.53)</td>
</tr>
<tr>
<td>Pipotazine</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Prochlorperazine</td>
<td>2100</td>
<td>3668</td>
<td>1.03 (0.97–1.10)</td>
</tr>
<tr>
<td>Promazine</td>
<td>227</td>
<td>134</td>
<td>3.26 (2.60–4.08)</td>
</tr>
<tr>
<td>Sulpiride</td>
<td>42</td>
<td>26</td>
<td>3.00 (1.84–4.90)</td>
</tr>
<tr>
<td>Thoridazine</td>
<td>1403</td>
<td>987</td>
<td>2.87 (2.63–3.13)</td>
</tr>
<tr>
<td>Trifluoperazine</td>
<td>277</td>
<td>248</td>
<td>2.06 (1.72–2.46)</td>
</tr>
<tr>
<td>Zuclopenhexil</td>
<td>27</td>
<td>14</td>
<td>3.36 (1.75–6.44)</td>
</tr>
</tbody>
</table>

Combination therapy

Fluphenazine and nor-triptypline

Trifluoperazine with various combinations

Atypical antipsychotics

Amisulpride | 0 | 2 |
| Clozapine | 0 | 0 |
| Olanzapine | 4 | 3 | 2.67 (0.60–11.91) |
| Quetiapine | 0 | 1 |
| Risperidone | 21 | 11 | 3.38 (1.57–7.24) |


|Louise Howard, PhD, MRCP, MRCPsych, Graham Kirkwood, MSc, Morven Leese, PhD, Health Services Research Department, Institute of Psychiatry, Kings College London, UK|

Correspondence: Dr Louise Howard, PO 29, Health Services Research Department, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK. Email: l.howard@iop.kcl.ac.uk

(First received 28 February 2006, final revision 25 July 2006, accepted 1 September 2006)
Phenomenology of delirium
Assessment of 100 adult cases using standardised measures

DAVID J. MEAGHER, MARIA MORAN, BANGARU RAJU, DYMPNA GIBBONS, SINEAD DONNELLY, JEAN SAUNDERS and PAULA T. TRZEPACZ

Background Delirium phenomenology is understudied.

Aims To investigate the relationship between cognitive and non-cognitive delirium symptoms and test the primacy of inattention in delirium.

Method People with delirium (n=100) were assessed using the Delirium Rating Scale – Revised –98 (DRS – R98) and Cognitive Test for Delirium (CTD).

Results Sleep – wake cycle abnormalities and inattention were most frequent, while disorientation was the least frequent cognitive deficit. Patients with psychosis had either perceptual disturbances or delusions but not both. Neither delusions nor hallucinations were associated with cognitive impairments. Inattention was associated with severity of other cognitive disturbances but not with non-cognitive items. CTD comprehension correlated most closely with non-cognitive features of delirium.

Conclusions Delirium phenomenology is consistent with broad dysfunction of higher cortical centres, characterised in particular by inattention and sleep – wake cycle disturbance. Attention and comprehension together are the cognitive items that best account for the syndrome of delirium. Psychosis in delirium differs from that in functional psychoses.

Declaration of interest P.T. is an employee of Eli Lilly. D.M. has an unrestricted educational grant from Astra Zeneca Pharmaceuticals.

Although our understanding of the clinical epidemiology of delirium has advanced considerably over the past decade, greater phenomenological study should allow more targeted studies of underlying mechanisms and therapeutic response. Delirium involves a constellation of symptoms reflecting widespread disruption of higher cortical functions that characteristically occur with an acute onset and fluctuating course. However, the interrelationship of delirium symptoms and their relevance to aetiology, treatment experience and outcome are poorly understood. Moreover, there is a dearth of research using validated instruments designed to assess the phenomenological breadth and complexity of this disorder (Turkel et al., 2006).

Two validated tools open the way for more detailed phenomenological study of delirium. The Cognitive Test for Delirium (CTD; Hart et al., 1996) measures five cognitive domains using standard neuropsychological methods. The Delirium Rating Scale – Revised –98 (DRS – R98; Trzepac et al., 2001a,b) covers a broad range of delirium symptoms not measured by other delirium instruments, including language, thought process abnormalities, visuospatial ability and both short- and long-term memory. We report a 2-year study of the frequency and severity of symptoms in 100 cases of delirium occurring in a palliative care setting using the DRS – R98 and the CTD. We explored the inter-relationship among delirium symptoms and, by measuring cognition carefully in conjunction with the DRS – R98, tested the primacy of inattention in delirium.

METHOD

Study design
We conducted a prospective cross-sectional study of delirium symptoms and cognitive performance in consecutive cases of DSM–IV delirium referred from a palliative care in-patient service. Patients assessed on daily ward rounds by the palliative care team as having altered mental state were screened with the Confusion Assessment Method (CAM; Inouye et al., 1990) – a four-item instrument based on DSM–III–R criteria. Patients were not included if they were near death or if circumstances were too difficult to allow assessment (in the opinion of the treating medical team), which resulted in a small number (less than 10%) being excluded. During the study period there were 434 new admissions to the unit, of which 100 (23%) are described here.

Delirium according to DSM–IV criteria (American Psychiatric Association, 1994) was confirmed by a research physician – (either the principal investigator (D.J.M.) or one of three specialist registrars trained to establish acceptable interrater reliability. Each case was then assessed by completion of the DRS–R98 followed by the CTD. The DRS–R98 rated the preceding 24 h period, whereas the CTD measured cognition at the time of its administration. Responses to the CTD were not used to rate DRS–R98 items. Both the DRS–R98 and the CTD are well-validated instruments, highly structured and anchored for rating and scoring.

Consent
The procedures and rationale for the study were explained to all patients, but because of their delirium at entry into the study it was presumed that most were not capable of giving informed written consent. Because of the non-invasive nature of the study, ethics committee approval was given to augment patient assent with proxy consent from next of kin (where possible) or a responsible caregiver for all participants in accordance with the Helsinki guidelines for medical research involving human subjects (World Medical Association, 2004).

Assessments
Demographic data, psychotropic drug exposure and the possibility of underlying dementia (suggested by history or investigation) were collected. Nursing staff were interviewed to assist rating of symptoms over the previous 24 h.

Delirium Rating Scale – Revised –98
The original Delirium Rating Scale (Trzepac et al., 1988) is widely used to measure symptom severity in delirium, but
has the limitations of grouping cognitive disturbances into a single item, not distinguishing motoric disturbances and not assessing thought process or language disorder. It has therefore been substantially revised to allow broad phenomenological assessment and serial ratings. The DRS–R98 is a 16-item scale with 13 severity items and 3 diagnostic items and it has high interrater reliability, sensitivity and specificity for detecting delirium in mixed neuropsychiatric and other hospital populations (Trzepacz et al, 2001a). It was validated both as a total scale (16 items) and a severity scale (13 items) for repeated measures. Each item is rated 0 (absent/normal) to 3 (severe impairment), with descriptions anchoring each severity level. Severity scale scores range from 0 to 39, with higher scores indicating more severe delirium. Delirium typically involves scores above 15 points (severity scale) or 18 points (total scale). For determination of item frequencies in this study, any item scoring at least 1 was considered present.

**Cognitive Test for Delirium**

The CTD (Hart et al, 1996) was specifically designed to assess patients with delirium – in particular those who are intubated or unable to speak or write. It assesses 5 neuropsychological domains (orientation, attention, memory, comprehension and vigilance), emphasising non-verbal (visual and auditory) modalities. Each individual domain is scored 0–6 in 2-point increments, except for comprehension which is scored in single-point increments. Total scores range between 0 and 30, with higher scores indicating better cognitive function. This measure reliably differentiates delirium from other neuropsychiatric conditions including dementia, schizophrenia and depression (Hart et al, 1997).

Performance on individual neuropsychological sub-tests (e.g. attention) can be scored on a 4-point scale (6 normal, 4 mild inattention, 2 moderate inattention, 0 severe inattention). Item severities were used to compare the relationship between individual items of the DRS–R98 to assess the relationship between cognitive and non-cognitive elements of delirium.

**Aetiology**

Attribution of aetiology based on all available clinical information was made by the palliative care physician according to a standardised delirium aetiology checklist (further information available from the authors upon request) with 12 categories: drug intoxication, drug withdrawal, metabolic/endocrine disturbance, traumatic brain injury, seizures, infection/inflammation (intracranial), infection (systemic), neoplasia (intracranial), neoplasia (systemic), cerebrovascular, organ insufficiency, other central nervous system disorder and other systemic disorder. The presence and suspected role of multiple potential causes were documented for each case of delirium, rated on a 5-point scale for degree of attribution to the delirium episode, ranging from ‘ruled out/not present/not relevant’ (0) to ‘definite cause’ (4).

**Statistical analyses**

Statistical analysis was conducted using the Statistical Package for the Social Sciences version 10.1. Demographic and rating scale data were expressed as means plus standard deviation. Continuous variables were compared by one-way analysis of variance (ANOVA). The severity of categorical and/or quasi-continuous variables such as the individual items of the DRS–R98 and CTD was compared with chi-squared analyses. Pearson correlations were performed between some individual items and between scale total scores. Level of significance was determined with a cut-off of 0.05, except where multiple comparisons were made when a Bonferroni correction ($P<0.001$) was applied.

**RESULTS**

Half of the 100 patients in the study were men, and the mean age of the group was 70.1 years (s.d. = 11.5). A mean of 3.5 (s.d. = 1.3) aetiological categories were noted per case, with neoplasia (67%), systemic infection (63%), metabolic–endocrine disorder (45%), organ failure (32%), drug intoxication (27%) and central nervous system lesions (26%) being the most common contributing causes. Patients had a mean DRS–R98 total score of 21.1 (s.d. = 5.5) and severity score of 16.6 (s.d. = 5.5), and a mean CTD score of 14.5 (s.d. = 8.1). The characteristics of patients with delirium only are compared with those of patients with comorbid dementia in Table 1.

Table 2 summarises the cognitive and non-cognitive disturbances assessed with the DRS–R98. Inattention (diagnostic criterion A of DSM–IV) was present in 97% of patients; other cognitive deficits were also common (76–89%), disorientation being the least frequent. Among the non-cognitive items, sleep disturbance (97%) and motoric disturbance (62%) each for hypoactive and hyperactive items, with 31 patients having evidence of both) were common, such that 94 patients had evidence of at least some degree of motoric disturbance (items 7 and 8 of DRS–R98). Language and thought process abnormalities were each present in over half the group but were less common than cognitive symptoms. Even when only more severe degrees of impairment were considered, attention and sleep–wake cycle deficits remained the most common, each at 73%.

Forty-nine patients had evidence of psychosis, as defined by a score of $\geq 2$ on item 2 (perceptual disturbances), item 3 (delusions) or item 6 (thought disturbance) on the DRS–R98. Eighteen of these patients scored 3 on one of these three items, indicating florid psychosis. The 49 patients with psychosis were not significantly different from the other 51 patients regarding motoric profile (DRS–R98 items 7 and 8) and overall severity of cognitive disturbance (measured by the CTD). They were younger ($t=1.9$, $P=0.05$) with higher total DRS–R98 scores ($t=−3.8$, $P<0.001$) and more severe affective lability ($t^2=16.1$, d.f.=2, $P<0.001$).

Patients with psychosis tended to have disturbance of a single psychotic component, with only 6 of these 49 patients scoring $\geq 2$ on more than one item. For the whole cohort, DRS–R98 items 2 (perceptual disturbance) and 3 (delusions) were not significantly correlated ($r=0.16$; item 6 (thought disturbance) was not significantly correlated with item 2 ($r=0.15$) or item 3 ($r=0.01$). Moreover, when the analysis was restricted to patients with psychosis ($n=49$), thought disturbance and perceptual disturbances were inversely correlated ($r=−0.49$, $P=0.001$) and both delusions ($r=−0.59$, $P=0.001$) and thought disturbance ($r=−0.35$, $P=0.01$) correlated positively with affective lability, whereas perceptual disturbance was negatively correlated with affective lability ($r=−0.41$, $P=0.003$).

Although neither delusions nor perceptual disturbances correlated significantly with any of the cognitive items of DRS–R98 or CTD, thought process disturbance correlated with impairments of attention ($r=−0.46$, $P=0.001$), memory ($r=−0.40$, $P<0.01$), orientation ($r=−0.30$, $P=0.03$) and comprehension ($r=−0.28$, $P=0.05$).
items on the CTD, and with attention ($r=0.59$, $P<0.001$), orientation ($r=0.33$, $P=0.03$) and long-term memory ($r=0.34$, $P=0.03$) items – but not short-term memory or visuospatial function items – on the DRS–R98.

Cognitive dysfunction rated with the CTD is shown in Table 3. This shows widespread impairment of neuropsychological function, with the most frequent (94%) and severest impairments in attention and vigilance. This parallels the DRS–R98 impairments, of which attention was most often impaired and orientation least impaired, even though these scales were rated independently of one another and for different time frames – DRS–R98 for the previous 24 h and CTD for current performance. The DRS–R98 attention item includes distractibility and therefore encompasses both attention and vigilance as assessed in the CTD. Corresponding items on the CTD and the DRS–R98 correlated highly: DRS–R98 orientation and CTD orientation ($r=−0.75$), DRS–R98 attention and CTD attention ($r=−0.73$), DRS–R98 attention and CTD vigilance ($r=−0.60$), and CTD memory with DRS–R98 short-term memory ($r=−0.47$) and long-term memory ($r=−0.61$). Interestingly, CTD comprehension correlated with the DRS–R98 item for language ($r=−0.42$, $P=0.001$) but not with thought process abnormalities ($r=−0.09$).

In view of the central role given to disturbed attention in current delirium descriptions, patients were divided into three categories according to the severity of attentional deficit measured using the CTD: score 4–6 ($n=32$), score 2 ($n=34$) and score 0 ($n=34$). These groups differed for many items (Table 4); however, when significance levels were corrected for multiple comparisons, the degree of inattention was associated with the level of impairment of other cognitive disturbances (rated on both CTD and DRS–R98) but not the non-cognitive DRS–R98 items, except for language ($x^2=19.5$, d.f. $=6$, $P=0.001$).

We further examined whether impairment on the other CTD items related to scores on DRS–R98 items as strongly as did CTD attention, to ascertain whether attention had a unique role. After corrections for multiple comparisons, the severity of vigilance impairment was closely related to all other aspects of cognition but not to non-cognitive items (except for language) and thus mirrored the findings with the CTD attention item. Orientation, memory and comprehension were less strongly associated with DRS–R98 cognitive items (Table 5). In contrast to attention, severity of comprehension disturbance was associated with the most non-cognitive DRS– R98 symptoms, including sleep–wake cycle disturbance, psychomotor retardation and language difficulties. These patterns suggest two different domains of delirium symptoms. Seventeen patients had documented evidence of pre-existing cognitive deficits, suggesting their delirium co-occurred with chronic cognitive impairment. These patients were significantly older, had a greater aetiological burden of underlying diseases, and had more severe disturbances on the DRS–R98 and CTD than patients with delirium only (see Table 1). This difference in severity of DRS–R98 scores was accounted for by greater disturbance on the five DRS–R98 cognitive items ($r=−2.8$, $P<0.01$) rather than the eight DRS–R98 neuropsychiatric and behavioural items.

Out of concern that the inclusion of patients ($n=17$) with comorbid pre-existing cognitive impairment might have influenced findings, analyses were repeated for the study population with delirium only ($n=83$). The findings regarding DRS–R98 item frequencies, patterns of psychosis and interrelationship of cognitive items on CTD and DRS–R98 phenomenology were essentially unaltered.

**DISCUSSION**

This work investigates a more comprehensive range and specificity of symptoms than previous studies of delirium. We assessed 100 consecutive cases of DSM–IV delirium using valid, sensitive and standardised
Delirium symptoms can be divided into 'core' features that are almost invariably present (disturbances of attention, memory, orientation, language, thought processes and sleep–wake cycle) and 'associated' features that are more variable in presentation (e.g. psychotic symptoms, affective disturbances, different motoric profiles) (American Psychiatric Association, 1999; Trzepacz, 1999). Disturbance of attention is a cardinal symptom of delirium and in our analysis associated strongly with all other cognitive deficits and language, but not with most of the non-cognitive features. Some neurologists have viewed delirium as a disorder of attention. However, the frequency of non-cognitive symptoms and their lack of association with the severity of objectively measured attentional impairment strongly support the view of delirium as primarily a disorder of cognitive, behavioural and psychopathological features. We assessed the frequency and severity of less studied symptoms including visuospatial impairment, disorganised thinking, language impairment and different components of attention, memory, and motoric presentations, as well as more detailed evaluation of characteristics of sleep–wake cycle abnormality, perceptual disturbances and thought process abnormality. Previous phenomenological work has generally classed symptoms as present or absent without proportioning severity. This can result in more minor disturbances (e.g. of sleep) that are common in all hospitalised patients being rated as equivalent to more significant major disturbances (e.g. sleep–wake cycle reversal) that occur in delirium.

Our findings support the concept of delirium as primarily a disorder of cognitive dysfunction and mean item scores assessed with the Cognitive Test for Delirium (n=100).

### Table 3

Frequency of different severity levels of cognitive dysfunction and mean item scores assessed with the Cognitive Test for Delirium (n=100)

<table>
<thead>
<tr>
<th>CTD item</th>
<th>Frequency, %</th>
<th>CTD score</th>
<th>Mean (s.d.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Score 5–6</td>
<td>Score 3–4</td>
<td>Score 1–2</td>
</tr>
<tr>
<td>Orientation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attention</td>
<td>27</td>
<td>21</td>
<td>30</td>
</tr>
<tr>
<td>Memory</td>
<td>16</td>
<td>34</td>
<td>34</td>
</tr>
<tr>
<td>Comprehension</td>
<td>35</td>
<td>17</td>
<td>39</td>
</tr>
<tr>
<td>Vigilance</td>
<td>14</td>
<td>27</td>
<td>26</td>
</tr>
</tbody>
</table>

CTD, Cognitive Test for Delirium.

1. Range 0–6; lower scores indicate poorer performance.

### Table 4

Item scores for the two delirium scales according to degree of inattention on the Cognitive Test for Delirium

<table>
<thead>
<tr>
<th>Item</th>
<th>Item score: mean (s.d.)</th>
<th>p 0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CTD attention score 4 or 6</td>
<td>CTD attention score 2</td>
</tr>
<tr>
<td></td>
<td>(n=32)</td>
<td>(n=34)</td>
</tr>
<tr>
<td>DRS–R98</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 Sleep–wake cycle disturbances</td>
<td>1.5 (0.6)</td>
<td>1.6 (0.7)</td>
</tr>
<tr>
<td>2 Perceptual disturbances and hallucinations</td>
<td>1.0 (1.0)</td>
<td>0.6 (0.9)</td>
</tr>
<tr>
<td>3 Delusions</td>
<td>0.4 (0.9)</td>
<td>0.5 (0.8)</td>
</tr>
<tr>
<td>4 Lability of affect</td>
<td>0.6 (0.7)</td>
<td>0.7 (0.8)</td>
</tr>
<tr>
<td>5 Language</td>
<td>0.4 (0.6)</td>
<td>0.9 (0.8)</td>
</tr>
<tr>
<td>6 Thought process abnormalities</td>
<td>0.4 (0.6)</td>
<td>0.9 (0.8)</td>
</tr>
<tr>
<td>7 Motor agitation</td>
<td>0.7 (0.8)</td>
<td>0.9 (0.8)</td>
</tr>
<tr>
<td>8 Motor retardation</td>
<td>0.9 (0.8)</td>
<td>0.9 (0.9)</td>
</tr>
<tr>
<td>9 Orientation</td>
<td>0.7 (0.7)</td>
<td>1.2 (0.9)</td>
</tr>
<tr>
<td>10 Attention</td>
<td>1.2 (0.6)</td>
<td>2.0 (0.5)</td>
</tr>
<tr>
<td>11 Short-term memory</td>
<td>1.3 (1.0)</td>
<td>1.5 (0.7)</td>
</tr>
<tr>
<td>12 Long-term memory</td>
<td>1.4 (1.0)</td>
<td>1.9 (0.9)</td>
</tr>
<tr>
<td>13 Visualspatial ability</td>
<td>1.2 (1.0)</td>
<td>1.7 (0.8)</td>
</tr>
<tr>
<td>Severity score</td>
<td>12.0 (4.2)</td>
<td>15.5 (4.3)</td>
</tr>
<tr>
<td>Severity score minus attention item</td>
<td>10.8 (3.9)</td>
<td>13.5 (4.2)</td>
</tr>
<tr>
<td>CTD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Orientation</td>
<td>4.6 (1.6)</td>
<td>2.9 (2.2)</td>
</tr>
<tr>
<td>Comprehension</td>
<td>5.5 (0.8)</td>
<td>4.7 (1.2)</td>
</tr>
<tr>
<td>Memory</td>
<td>4.5 (1.5)</td>
<td>2.5 (1.9)</td>
</tr>
<tr>
<td>Vigilance</td>
<td>4.0 (1.8)</td>
<td>2.7 (1.6)</td>
</tr>
<tr>
<td>Total minus attention item</td>
<td>18.1 (4.5)</td>
<td>12.6 (4.5)</td>
</tr>
</tbody>
</table>


1. Lower scores are worse on CTD; higher scores are worse on DRS–R98.
2. p-test for item comparisons and one-way analysis of variance for total scale scores.
3. Values after Bonferroni correction.
other symptoms, for example, sleep–wake cycle disturbance, altered motoric behaviours, and thought content and process abnormalities. Sleep–wake cycle disturbance may underlie the fluctuating nature of delirium severity over a 24 h period (Balan et al., 2003).

**Pattern of cognitive disruption in delirium**

This study confirms delirium as a disorder of global cognition characterised by a prominent disturbance of attention and vigilance. Disorientation was the least frequent cognitive symptom, even though many non-psychiatric physicians rely on bedside tests of orientation to time, place and person as their principal mental status evaluation. Almost a quarter of our delirious patients had no evidence of disorientation on the DRS-R98 and only 52% had evidence of greater than mild disturbance of orientation on the CTD. The use of disorientation as a key indicator of delirium is thus fraught with the likelihood of missed cases, and the use of other, more consistent symptoms (such as inattention) would be a more reliable way of screening for suspected delirium. The use of instruments such as the Mini-Mental State Examination (Folstein et al., 1975), which are heavily weighted towards orientation, to detect or monitor delirium is therefore not supported by these findings.

The cognitive impairment of delirium may represent a single construct or a constellation of elements with differing underpinnings. Poor performance on CTD attention and vigilance items was significantly related to the degree of disturbance on all other cognitive items on both the CTD and DRS-R98, but much less so for non-cognitive items. Because intact attention is required to recall new information, it is unclear whether the short-term memory deficits measured on the DRS-R98 (tested in verbal modality) and the visual memory deficits measured on the CTD are truly primary memory dysfunctions or secondary to attentional deficits. The DRS-R98 long-term memory impairments may be more related to retrieval problems and perhaps less affected by inattention than short-term memory for new material.

Performance on CTD orientation, memory and comprehension items was significantly related to fewer cognitive items compared with CTD attention. The CTD comprehension item (comprising a combination of language and executive function) was associated with more non-cognitive DRS-R98 items than the other CTD items and may denote a different domain of delirium symptoms than does attention. The combination of disturbed attention and comprehension may best represent the underlying disturbances central to overall delirium phenomenology.

Visuospatial abnormalities are not usually measured in delirium assessments even though they may underlie problems of wandering and poor environmental interactions. Mean visuospatial ability scores were almost as impaired as attention, and CTD attention is measured in a visuospatial modality. This overlap may reflect the shared role of the non-dominant posterior parietal cortex in both attention and visuospatial functions (Trzepacz, 1999).

Despite an enduring emphasis on the characteristic fluctuating nature of delirium, this has not been directly studied. Ratings of equivalent cognitive items on the DRS-R98 and CTD were highly correlated (inversely as expected), despite one being a symptom rating scale evaluating a 24 h period and the other a cognitive test measuring current status. This suggests that certain delirium symptoms – cognition and language – are not as fluctuant as previously described, although this requires further scrutiny with serial measurement over relatively short periods.

**Psychotic symptoms**

The significance of psychotic symptoms in delirium remains unclear. It is not known whether patients develop these features due to specific physiological causes, cognitive impairment with misunderstanding of the external environment, misperceptions, as part of mood disturbances, or through some other aspect of individual patient vulnerability (Francis, 1992). We found that thought process abnormalities – but not delusions or perceptual disturbances – correlated with overall cognitive impairment. Both delusions and thought disorder correlated with affective lability, although perceptual disturbance was inversely correlated to both thought disorder and affective lability. Previous work comparing the psychosis of delirium with that of schizophrenia found that in delirium thought content disturbances tended to involve themes from the immediate environment and circumstances, hallucinations were frequently visual rather than auditory, and formal thought disorder typically comprised poverty of thinking and illogicality.
sleep–wake cycle (Fann et al., 2005), and that orientation difficulties, inattention, poor memory, emotional lability and sleep disturbances are more persistent symptoms (Lekkoff et al., 1994; McCusker et al., 2003).

Second, the inclusion of patients with dementia might affect the clinical profile but there was little discernible effect when our study analyses were repeated for the pure-delirium study population. It appears that delirium phenomenology is altered little by the presence of dementia (Trzepacz et al., 1998), such that delirium symptoms tend to overshadow dementia when they co-exist although these symptoms do occur in the context of greater overall cognitive impairment. Equally, it should be recognised that in order to be truly representative of delirium, studies need to include patients who also have dementia, in recognition of the substantial comorbidity between the two conditions.

This study describes delirium phenomenology in a palliative care population, which may restrict its generalisability to other groups with this condition. Delirium is considered a unitary syndrome with a stereotyped constellation of symptoms thought to reflect disturbance of a final common neural pathway (Trzepacz, 1999). Moreover, the term has subsumed the many synonyms that have been used to denote acute generalised cognitive disturbances in various settings but were not based on scientific evidence. Nonetheless, clinical profile may be influenced by factors that characterise different aetiological or treatment settings, but single studies have not compared symptom profiles across patient groups. Delirium occurring in cancer patients tends to be particularly multifactorial in causation, with hypoactive motoric presentations especially common (Morita et al., 2001; Centeno et al., 2004; Spiller & Keen, 2006). Our sample included patients with a broad range of relevant aetiologies and medications, many with significant psychotropic effects that could alter clinical presentation. Further studies are needed to explore the impact of aetiological, treatment and other individual patient factors on the clinical presentation of delirium.

Advancing the concept of delirium

The concept of delirium has evolved considerably over the past 25 years. This is reflected in recent studies comparing diagnostic frequency when DSM-III, DSM-III-R, DSM-IV and ICD–10 criteria are applied to single populations (Laurila et al., 2003; Cole et al., 2003). Future descriptions will allow further refinement of the syndrome in keeping with emerging evidence and need to account for key phenomenological issues, including the following:

- delirium detection and diagnosis are confounded by inadequate appreciation of variations in presentation and breadth of symptoms;
- core features used to define delirium should be readily detectable and occur with consistency; over-reliance on less common symptoms contributes to non-detection, which in turn hampers clinical and research efforts;
- core defining features should differentiate delirium from other neuropsychiatric disorders, especially dementia.

Study limitations

Studies with cross-sectional designs do not examine symptom evolution or whether domains of symptoms vary as overall severity changes. Longitudinal studies suggest that early delirium is characterised by psychomotor disturbances and a disrupted sleep–wake cycle (Fann et al., 2005), and that orientation difficulties, inattention, poor memory, emotional lability and sleep disturbances are more persistent symptoms (Lekkoff et al., 1994; McCusker et al., 2003).

Second, the inclusion of patients with dementia might affect the clinical profile but there was little discernible effect when our study analyses were repeated for the pure-delirium study population. It appears that delirium phenomenology is altered little by the presence of dementia (Trzepacz et al., 1998), such that delirium symptoms tend to overshadow dementia when they co-exist although these symptoms do occur in the context of greater overall cognitive impairment. Equally, it should be recognised that in order to be truly representative of delirium, studies need to include patients who also have dementia, in recognition of the substantial comorbidity between the two conditions.

This study describes delirium phenomenology in a palliative care population, which may restrict its generalisability to other groups with this condition. Delirium is considered a unitary syndrome with a stereotyped constellation of symptoms thought to reflect disturbance of a final common neural pathway (Trzepacz, 1999). Moreover, the term has subsumed the many synonyms that have been used to denote acute generalised cognitive disturbances in various settings but were not based on scientific evidence. Nonetheless, clinical profile may be influenced by factors that characterise different aetiological or treatment settings, but single studies have not compared symptom profiles across patient groups. Delirium occurring in cancer patients tends to be particularly multifactorial in causation, with hypoactive motoric presentations especially common (Morita et al., 2001; Centeno et al., 2004; Spiller & Keen, 2006). Our sample included patients with a broad range of relevant aetiologies and medications, many with significant psychotropic effects that could alter clinical presentation. Further studies are needed to explore the impact of aetiological, treatment and other individual patient factors on the clinical presentation of delirium.

REFERENCES


Meagher, D. J., Quinn, J. F., Bourke, S., et al. (2004) Longitudinal assessment of psychopathological domains over late-stage schizophrenia in relation to duration of initially untreated psychosis: 3-year prospective study in


Reducing emotional distress in people caring for patients receiving specialist palliative care

Randomised trial

KIRI WALSH, LOUISE JONES, ADRIAN TOOKMAN, CHRISTINA MASON, JOANNE MCLoughlin, ROBERT BLIZARD and MICHAEL KING

Background  Caring for relatives with advanced cancer may cause psychological and physical ill health.

Aims  To evaluate the effectiveness of increased support for distressed, informal carers of patients receiving palliative care.

Method  The sample was composed of 271 informal carers who scored over 5 on the 28-item General Health Questionnaire (GHQ–28). The intervention comprised six weekly visits by a trained advisor. Primary outcome was carer distress (GHQ–28) at 4-week, 9-week and 12-week follow-up. Secondary outcomes were carer strain and quality of life, satisfaction with care, and bereavement outcome.

Results  Scores on the GHQ–28 fell below the threshold of 5/6 in a third of participants in each trial arm at any follow-up point. Mean scores in the intervention group were lower at all time points but these differences were not significant. No difference was observed in secondary outcomes. Carers receiving the intervention reported qualitative benefit.

Conclusions  The intervention might have been too brief, and ongoing help might have had accruing benefits. Alternatively, informal carers of patients with cancer may already receive considerable input and the advisor’s help gave little additional advantage; or caring for a dying relative is extremely stressful and no amount of support is going to make it much better.

Declaration of interest  None.

Family members and friends who care for patients with advanced cancer living at home are at risk of psychological and physical ill health (Field et al., 1993; Chapman & Pepler, 1998; Payne et al., 1999; Rhodes & Shaw, 1999; Soothill et al., 2001; Thomas et al., 2002). Specialist palliative care services working in the community developed to respond to the complex problems experienced by these patients. Although these professionals may be in contact with patients’ families and friends, their main locus is the patient (Higginson et al., 2003). The National Institute for Health and Clinical Excellence (NICE) in Britain recently recommended expansion of specialist palliative care services to multiprofessional support for carers, independent of patient care (National Institute for Clinical Excellence, 2004). However, there is no consensus on what sort of intervention would ease carers’ burden, or its effectiveness. This trial was conceived and completed before the publication of the NICE recommendations. Specialist palliative care teams across London were actively involved in the planning, piloting and conduct of this research. In the summer of 1998 we asked 60 informal carers of patients with cancer, under the care of three local palliative care teams, to indicate their preferred mode of extra support from a number of options which included respite care, other practical help, more written information and telephone advice. Over 80% of respondents chose a weekly visit by a trained advisor.

We aimed to evaluate the effectiveness of an intervention to reduce symptoms of anxiety and depression and carer burden, improve quality of life and satisfaction with care, and reduce the intensity of grief reactions in distressed informal carers of patients with cancer. We predicted that a brief, carer-focused intervention, in addition to usual specialist palliative care, would be more effective at reducing distress than usual specialist palliative care alone.

METHOD

Study setting

Ethical approval was granted by the London Multi-Centre Regional Ethics Committee in February 2000, and subsequently by local research ethics committees. Seven specialist palliative care teams in three London cancer networks, serving a combined population of almost 2 million people, took part in the study.

Recruitment and randomisation

From January 2001 to April 2003 people providing informal care to patients in all new referrals to six of the participating teams were screened for psychological distress using the 28-item version of General Health Questionnaire (GHQ–28; Goldberg, 1970). The seventh team joined the trial in June 2001. The informal carer was identified by patients and palliative care teams as the main person who provided unpaid practical and emotional support to the patient on a regular basis and was in contact with the palliative care team. Palliative care staff introduced the study at the earliest opportunity, usually on second or third contact. The GHQ–28 was then completed immediately or if the informal carer was not present for the palliative care team visit, left for completion later. Carers returned questionnaires in pre-paid envelopes to the research team. The research team was informed if the carer declined to fill in the GHQ–28, if the patient was unlikely to survive the time it would take to introduce the intervention, or if the carer’s English skills would mean they could not gain full benefit from the advisor visits. Informal carers who scored above the threshold of 5/6 (Goldberg & Williams, 1988) on the GHQ–28 were approached to obtain informed consent and complete baseline assessments, which included a measure of carer burden (Robinson, 1983) and quality of life (Weitnner et al., 1999). Patients’ physical performance status was assessed using the criteria of the Eastern Cooperative Oncology Group (ECOG; Oken et al., 1982). On completion of the baseline assessments, participants were randomised using a block randomisation design, stratified according to the seven participating teams. Interviewers were masked to the block size of 12. Randomisation took place at the trial centre under the supervision of the trial statistician (R.B.).
**Trial arms**

**Usual care**

Specialist palliative care was provided by a team of clinical nurse specialists, with specialist medical support and sometimes specialist social work support, giving advice to patients at home, to their families and to their primary healthcare teams. Patients were assisted with control of pain and other physical symptoms as well as with social, psychological, emotional and spiritual issues. Some people are referred for palliative care close to death in the context of a rapidly changing clinical picture, whereas others remain in contact with palliative care services for many months.

**The carer advisor intervention**

The intervention was developed by the research team. Two part-time carer advisors with experience in community nursing and social work delivered the intervention, which consisted of six visits over a 6-week period. The advisors aimed to meet the carer alone, if necessary arranging meetings outside the home or at the carer’s workplace, sometimes during evenings or weekends. A comprehensive assessment of domains of need was made; past, present and future issues were discussed and advice, information and emotional support provided. The intervention was kept to giving advice and support rather than taking action on behalf of carers; however, advice might go as far as (for example) helping carers to calculate their benefit entitlements. In the event of a patient’s death during the intervention period, the advisors continued to offer visits, up to a total of six. Sometimes a telephone call took the place of a visit. Telephone calls enabled flexibility in the intervention and helped some carers to broach issues that were difficult to discuss face-to-face. Such calls were discussed with the research team to decide whether they were equivalent to a full intervention visit. The mean number of advisor contacts was 5.0 (s.d.=2.0), and the mean number of contacts up to the death of the patient was 3.6 (s.d.=2.6).

Both advisors undertook 1 month’s training, involving fieldwork in palliative care in the community, a hospice and a hospital setting. The advisors met weekly with the research team for debriefing, for advice on any issues that arose and to ensure that all domains of carer need were covered in the intervention. These domains were:

(a) patient care;
(b) physical health needs;
(c) need for time away from the patient in the short term and longer term;
(d) need to plan for the future;
(e) psychological health, relationships and social networks;
(f) relationships with health and social service providers;
(g) finances.

After 1 year, a further half-day in-service training session took place in which the delivery of the intervention was reviewed.

**Study outcomes**

Informal carers completed postal questionnaires at 4 weeks, 9 weeks and 12 weeks after randomisation (see Fig. 1). The first follow-up, part-way through the 6-week intervention, was chosen to achieve at least one research assessment in most cases before the patient died. When a patient died, the study participant was sent a sympathy card and contacted again 4 months later for the final follow-up. A patient’s death therefore necessarily ended the carer’s participation in the subsequent follow-up assessments. Our primary outcome was the proportion scoring above threshold (5/6) on the GHQ–28 at follow-up. Secondary outcomes were GHQ–28 score, Carer Strain Index (Robinson, 1983) and Care-Giver Quality of Life Index (Cancer) (Weitzner et al., 1999) scores 4 weeks, 9 weeks and 12 weeks after randomisation, and scores on Core Bereavement Items (CBI; Burnett et al., 1997) and satisfaction with care 4 months after the death of the patient. Brief, semi-structured interviews at the final follow-up provided a qualitative assessment of acceptability and helpfulness of the support provided by the intervention.

**Power and statistical analysis**

**Power and sample size**

All carers scored above the threshold 5/6 on the GHQ–28 at entry to the trial. Prospective...
research in other settings (Weich et al., 1997) indicated that, given the stresses involved, 70% of the usual care group would be likely to score above this threshold at follow-up. Thus our per protocol power calculation indicated that in order to detect a drop to 50% caseness in the experimental group at 90% power and the 5% level of significance, 124 carers would be required in each arm. To cover an expected 10% attrition from the trial we needed to recruit 280 carers, a sample that would also provide sufficient power for examination of GHQ–28 score as a continuous measure.

**Analysis**

Treatment success was defined as any drop in GHQ–28 score to below threshold, measured 4 weeks, 9 weeks or 12 weeks after randomisation. More detailed analyses were performed on GHQ–28, Carer Strain Index and quality of life scores from baseline, 4 weeks, 9 weeks and 12 weeks by a mixed model approach using the random intercept random slope facility provided by the generalised linear latent and mixed models (GILLAMM) procedure in Stata release 8 (Rabe-Hesketh & Everitt, 2004). The model was built in the following order:

(a) effect of treatment to detect overall difference between the groups;
(b) effect of time to detect linear change over time as a result of taking part in the trial;
(c) linear interaction to detect whether treatment groups changed over time in a different linear fashion;
(d) quadratic term for time to detect whether the change was curvilinear;
(e) quadratic interaction to detect whether the groups differed in their curvilinear change over time.

The most parsimonious model was selected, conditional on the inclusion of the main effect of the intervention. Group means on the CBI and satisfaction with care were compared in a one-way analysis of variance (ANOVA).

### RESULTS

#### Results of screening and recruitment

During the 28 months of recruitment 1577 new referrals were reported by the participating teams (Fig. 1). Referral details were sometimes lost if the informal carer did not meet the palliative care team and the GHQ–28 form had to be passed on by the patient, or when informal carers agreed to complete the GHQ–28 at a later time but failed to do so. In total 669 carers completed the GHQ–28 of whom 411 (61%) scored above the threshold. Fifty-five patients died before carer consent could be obtained. We invited 356 carers to take part in the trial and 271 (76%) of them agreed.

#### Follow-up rates at 4, 9 and 12 weeks

As expected, a number of participants were lost through the death of the patient. At 4 weeks 43 (16%) patients had died, by 9 weeks 85 (31%) had died and by 12 weeks 109 (40%) had died. Refusal rates at each follow-up point where the patient remained alive were 19% (43/228), 27% (50/186) and 24% (39/162) respectively (Fig. 1).

#### Follow-up rates at 4 months after death

Two hundred and twenty-one patients had died by end of data collection at the end of July 2003. Ninety-seven of 113 carers (86%) in the usual care arm and 84 of 108 in the intervention arm (78%) participated in the 4-month follow-up (82% overall).

### Characteristics of the study group by trial arm

Four-fifths of trial participants were women, 86% were White and 64% were spouses or partners of patients. Their mean age was 56.3 years (range 16–92) (Table 1). No major difference occurred between the...
randomised groups at baseline on demographic variables, GHQ–28 score or the patient’s physical performance status assessed using the criteria of the ECOG (Oken et al., 1982). However, there was some imbalance in carer strain and quality of life. There was no difference between trial arms in willingness of participants to complete follow-up assessments or in the patients’ life expectancies; median survival time from trial entry was 12 weeks (Table 1).

### Primary outcome

Approximately a third of carers in each trial arm reduced their distress enough to record a GHQ–28 score below the threshold of 5/6 at any follow-up point (Table 2).

### Secondary outcomes

We examined GHQ–28 scores in more detail. The GLLAMM models assume that data are missing at random. There was no difference in the follow-up GHQ–28 scores of those attending and those not attending their next follow-up assessment and the assumption is justified. Mean scores dropped by the 4-week and 9-week follow-up assessments but increased again by 12 weeks (Fig. 2, Table 3). Although the intervention group appears to experience greater improvement, the results of GLLAMM did not reach significance for the interaction effects. The most parsimonious model included significant values for time ($z = -4.70, P < 0.001$), which was curvilinear ($z = 3.00, P < 0.004$); the treatment effect was not significant ($z = -1.10, P = 0.272$). Carer quality of life deteriorated over time (Table 3) but there was no significant interaction between time and trial arm on this outcome or carer strain; nor were differences found in bereavement phenomena or satisfaction with care 4 months after the patients’ death (Table 4).

### Carers’ views

The most valued aspect of the service was the additional emotional support, with fewer carers reporting value from the added information, advice or practical or financial help. One-fifth of respondents felt the allocation of an advisor came too late in the patient’s illness and almost a third thought more sessions with the advisor would have been helpful (Table 5).

### DISCUSSION

We failed to support our main hypothesis that a brief intervention by a carer advisor would reduce psychological symptoms in distressed informal carers of cancer patients. Although a small treatment effect (Cohen, 1988) for the carer advisor intervention in reduction of psychological distress was observed in our secondary analysis, it was short-lived and did not reach statistical significance. To demonstrate that the treatment effects detected at 4 weeks and 9 weeks (s.d.=0.22 and s.d.=0.20 respectively) were statistically significant would have required 323 and 382 participants respectively in each trial arm.

### Strengths and limitations

Recruitment to this trial demonstrates that large-scale randomised controlled trials are possible in palliative care. Follow-up rates were acceptable, with face-to-face contact after the death of the patient exceeding 80%. Sixty-one per cent of carers scored above threshold on the GHQ–28 and were eligible for the trial, which means that there is considerable psychological morbidity in

---

**Table 2** Outcome in terms of threshold scoring on the 28-item General Health Questionnaire

<table>
<thead>
<tr>
<th></th>
<th>Usual care group</th>
<th>Care advisor group</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n/N (%)</td>
<td>n/N (%)</td>
<td></td>
</tr>
<tr>
<td>Below GHQ–28 threshold at any follow-up point with no relapse</td>
<td>21/91 (23)</td>
<td>21/100 (21)</td>
<td>$\chi^2(1)=0.73, P=0.73$</td>
</tr>
<tr>
<td>Below GHQ–28 threshold at any follow-up point</td>
<td>29/91 (32)</td>
<td>35/100 (35)</td>
<td>$\chi^2(1)=0.65, P=0.76$</td>
</tr>
</tbody>
</table>

GHQ–28, 28-item General Health Questionnaire.

---

**Table 3** Mean scores on primary and secondary outcomes over the study period

<table>
<thead>
<tr>
<th>Outcome measure</th>
<th>Pre-randomisation</th>
<th>Follow-up assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4 weeks</td>
<td>9 weeks</td>
</tr>
<tr>
<td>GHQ–28$^1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usual care group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score: mean (s.d.)</td>
<td>13.0 (5.2)</td>
<td>11.9 (6.4)</td>
</tr>
<tr>
<td>n</td>
<td>133</td>
<td>85</td>
</tr>
<tr>
<td>Intervention group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score: mean (s.d.)</td>
<td>12.8 (5.1)</td>
<td>10.5 (6.3)</td>
</tr>
<tr>
<td>n</td>
<td>137</td>
<td>97</td>
</tr>
<tr>
<td>Carer strain$^1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usual care group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score: mean (s.d.)</td>
<td>30.2 (11.5)</td>
<td>27.8 (11.5)</td>
</tr>
<tr>
<td>n</td>
<td>134</td>
<td>86</td>
</tr>
<tr>
<td>Intervention group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score: mean (s.d.)</td>
<td>27.1 (10.6)</td>
<td>27.7 (11.6)</td>
</tr>
<tr>
<td>n</td>
<td>137</td>
<td>99</td>
</tr>
<tr>
<td>Carer quality of life$^1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Usual care group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score: mean (s.d.)</td>
<td>66.4 (21.1)</td>
<td>62.9 (19.3)</td>
</tr>
<tr>
<td>n</td>
<td>132</td>
<td>82</td>
</tr>
<tr>
<td>Intervention group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score: mean (s.d.)</td>
<td>72.8 (21.1)</td>
<td>69.6 (22.4)</td>
</tr>
<tr>
<td>n</td>
<td>130</td>
<td>93</td>
</tr>
</tbody>
</table>

GHQ–28, 28-item General Health Questionnaire.

1. Higher scores indicate more psychological distress, greater carer strain and lower quality of life.
Table 4  Grief scores and satisfaction with care by trial arm

<table>
<thead>
<tr>
<th></th>
<th>Usual care group (n=97)</th>
<th>Intervention group (n=84)</th>
<th>Total (n=181)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>CBI score: mean (s.d.)</td>
<td>45.6 (11.6)</td>
<td>47.1 (11.2)</td>
<td>46.3 (11.4)</td>
<td></td>
</tr>
<tr>
<td>Considered care poor, n (%)</td>
<td>21 (22)</td>
<td>16 (19)</td>
<td>37 (21)</td>
<td></td>
</tr>
</tbody>
</table>

CBI, Core Bereavement Items.
1. One CBI questionnaire incomplete.
2. Two CBI questionnaires incomplete.
3. Data missing for two people.
4. Data missing for one person.

Table 5  Carers’ views of the content and timing of the carer advisor intervention (n=81)

<table>
<thead>
<tr>
<th></th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Content of the intervention</strong></td>
<td></td>
</tr>
<tr>
<td>Carer received additional practical or financial help</td>
<td>20 (25)</td>
</tr>
<tr>
<td>Carer found the additional advice useful</td>
<td>45 (56)</td>
</tr>
<tr>
<td>Carer found the additional information useful</td>
<td>47 (58)</td>
</tr>
<tr>
<td>Carer felt added emotional support</td>
<td>68 (84)</td>
</tr>
<tr>
<td>Overall the help was very/fairly useful</td>
<td>69 (85)</td>
</tr>
<tr>
<td><strong>Timing of the intervention</strong></td>
<td></td>
</tr>
<tr>
<td>Carer thought more sessions would have been useful</td>
<td>26 (32)</td>
</tr>
<tr>
<td>Carer felt the sessions with the advisor came at the right time</td>
<td>46 (57)</td>
</tr>
<tr>
<td>Carer felt the sessions with the advisor came too early in the patient’s illness</td>
<td>6 (7)</td>
</tr>
<tr>
<td>Carer felt the sessions with the advisor came too late in the patient’s illness</td>
<td>17 (21)</td>
</tr>
</tbody>
</table>

Interpretation

There are several possible reasons for our negative result. First, the intervention might have been too brief. Qualitative data collected after the death of the patient suggested that carers experienced some subjective benefit from the advisor visits, but also a sense that the intervention was too brief. Second, informal carers of patients with cancer might already have been receiving considerable input from specialist palliative care services and the care advisor’s extra help might have had little additional advantage; for example, our intervention might have had greater impact in cases of chronic cardiac failure where routine support for patients and carers is less well developed. Third, caring for a dying relative is extremely stressful and no amount of support is going to make it much better. Fourth, our intervention might simply have been wrongly planned and thus unhelpful; however, our qualitative results
do not support this possibility. Fifth, our outcome measures might have been insensitive to change or there was simply too much variance in the trial to allow detection of important change. Finally, given that nurses in the ‘treatment as usual’ group were aware of the nature of the trial and the intervention under evaluation, they might have tried harder to provide carer support. Given what we know about the workloads for nurses in these teams, we believe the last possibility is unlikely.

Implications and challenges for health service research

National guidance published since the start of this trial acknowledges the central role of families and carers in the informal care of cancer patients, particularly in the palliative phase (National Institute for Clinical Excellence, 2004). Each domain of care addressed by our intervention is referred to in the guidance, which contains a chapter specifically devoted to carer issues. Transitory benefits are not unusual in studies of brief service interventions and highlight a paradox in our concept of the effectiveness of such interventions (Bower et al., 2003): when a medication is seen to be effective in treating a medical or psychological condition, it is not regarded as ineffective if the condition relapses on withdrawal of that drug; in psychological or supportive interventions, however, loss of benefit when the intervention is withdrawn is often interpreted as indicating that the intervention is ineffective. Measuring change once the agent of change has been removed may be inappropriate in supportive care, especially near the end of life in rapidly progressive clinical and emotional circumstances. Our quantitative and qualitative results suggest that the carer advisor intervention was too brief and that ongoing help might have had accruing benefits. This would mirror the policy direction of earlier referral for palliative care services and contribute to more effective supportive care (National Institute for Clinical Excellence, 2004). Nevertheless, rigorous analysis of the effectiveness of care for patients and carers in trials such as this provides valuable evidence for service development in palliative and supportive care and responds directly to the recommendations and requirements set out in the NICE guidance (National Institute for Clinical Excellence, 2004).

ACKNOWLEDGEMENTS

We thank the carers who participated, staff of the specialist teams and carer advisors Andrea Beetson and Margaret Perkins. Participating specialist palliative care teams were: Royal Free; Camden and Islington; Whittington Hospital (Islington); North London Hospice; Harrow; Haringey Macmillan Team, and St Joseph’s Hospice, Hackney. We also thank Ula Nur and Gerhart Knener for their statistical advice and the Camden and Islington Mental Health and Social Care Trust and the Royal Free Hampstead NHS Trust for their support. The trial was funded by the former Cancer Research Campaign, now Cancer Research UK (grant number C1432/A179).

REFERENCES

Best practice when service users do not consent to sharing information with carers

Background Service users with psychosis may not consent to sharing information with carers. However, carers require access to relevant information to support them in their role.

Aims To inform clinical practice when service users withhold consent to share information with their carer.

Method Study data were derived from a synthesis of policy review (n=91), national survey (n=595) and individual interviews (n=24).

Results Key principles to guide information-sharing practices were identified. Service users highlighted confidentiality being guaranteed by consent processes. Carers suggested a ‘culture shift’ was required, with professionals trained to work with carers. Professionals emphasised mental capacity, professional judgement and the context of care. A best practice framework is proposed.

Conclusions An important distinction is between general information, which can always be shared without consent, and personal information, which is new to the carer and where consent needs to be considered. Clinical judgement is central to balancing conflicting ethical imperatives in this area.

Declarations of interest None.

Health professionals give information to carers to support them in their caring role (Department of Health, 2002), but the carer’s need for information must be balanced with the service user’s rights to privacy (Szmukler & Bloch, 1997). When carer involvement seems justified but the service user is withholding consent, professionals face an ethical dilemma between non-malfeasance (i.e. not doing harm, through failing to disclose) and beneficence (doing good, by respecting patient confidentiality) (Furlong & Leggatt, 1996; Beauchamp & Childress, 2001). This dilemma is especially complex in psychiatry (Arksey et al, 2002), where capacity to give informed consent may not be present, and where the relationship with the carer can in itself influence the course of the disorder (Raune et al, 2004). Where consent is withheld, professionals may still need information from the carer for a full assessment, and carers retain the right to have their own needs assessed (Department of Health, 2000).

There is a lack of research-based evidence in this area. We therefore completed a national study with the aim of developing a framework for best clinical practice where service user consent for sharing information with their carer is withheld.

METHOD

Data presented here were collected as part of a UK study assessing mental health information-sharing practices across the life course, including children and adolescents, adults of working age and older people. The final report containing a more detailed methodological description is available at http://www.sdo.lshm.ac.uk/sdo542003.html. The data presented here are focused on adults of working age with psychosis.

Study design

Data were synthesised from a consecutive policy review, a national survey of current practice and individual qualitative interviews. Each stage informed the next. The larger study also included facilitated stakeholder groups holding informal discussions and large group workshops, but these components did not address the subject of non-consent by people with psychosis, and so are not included in the analyses here. Multiple methods of data collection were used to allow triangulation – the use of different data sources to reach the results. Synthesising quantitative and qualitative methods is the right approach in an area characterised by a complex and often conflicting set of polarised beliefs from different groups: service users, carers and staff.

Three groups informed the design: a core research group (n=10), an expert panel who met three times (n=19) and a virtual panel who communicated electronically (n=14). Both panels comprised service users, carers, multidisciplinary professionals, carer support workers and academics. All groups contributed to the sequential stages of data collection, the analysis of the data and the development of emergent frameworks.

Setting and participants

Policy review

We collated policy documentation by surveying professional, service user and carer organisations, including the Mental Health Alliance (comprising 60 organisations) and the Care Programme Approach Association; directors of mental health trusts and social services in England; websites of professional and voluntary organisations; and international contacts.

Current practice survey

Electronic and paper surveys were developed in three different versions (service user, carer, and professional) and were piloted with relevant stakeholders (n=14). Each version comprised similar core questions (demographic details, experiences of information sharing, examples of good practice) and stakeholder-specific questions. The surveys were advertised through research partner organisations’ websites (n=13), group e-mails (n=7), promotion at conferences (n=5), targeted mailings (n=3), magazine advertisements (n=3), targeted promotion to Black and minority
ethnic groups (n=5) and individual contacts from existing databases (n=290).

**Individual qualitative interviews**

Quota sampling was used to maximise representativeness, by balancing location, gender, ethnicity of participants and experience. Two researchers piloted and used an in-depth interview schedule for professionals, carers and service users to assess involvement in mental health; how confidentiality and information-sharing practices have affected roles; where information-sharing has worked well; issues in information-sharing; and how information sharing could be improved. Interviews were conducted by telephone (except for five, which were conducted face-to-face on request), and lasted 25–90 min. Detailed manual notes were made during interviews and these were typed immediately afterwards to provide an accurate record of the discussion.

**Analysis**

Policies were categorised using an existing framework (Surrey-Wide Operational Partnership Group in Mental Health, 1999) by one researcher, with a subsample categorised by a second researcher to check the coding accuracy. Quantitative differences between survey groups were assessed using the Statistical Package for the Social Sciences Version 12 for Windows. Qualitative survey responses were analysed using content analysis (Weber, 1990). Qualitative interview data were stored and managed using NVivo version 2 (http://www.qsrinternational.com). The transcripts were analysed manually, following good practice principles to identify emergent themes (Silverman, 2001). Four researchers generated a preliminary coding framework, which was then applied to the full data in NVivo by one researcher, with reliability checks carried out by two other researchers. The data presented in this paper are extracted from two interview analysis themes: information-sharing principles and information-sharing strategies.

Data from the three sources were used to produce an emergent framework for information-sharing when consent is withheld by people with psychosis. It was developed by clustering recommendations to remove duplicates, prioritising those generated from more than one source and/or those more strongly present (either numerically in the quantitative data, or as strength-of-theme data in either qualitative source), separating them into different points in the information-sharing pathway (e.g. obtaining consent, exploring decisions with the service user), developing a draft emergent framework, and then refining through feedback from the core research group, expert panel and virtual panel.

**RESULTS**

**Policy**

The review identified 56 policies and 35 supporting documents, although many included conflicting statements. Only 5 policies provided practical guidance on how to appropriately share information. Eleven policies (20%) specifically addressed information-sharing with carers: 5 from National Health Service trusts, 5 from carer bodies and 1 from the National Institute of Mental Health in England. Statutory sector policies emphasised professional responsibilities: to assess mental capacity and where present to seek the service user’s consent to disclose personal information to the carer on a need-to-know basis; to review consent regularly; and to ensure accurate recording of information. Policies co-authored with carer groups also highlighted the use of advance statements to record preferences for crisis management (Henderson et al, 2004) and the promotion of inclusive approaches in respect of carers.

**National survey**

Survey participants comprised mental health service users (n=91), carers (n=329) and professionals (n=175). In the service user group 44 (48%) were male, 85 (93%) were White, 39 (42%) lived with their carer and 21 (23%) had been compulsorily detained in the previous year. In the carer group 64 (20%) were male, 309 (94%) were White, and 161 (49%) lived with the service user. Professionals included 66 (38%) psychiatric nurses, 29 (17%) social workers, 23 (13%) psychiatrists and 16 (9%) psychologists. Work settings comprised community teams (n=92; 53%), in-patient units (n=47; 27%), day care (n=26; 15%) and primary care (n=10; 6%). The combined sample provided 595 responses.

In the carer sample (n=329) the majority reported they were well supported in terms of access to ‘general information’. Ninety-two per cent understood the service user’s diagnosis and 69% had access to sufficient ‘general information’ which they gained from voluntary sector organisations (49%), carer support groups (47%), other carers (35%), the internet (32%) and community psychiatric nurses (CPNs) (30%).

A total of 186 carers (60%) had been given the opportunity to discuss the information they came across from a variety of sources with mental health professionals; these carers were significantly more likely to live separately from the service user (77% v. 87%, P=0.023) to be aged less than 61 years (74% v. 85%, P=0.023) and not be providing 24 h care, 7 days per week (86% v. 72%, P=0.002).

Considering personal information, 261 (82%) stated they needed access to personal information in order to care both effectively and safely. The types of personal information required included details of whom to contact in a crisis (79%), possible future treatment options (68%), likely progress of the service user’s mental health problems (65%), what the care plan says (59%), early signs of relapse (52%) and what treatments the service user is currently receiving (50%). In the previous year, 145 carers (46%) had received personal information to support their role and 171 (54%) had not. Table 1 shows that carers identified both professional practice and service user-based explanations for professionals not sharing personal information with them.

In the service user sample (n=91), more than half (59%; n=51) stated that their carers should have access to some personal information, with 47 (55%) reporting feeling ‘comfortable’ with their carer being involved and 47 (55%) believing carers should be offered separate time with professionals as a source of support. Service users and carers highlighted the absence of regular collection of consent to disclose authorisation – 51 (67%) of 76 service users with a named carer had not been asked to sign a disclosure consent form. Updating of consent authorisation was variable: 13 service users (14%) reported always being asked before information was disclosed, 19 (21%) sometimes, 20 (22%) rarely, 15 (16%) never and 24 (27%) did not know.

In the professional sample (n=175), half (50%) identified that their employer had a policy regarding sharing confidential information with carers. Among those with policies (n=88), 23% found these very helpful and 63% quite helpful. Professionals also identified why information is not shared with carers: 79% service user
withheld consent; 55% carers not accessible; 48% they had insufficient time; 42% not asked service user for consent; 29% service user unable to provide consent; 23% service user lacked capacity to provide consent. Table 2 presents the perspectives of each stakeholder group on whether and when information should be disclosed without consent. Views on potential problem resolution strategies are shown in Table 3.

Content analysis of qualitative data provided in the surveys identified principles to underpin good information-sharing practices. For service users (n=37) the key principles were establishing and maintaining better dialogue between all parties; routine collection of informed consent; positive attitude of professionals towards service users and carers; and flexible and creative approaches to information-sharing.

Table 1  Carer perspectives on reasons why professionals did not share personal information (n=171)

<table>
<thead>
<tr>
<th>Reason for breaking patient confidentiality</th>
<th>Personal information not shared with carer because:</th>
</tr>
</thead>
<tbody>
<tr>
<td>I have not asked for any</td>
<td>59 (35)</td>
</tr>
<tr>
<td>Service user did not provide consent</td>
<td>35 (21)</td>
</tr>
<tr>
<td>Service user was unable to give consent</td>
<td>9 (5)</td>
</tr>
<tr>
<td>Service user was not asked to provide consent</td>
<td>32 (19)</td>
</tr>
<tr>
<td>Patient confidentiality was given as the reason but without a supportive explanation</td>
<td>47 (28)</td>
</tr>
<tr>
<td>Patient confidentiality was given as the reason but with a supportive explanation</td>
<td>20 (12)</td>
</tr>
<tr>
<td>Don't know</td>
<td>18 (11)</td>
</tr>
</tbody>
</table>

Specific comments:
- For example, 'service user provides consent and then changes mind'; 'out of respect for service user would like to know but respect their wishes so don't persist'; 'carer is not next of kin'; 'language barriers'

For carers (n=107), good practice principles were identified as carer proactivity; recognition of carers' needs and rights; improved communication between all parties; improved professional attitudes towards carers; and collection of informed consent.

- 'Generally, any resolution has come about because of reminding, pushing, and demanding on our parts. This should not be.' (Carer 100)
- 'My son has frequently withdrawn consent for me to have information about him. His care team have gone to great lengths to explain to him exactly what they would tell me and why they feel I need to know it. Usually this works. When it doesn't (i.e. when he refuses) they revisit his decision regularly with him.' (Carer 99)
- 'A lot is down to individuals. I cannot tell you what a difference it has made since a new CPN has taken over care for my son. I can phone her at any time and she follows this up with appropriate action.' (Carer 115)

Individual qualitative interviews

The 24 interview participants comprised mental health service users (n=5), carers for people with severe mental illness (n=7), professionals (n=9) and carer support workers (n=3). Interviewees identified both governing principles and specific strategies to guide information-sharing. They emphasised the core role of individual judgement, relationships built upon openness, knowledge and trust, and the process

<table>
<thead>
<tr>
<th>Reason for breaking patient confidentiality</th>
<th>Service users (n=91)</th>
<th>Carers (n=326)</th>
<th>Professionals (n=175)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are there any occasions when information should be shared without service user consent?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Respondents (number stating there are occasions), n (%)</td>
<td>59 (65)</td>
<td>312 (96)</td>
<td>170 (97)</td>
</tr>
<tr>
<td>Reason for breaking patient confidentiality and sharing information without consent, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When the service user is very unwell</td>
<td>35 (59)</td>
<td>274 (88)</td>
<td>74 (44)</td>
</tr>
<tr>
<td>When the service user has agreed in advance</td>
<td>40 (67)</td>
<td>185 (59)</td>
<td>114 (67)</td>
</tr>
<tr>
<td>If people are worried about the service user's safety</td>
<td>37 (63)</td>
<td>237 (76)</td>
<td>132 (78)</td>
</tr>
<tr>
<td>If there are concerns about the service user harming other people</td>
<td>35 (59)</td>
<td>209 (67)</td>
<td>152 (90)</td>
</tr>
<tr>
<td>If carer lives with service user</td>
<td>14 (24)</td>
<td>180 (58)</td>
<td>40 (24)</td>
</tr>
</tbody>
</table>
of collecting informed consent. The importance and complexity of information-sharing decisions were highlighted by each stakeholder group.

'Possibly the most important thing about [information] sharing is once you have, you can't change things. You only get one chance so it has to be right.' (service user 1)

'I think the information that carers need varies from one case to another. Professionals need to talk to carers about confidentiality. I have never come across a carer who knows their rights and the procedures involved in confidentiality. Carers need to be given sufficient information to do their job well.' (carer 12)

'[Black and minority ethnic] families are very distrustful of services. Having been in mental health for so long I can understand why. There's something about stigma. They are frightened to share information in case they are pre-judged. My approach has been about acknowledging their anger and distrust, and not being defensive about the services we offer.' (carer support worker 2)

'Every party wants to have their voice heard.' (psychiatrist 1)

The service user interviews were dominated by one issue: the importance of patient confidentiality. All stressed how consent to disclose should be obtained before information is shared with carers. The requirement for consent was strongly linked to self-esteem, privacy, personal choice, independence, autonomy, general well-being and empowerment.

Carers accepted the service user's right to withhold consent, but (like service users) acknowledged this might have an impact on the standard of care they can provide. They emphasised the importance of information relevant to their support role, but did not need or want to know everything about the person supported. Carers viewed professionals as often lacking the confidence, empathy, skills, time and organisational backing to fulfil a carer support role alongside provision of health and social care treatment for the service user.

The perspectives of professionals on information-sharing were largely consistent with carers and service users in emphasising confidentiality; context of care (length of relationship, type of illness, stage of recovery, living arrangements, past history); mental capacity and consent; and establishing service user and carer confidence in professionals. In addition, professionals identified that they had a duty to assess risk, to avoid harm and to use professional discernment for decision-making. Appendix 1 provides illustrative quotations from each stakeholder group about patient confidentiality and stakeholder responsibilities.
Suggestions for good practice in information-sharing were made. Service user recommendations included effective communication, whereby all parties are kept informed of decisions; professional assessment of appropriate level of information-sharing; use of advance agreements; service improvements for both service users and carers, improving quality of care in mental health; and service user involvement in the local development of information-sharing guidelines and procedures. Carer recommendations included open and honest communication between stakeholders; improved recognition of the role of carers and their relevant knowledge; and a reduction in the perceived lack of engagement with or respect for those providing informal support. Professionals considered that an assessment of the carer’s and service user’s personal circumstances was a vital part of information-sharing with carers. Most professionals highlighted the benefits of bringing together parties to discuss care and treatment plans when an identified carer was involved. Community-based professionals were particularly aware of the importance of spending time separately with service users and carers.

**Framework for best clinical practice**

Based on the above results, a framework for best practice was developed for information-sharing with carers where professionals are dealing with service user non-consent. A key distinction to emerge was between two types of information: general and personal. General information is defined as information that supports carers in their role, without providing new details specific to the service user. In contrast, personal information is new and specific to the service user. Whether information is general or personal is case-specific: providing information about schizophrenia would be general information if the diagnosis were known by the carer but personal information if it were not known. By distinguishing between general and personal information the framework emphasises that support and some information can be provided to carers without patient confidentiality being broken. For example, carers might need support to deal with being excluded from an information-sharing dialogue, usually as a consequence of the service user developing increased independence. Equally, service users may need support to involve carers within prescribed boundaries, while revisiting their non-consenting decision regularly.

Two levels of action were identified: organisational and clinical. Organisational actions are recommendations that require organisational planning and implementation, (see Appendix 2). Clinical responsibilities are actions that individuals working in mental health can take to support service users and carers through information-sharing. An emergent framework for best practice by clinicians is shown in Fig. 1. Contrasting opinions were found on questions such as whether information can ever be shared without consent, and whether carers should meet professionals without the service users being present. Therefore the best practice framework emphasises the central role of clinical judgement in decision-making.

---

**Fig. 1  Framework for best clinical practice when consent is not given to share information with carers**

![Diagram](image-url)
DISCUSSION

This study found that there is not yet consensus in the UK as to best practice when service users refuse or give only partial consent to information-sharing with carers. However, it was possible to synthesise the identified good practice points into a clinically applicable framework for best practice. A central distinction to emerge was between general information and personal information.

Strengths and weaknesses of the study

Limitations include the self-selection of respondents, who may therefore not be representative. In particular, the sample included few people from Black and minority ethnic backgrounds. This means that any differences in perspective will not have been captured, which compromises the generalisability of the emergent framework. There were also difficulties in accessing policy documents. The identification of policy about carers was problematic given the lack of an agreed definition of a ‘carer’ – the term sometimes was used to include paid staff. The main strength of this study was the use of multiple sources of data. This was facilitated by active collaboration between researchers and relevant voluntary sector groups, with the intention of making participation in the study simple so as to minimise access barriers. Our multiple methods allowed for validation of themes, and the large set of respondents overall is also a positive feature.

The importance of support for carers

We heard from professionals, service users and carers about the importance of providing carers with timely and appropriate information. However, carers reported that in practice they experienced a lack of confidence, skills and organisational backing for staff to engage with them. This finding is consistent with other studies. For example, a qualitative study of 27 Australian primary carers found that lack of engagement by professional staff led to increased levels of distress among the carers, and left them feeling resentful and frustrated (Wynaden & Orb, 2005).

Carers are likely to benefit from good information-sharing practices. Research by the mental health charity Rethink found that carers with access to information and support were likely to rate fewer adverse affects from caring, including mental or physical health problems, financial pressures, and impact on family relationships (Pinfield & Corry, 2003). Furthermore, there are adverse clinical consequences for the service user where the carer is inadequately supported, since high expressed emotion in carers can predict relapse in psychosis (Raune et al, 2004).

Implications for health and social care organisations

Policy documents on information-sharing practices in mental health must be translated into practice on the ground with the support of local organisational structures. Professionals reported that when policies on information-sharing were well published in the workplace they were useful documents. In particular such documents were helpful when they provided practical guidance on how to resolve information-sharing dilemmas, and when they outlined the legal and ethical boundaries of professional responsibility. The best practice framework (see Fig. 1) is recommended as an evidence-based and multiprofessional approach suitable for incorporation into local policy. Dissemination of collaboratively authored local protocols based on this framework to staff and through carer groups might improve practice in this area.

Other factors to consider include organisational barriers, such as insufficient time to work with carers and a professional tendency to avoid working with carers. Carers’ rights to a needs assessment provided one route to support the carer. All these considerations informed the development of the organisational checklist (see Appendix 2). The checklist is intended to be used as an audit tool for mental health services to monitor and improve their organisational approaches to supporting information-sharing with carers.

Implications for healthcare professionals

Dealing with situations in which service users do not consent to information-sharing poses clinical and ethical dilemmas. On the one hand, the training of health professionals is oriented towards patient confidentiality rather than information-sharing, and they are concerned to keep the trust of the service user (British Medical Association, 1999). On the other hand, legal rights to confidentiality are not absolute (Department of Health, 1995; House of Commons, 1998a,b). Such complexity is not amenable to simple deterministic solutions, and clinical judgement must remain at the heart of decision-making in this area. This is particularly true when working with people with psychosis, whose capacity to provide informed consent can fluctuate, and the process – as opposed to the event – of providing consent requires continuity of care and strong therapeutic relationships.

The development of an empirically justified best practice framework is important for several reasons: first, to ensure that there is a shared understanding between service users, their carers and professionals about the centrality of service user consent, and the situations in which it can be justified to share information without consent; second, to highlight that carers can be supported even when consent is not given for sharing of personal information; and third, to support professional accountability in clinical practice.

For healthcare professionals, several strategies emerged. A change in attitude towards carers is indicated, to value more fully the carer’s role. Ongoing communication with both patient and carer is vital, covering the aspects in Boxes A–D in Fig. 1. Fluctuating mental capacity – a problem not restricted to psychiatry (Raymont et al, 2004) – means that advance statements and regular review of consent should be routine practice.

General information that builds on the carer’s existing knowledge can always be shared without consent, so the distinction between general and personal information needs to be understood by service users, carers and professionals. Clinical skills are also needed to identify what the carer already knows before any information is disclosed: what would be general information (which can be shared without considering consent issues) and what would be personal information (where consent needs to be considered) for this carer? Supportive communication with carers is desirable, even when consent for personal information-sharing has not been given. Clinical strategies might include viewing non-consent as a positive indicator of recovery and increased autonomy; emphasising that the refusal of consent is the current stance of the service user which will be regularly reviewed by the clinician; and providing as much general information as possible.
Future research

The benefits and difficulties for both carers and service users of dealing actively with situations in which information-sharing consent is withheld could be investigated. There is evidence that when the need of carers for timely and appropriate information to fulfill their role is met, they experience lower carer burden (Pinfold & Corry, 2003). Interventions targeted at reducing expressed emotion have been developed to support families, including carer education and psychosocial services (Barrowclough et al., 1999; Szukler et al., 2003). In part, these interventions involve the provision of information to help the carer interpret the service user’s behaviour in ways that do not lead to criticism or emotional overinvolvement. It is plausible that better information-sharing is one of the active ingredients of the intervention. Future research should investigate whether the best clinical practice framework leads to a more positive impact of caring on the carer and reduces relapse rates by lowering expressed emotion.

Sharing information with carers is a complex process which is increasingly an international focus of policy (Department of Family and Community Services, 1999) and research (Marshall & Solomon, 2000). The joint interests of service users and their carers are best balanced when clinical judgement about the individual context remains central to balancing the implications of sharing or of not sharing information.

ACKNOWLEDGEMENTS

This research was a collaboration between the Institute of Psychiatry (King’s College London) and Rethink. We thank the research participants, panel members, and all the organisations who supported the study. The study was funded by the NHS Service Delivery and Organisation Research and Development Programme. The views expressed are those of the authors.

APPENDIX I

Stakeholder views on patient confidentiality in practice

Service user perspective

1 can see it [patient confidentiality] is a difficult issue for carers, but there is also the problem of involvement of the carer if the service user finds it unhelpful or distressing. I know there is a lot of talk about abuse but it doesn’t have to be that extreme. Whether the relationship is happy, healthy, harmonious at a particular traumatic time or not, if carer involvement is not welcomed this needs to be addressed.’ (service user 4)

1 definitely think that the service user should be the one to have the final say on how much information to share if they have the capability.’ (service user 1)

1 am fond of advance directives, I think they are a very good idea. Everyone should have a statement for when or if they are ill again. This would mop up all the issues, and should help a carer’s involvement as well.’ (service user 5)

Carer perspective

1 have a reasonable relationship with my son, I always seek his permission first before I look at his care plan or medical notes. I agree with patient consent as I would not be happy for my son to know information about me without my consent.’ (carer 12)

‘When the service user doesn’t want the carer involved they need to know that when they leave hospital and expect care, the carer can’t provide the best service without the relevant information.’ (carer 13)

‘If the carer don’t have the full knowledge this is very dangerous. You cannot care fully unless you have full knowledge, mistakes will be made and this could be harmful.’ (carer 14)

Professional perspective

1 think there are times when you really want to tell relatives about the patient’s behaviour. Sometimes you have to say, ‘I can’t let you go home [from hospital] unless I tell them this.’ (consultant psychiatrist 26)

‘The capacity of the patient to make decisions is the key factor in determining what information is shared with others.’ (psychiatrist 39)

‘Risk drives information sharing in mental health.’ (social work manager 33)

‘Often my role in hospital would be explaining rights to people, giving a factual explanation about legal aspects rather than sensitive information. But if they were asking me to confirm something I would still be guarded and not to be seen to be confirming to avoid breaches of confidentiality and the patient coming back to me. A lot of health staff have this attitude.’ (approved social worker 28)

APPENDIX 2

Organisational actions to support information-sharing with carers

☐ Positive approach towards working with carers, including a programme of support, e.g. carers’ assessments, information resource packs, carer-involvement opportunities.

☐ Easy access to information-sharing policy document and implementation guidelines which have been developed in partnership by mental health professionals, service users and carers.

☐ ‘Know your rights’ resources available for carers in a language of their choice

☐ Training for mental health professionals on how to work effectively with carers, including guidance on information-sharing.

☐ Specific training for dealing with situations where the service user withholds full or partial consent to share information, e.g. using the best clinical practice framework (Fig. 1).

☐ Acknowledgement of the complexity of information-sharing decisions and support structures to assist staff in applying professional discernment.

☐ Awareness of the culturally sensitive approaches required to support carers from diverse communities.

☐ Organisational validation of information-sharing, with carers being part of the clinical role.

☐ Develop and audit use of a system for collecting and reviewing ‘patient consent to disclose’ documentation and advance statements.

☐ Develop and audit use of a system for recording what information has already been shared with carers, so that the distinction between general and personal information can be maintained.

☐ Continuity of care for service users (and carers) which supports information-sharing practices through development of a strong therapeutic relationship and in-depth knowledge of caring context.

☐ Promotion of effective, open and honest communication between professionals, carers and service users.

☐ Carer’s needs assessment process includes consideration of information needs.

REFERENCES


British Medical Association (1999) Confidentiality and Disclosure of Health Information. BMA.

Department of Family and Community Services (1999) Mental Health Promotion and Prevention National Action Plan. DFCS.


Furlong, M. & Leggatt, M. (1994) Reconciling the patient’s right to confidentiality and the family’s need to...
Under Pressure: The Impact of Caring on People Supporting Family Members with Mental Health Problems: a Survey of NAMI Members.

Prevalence of mental incapacity in medical inpatients and associated risk factors: cross-sectional study.

Effect of joint crisis plans on use of compulsory treatment in psychiatry: single blind randomised controlled trial.

Impact of caring on people supporting family members with severe mental illness: a survey of NAMI members.

Analysis for Analysing Talk, Text and Interaction

Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction

An exploratory randomised controlled trial of a support programme for carers of patients with a psychosis.

Impact of patient confidentiality on carers of people who have a mental disorder.

SHARING INFORMATION WITH CARERS
Cancer risk among parents and siblings of patients with schizophrenia

I. LEVAV, I. LIPSHITZ, I. NOVIKOV, I. PUGACHOVA, R. KOHN, M. BARCHANA, A. PONIZOVSKY and H. WERNER

Background  A reduced risk of cancer has been noted among people with schizophrenia. Given that genetic causes have been proposed as an explanation of this finding, one would expect that the risk of cancer among first-degree relatives would be equally reduced.

Aims  To investigate the risk of cancer among the biological parents and full siblings of people receiving in-patient care for schizophrenia.

Method  Linkage analysis was conducted between national population, psychiatric and cancer databases. Standardised incidence ratios for all cancer sites were calculated by comparing the incident rates among first-degree relatives with national incidence rates.

Results  A reduced cancer risk was found across all groups examined. Among parents, whose numbers were adequately large, the findings reached statistical significance. For index cases and siblings – a markedly younger population – only a trend was elicited.

Conclusions  The genetic hypothesis – namely, the presence of a gene with the dual effect of reducing the cancer risk and disrupting neurodevelopment – is a plausible explanation for these findings.

Declaration of interest  None. Funding detailed in Acknowledgements.

The reduced risk of cancer among patients with schizophrenia is a research puzzle (Jablensky & Lawrence, 2001). Indeed, this finding – which does not apply to other psychiatric disorders (Carney et al, 2004) – is totally unexpected, since people with schizophrenia have additional health risks, such as heavy smoking (Dalack et al, 1998), other unhealthy lifestyle habits (Brown et al, 1999) and often medical neglect (Craddock-O’Leary et al, 2002). Several recent studies (e.g. Barak et al, 2005; Grinshpoon et al, 2005), although not all (Goldacre et al, 2005), have replicated this puzzling result. Among the hypotheses proposed to explain the reduced risk is that the p53 gene generates, through apoptosis, the dual effect of disrupting neurodevelopment and reducing the risk of cancer (Catts & Catts, 2000; Park et al, 2004; Yang et al, 2004; Cui et al, 2005).

Confirmation of the genetic hypothesis would require that the first-degree relatives of patients with schizophrenia be found to have a similar reduced cancer risk, compared with suitable populations. The epidemiological evidence remains inconclusive: both a lower risk (Lichtermann et al, 2001) and no risk differential (Dalton et al, 2004) have been found. The Finnish study (Lichtermann et al, 2001) reported a lower risk among fathers, with a standardised incidence ratio (SIR) 0.93 (95% CI 0.90–0.96), and also among mothers (SIR=0.89, 95% CI 0.86–0.92), brothers (SIR=0.85, 95% CI 0.76–0.93) and sisters (SIR=0.92, 95% CI 0.84–0.99). Curiously, the cancer risk for patients with schizophrenia in that study was higher than the cancer risk for the general reference population (SIR=1.17, 95% CI 1.09–1.25). In contrast, when the Danish researchers compared the parents of patients with schizophrenia with the general population, they found not only no overall risk differential, but a 20% increased risk of lung cancer among the mothers (Dalton et al, 2004). These results were contested by Lichtermann (2005), who argued that the observation period in the Danish study did not cover the whole lifetime of the parents, and that there was insufficient empirical evidence in the published literature for the ‘healthy parent’ effect as an explanation for the reduced cancer risk found in the parents of patients with schizophrenia in his own study. In their reply, Dalton et al (2005) supported their thesis by reporting that they did find a reduced cancer risk in parents, but only when the comparison population included childless adults as in the Finnish study.

Like Denmark and Finland, Israel also has a continuously updated population register as well as registers for cancer cases and psychiatric hospital admissions. This has enabled us to replicate the attempt to confirm a reduced cancer risk among first-degree relatives, a project that continues our earlier study in which we found a lower risk of cancer in patients with schizophrenia (Grinshpoon et al, 2005).

METHODS
Identification of patients with schizophrenia
The following service procedures and policies facilitated the identification of people with schizophrenia for inclusion in our study

(a) primary care physicians usually refer patients with severe psychiatric disorders to specialist services; 
(b) direct and relatively easy access to these services is also freely available on a drop-in basis (Levav & Grinshpoon, 2004); 
(c) the Israel Defense Forces’ medical examinations prior to recruitment (both genders) and during reserve duties (men only) serve as a universal screening procedure which is followed by a psychiatric assessment whenever deemed necessary; 
(d) patients with psychosis are usually admitted to in-patient care; 
(e) Israelis rarely seek hospitalisation for a schizophrenic disorder abroad.

This was confirmed in a community-based study of an Israeli-born birth cohort, which found that all respondents identified with schizophrenia in the cohort were known to the in-patient psychiatric services (Levav et al, 1993).
We used the nation-wide psychiatric case register to identify patients with schizophrenia for our sample. This 50-year-old register is mandated by law to maintain a cumulative record of all psychiatric hospital admissions (Lichtenberg et al., 1999; Mental Health Services, Department of Information and Evaluation, 2004). The identity number used to record all patient movements is the same as that used by the national population register and the cancer register. The psychiatric case register also provided the patients’ diagnoses upon admission and discharge as well as socio-demographic information. Diagnoses follow the ICD–10 (World Health Organization, 1992); those made prior to the introduction of ICD–10 have been updated. A test of the agreement between research diagnoses and those recorded in the register found a satisfactory match (Rabinowitz et al., 1994; Weiser et al., 2005). Although the recording of cases is complete for Jewish Israelis, this is not so for the Arab Israeli minority, particularly women, who use psychiatric in-patient services considerably less than Jewish Israelis (Mental Health Services, Department of Information and Evaluation, 2004). To avoid biasing the sample, the Arab Israeli minority was excluded.

Identification of cancer cases

The Israel National Cancer Registry (http://www.health.gov.il/ncr) was established in 1960 (Freedman et al., 2001). Reporting has been mandatory since 1982 for all medical facilities, public and private. The registry also collects data on cancer deaths from district health authorities and from the Ministry of the Interior’s population register. As in the psychiatric case register, the information is organised using a personal identity number. Information completeness exceeds 95%. Continuous efforts are made to improve reporting and accuracy (Fishler et al., 2002). Multiple tests, as prescribed by the International Agency for Research on Cancer (Parkin et al., 2002), for example the percentage of cases with morphological verification, the mortality to incidence ratio and the percentage of cases ascertained by death certificate only, are conducted regularly to check data quality.

The number of residents with cancer who seek diagnosis and treatment abroad is probably small, since medical services in Israel are free and adequate. However, the exact number of those going abroad for care, as well as the number of Israelis living abroad who might have returned home to avail themselves of free medical care, is unknown. Neither figure is likely to be large.

Linkage procedure

The two case registers and the population register can be linked by means of the personal identity number. This identity number is supplemented automatically with the person’s full name, gender, date of birth and place of origin and the father’s first name, to ensure a reliable linkage. The process of identification and linkage comprised four steps. First, through the psychiatric case register, we identified a cohort of persons discharged from their last or only in-patient episode with a diagnosis of schizophrenic disorder (ICD–10 codes F20–29). To do this we made use of an existing database of family-linked files of individuals with schizophrenia. In this database ($n=6132$) almost all patients with schizophrenia (about 95%) were born in Israel in the years 1970–1988, either to immigrants or to native-born parents. The few who were born abroad were 5 years old or less on immigration. Second, we identified their biological parents (mothers $n=5756$; fathers $n=5741$) and their full siblings (brothers $n=9445$; sisters $n=9846$) using the population register’s computerised family files. Third, we ran the files of both parents and siblings through the psychiatric case register to identify all patients discharged from their last or only in-patient episode with a diagnosis of schizophrenic disorder (F20–29), as noted above (fathers $n=224$; mothers $n=393$; brothers $n=508$; sisters $n=354$). Fourth, the files of index cases, parents and siblings were run through the cancer register to locate cancer cases. This four-stage process generated three sub-populations:

(a) index cases with and without cancer;
(b) parents never hospitalised for schizophrenia, or hospitalised for schizophrenia at least once, with and without cancer;
(c) siblings never hospitalised, or hospitalised at least once for schizophrenia, with and without cancer (Table 1). Dates of death and emigration were obtained from the population registry.

The case registers (psychiatric and cancer) owned and maintained by the Ministry of Health are administered under strict legislatively defined procedures. To preserve confidentiality, linkages are made by methods that ensure researchers are not given files with the individual’s real identity number. Internal review board approval to build the study family database was obtained from Butler Hospital, Providence, Rhode Island, USA.

Statistical analysis

The cancer incidence rates in the above three sub-populations were compared with the rates in the Jewish Israeli population using standardised incidence ratios (and their 95% confidence intervals), defined as the ratio of the observed to the expected number of cancer cases. The expected number of cases during the observation period was calculated by gender, area of origin (Africa–Asia, Europe–America or Israel) and age. Period-specific cancer incidence rates were used.

The person-years of exposure to cancer risk were defined as follows: index cases, from date of birth or date of immigration; parents, from date of immigration or from 1960, whichever was the later; and siblings, from date of birth or immigration or from 1960, whichever was the later (Table 2). The observation ended on death, diagnosis of cancer, emigration, or at the end of 2003.

Analysis by cancer site, gender and area of origin was conducted when the number of cancer outcomes in the sub-population was 10 or more. The test was performed for all index cases with a schizophrenic

<p>| Table 1 Cancer cases among patients with schizophrenia (index cases) and their first-degree relatives |
|---------------------------------------------|----------------|----------|</p>
<table>
<thead>
<tr>
<th></th>
<th>Group size</th>
<th>Cancer cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index cases</td>
<td>Group size</td>
<td>Cancer cases</td>
</tr>
<tr>
<td>Male</td>
<td>4073</td>
<td>28</td>
</tr>
<tr>
<td>Female</td>
<td>2059</td>
<td>14</td>
</tr>
<tr>
<td>Total</td>
<td>6132</td>
<td>42</td>
</tr>
<tr>
<td>Parents</td>
<td>Group size</td>
<td>Cancer cases</td>
</tr>
<tr>
<td>Male</td>
<td>5741</td>
<td>501</td>
</tr>
<tr>
<td>Female</td>
<td>5756</td>
<td>549</td>
</tr>
<tr>
<td>Total</td>
<td>11 497</td>
<td>1050</td>
</tr>
<tr>
<td>Siblings</td>
<td>Group size</td>
<td>Cancer cases</td>
</tr>
<tr>
<td>Male</td>
<td>9846</td>
<td>79</td>
</tr>
<tr>
<td>Female</td>
<td>9445</td>
<td>93</td>
</tr>
<tr>
<td>Total</td>
<td>19 291</td>
<td>172</td>
</tr>
</tbody>
</table>
Table 2  Exposure to cancer risk, calculated for age group and gender, for index cases and their first-degree relatives, excluding those diagnosed with schizophrenia

<table>
<thead>
<tr>
<th>Age, years</th>
<th>Exposure to cancer risk, person-years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Parents</td>
</tr>
<tr>
<td></td>
<td>Male</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0–4</td>
<td>540</td>
</tr>
<tr>
<td>5–9</td>
<td>2924</td>
</tr>
<tr>
<td>10–14</td>
<td>8043</td>
</tr>
<tr>
<td>15–19</td>
<td>14 178</td>
</tr>
<tr>
<td>20–24</td>
<td>19 187</td>
</tr>
<tr>
<td>30–34</td>
<td>25 331</td>
</tr>
<tr>
<td>35–39</td>
<td>26 316</td>
</tr>
<tr>
<td>40–44</td>
<td>26 609</td>
</tr>
<tr>
<td>45–49</td>
<td>25 743</td>
</tr>
<tr>
<td>50–54</td>
<td>22 332</td>
</tr>
<tr>
<td>55–59</td>
<td>15 900</td>
</tr>
<tr>
<td>60–64</td>
<td>9 243</td>
</tr>
<tr>
<td>65–69</td>
<td>4 720</td>
</tr>
<tr>
<td>70–74</td>
<td>1 767</td>
</tr>
<tr>
<td>75+</td>
<td>650</td>
</tr>
<tr>
<td>Total</td>
<td>226 648</td>
</tr>
</tbody>
</table>

Table 3  Standardised incidence ratios of cancer among index cases and first-degree relatives, compared with the general population, 1960–2004

<table>
<thead>
<tr>
<th>Cases</th>
<th>SIR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expected</td>
</tr>
<tr>
<td>Index cases</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>31.96</td>
</tr>
<tr>
<td>Female</td>
<td>17.98</td>
</tr>
<tr>
<td>All parents</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>597.69</td>
</tr>
<tr>
<td>Asia–Africa</td>
<td></td>
</tr>
<tr>
<td>Europe–America</td>
<td>149.39</td>
</tr>
<tr>
<td>Israel</td>
<td>168.29</td>
</tr>
<tr>
<td>Female</td>
<td>634.85</td>
</tr>
<tr>
<td>Asia–Africa</td>
<td></td>
</tr>
<tr>
<td>Europe–America</td>
<td>140.99</td>
</tr>
<tr>
<td>Israel</td>
<td>238.01</td>
</tr>
<tr>
<td>Parents without schizophrenia</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>577.94</td>
</tr>
<tr>
<td>Female</td>
<td>592.00</td>
</tr>
<tr>
<td>All siblings</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>85.59</td>
</tr>
<tr>
<td>Female</td>
<td>111.95</td>
</tr>
<tr>
<td>Siblings without schizophrenia</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>83.09</td>
</tr>
<tr>
<td>Female</td>
<td>107.13</td>
</tr>
</tbody>
</table>

SIR, standardised incidence ratio.

RESULTS

Both parents, including those with schizophrenia, had reduced cancer risk: mothers, SIR = 0.86 (95% CI 0.79–0.94); fathers, SIR = 0.84 (95% CI 0.76–0.91). The cancer risk for parents after excluding those hospitalised with schizophrenia remained almost unchanged: mothers, SIR = 0.87 (95% CI 0.79–0.95); fathers, SIR = 0.85 (95% CI 0.77–0.92). Lower ratios were found for gender-concordant pairs of offspring and parent: female index case and mother, SIR = 0.74 (95% CI 0.63–0.86); male index case and father SIR = 0.81 (95% CI 0.73–0.90). For gender-discordant pairs, the ratios were: father of female index case, SIR = 0.89 (95% CI 0.76–1.02); mother of male index case, SIR = 0.93 (95% CI 0.84–2.02).

Standardised incidence ratios among parents by restricted or extended type of schizophrenia showed no statistically significant difference. By all areas of origin, mothers and fathers showed a statistically

disorder (F20–29), and separately for 'restricted' cases (F20, F22, F25) (e.g. paranoid schizophrenia) and 'extended cases' (F21, F23, F24, F28, F29) (e.g. acute schizophrenia-like psychotic disorder), assuming a differential weight for the imputed genetic component in the aetiology of each group of disorders.

Since we could not know who in the general population sample was a parent, in order to check the 'healthy parent' hypothesis (Dalton et al, 2004) we performed a sensitivity analysis restricted to women, taking into account that in the latest national census about 7.0% of all women aged 35 or over had no children. For all cancers, we assumed that childless women had increased relative risks of 1.0, 1.5 and 2.0 as compared with mothers. The corrected expected number of cases among mothers, $E_c$, was calculated as $E_c = E/(0.93 + k 0.07)$ where $E$ is calculated using the official age group x gender x area of origin x period-specific cancer incidence rates, $E_c$ is the corrected expected number of cases and $k$ = 1.0, 1.5 and 2.0. The standardised incidence ratio (SIR) and the corrected ratio (SIRC) and their 95% confidence intervals were calculated using Poisson regression, with $E$ and $E_c$ as the respective offset. All calculations were performed using SAS version 9.1.3 for Unix software.
significant risk reduction (Table 3). With regard to the four leading cancer sites among parents, two sites showed reduced risks: breast cancer, expected 239, observed 204, SIR = 0.85 (95% CI 0.74–0.97); prostate cancer, expected 60, observed 46, SIR = 0.77 (95% CI 0.55–0.99). For lung and colorectal cancers the ratios were lower than unity in both parents, but failed to reach statistical significance. Index cases and siblings had a non-significant risk reduction, probably because of their relatively younger age. Index cases: females, SIR = 0.78 (95% CI 0.37–1.19), males, SIR = 0.88 (95% CI 0.55–1.20). Siblings: brothers, SIR = 0.92 (95% CI 0.72–1.13); sisters, SIR = 0.83 (95% CI 0.66–1.00). The cancer risk among siblings, after excluding those with schizophrenia, remained almost unchanged: brothers, SIR = 0.93 (95% CI 0.72–1.13); sisters, SIR = 0.83 (95% CI 0.66–1.00) (Table 3).

The sensitivity analysis performed on mothers to check for the ‘healthy parent’ (mother) effect did not alter the findings; mothers of people with schizophrenia retained a significantly reduced cancer risk (Table 4).

**DISCUSSION**

The finding of a reduced risk for cancer was consistent across all groups, but particularly marked with respect to parents, and especially for gender-concordant parent–offspring pairs. For index cases and siblings, the risk differential did not reach statistical significance, most probably owing to lack of statistical power. In an earlier study of ours based on a much larger cohort (Grinspoon et al., 2005), we did find a statistically significant reduced cancer risk for index cases. With respect to index cases, the problem in the study reported here might also be compounded by premature mortality (Goff et al., 2005).

With regard to specific cancer sites, the test in this study for breast cancer was the most adequate, since the relatively large sample size generated enough statistical power to demonstrate a statistically significant reduction. For other cancer sites, except for prostate cancer, the results were less definite.

This study has several methodological limitations and strengths. A first possible limitation was pointed out by Dalton et al. (2004, 2005) – namely, that to control for the ‘healthy parent’ effect, parents of children without schizophrenia constitute a better reference group than the general population. We doubt, however, that their objection to the Finnish study (Lichtermann et al., 2001) for using the general population as a reference group is relevant in our case, because of our different demographic patterns. Whereas in Denmark the total fertility rate was 1.7 and only 24.0% of households had three or more members (Danmark Statistik, http://www.dst.dk), among Jewish Israelis, the respective figures reached 2.6 and 60% (Central Bureau of Statistics, http://www.cbs.gov.il). In 1997 the Israeli census found only 7.2% of women aged 35 and older to be childless, whereas in Denmark the proportion was nearly 18.0%. Moreover, our *ad hoc* and conservative sensitivity analysis (Lichtermann, 2005) generated no indication of the healthy parent (in our case, mother) effect.

A second factor that might undermine our conclusions is that whereas patients with schizophrenia are more frequently found in the low socio-economic groups (Dohrenwend et al., 1992), cancer in Israel is more frequent in the higher socio-economic groups (Israel National Cancer Registry, 1990; Israel Center for Disease Control, 1998). Curiously, socio-economic status as a confounding variable has seldom been discussed in the research literature. In our study, however, it may not constitute a problem since parents of patients with schizophrenia are not found in the lower socio-economic groups at a higher proportion than in the general population (Goldberg & Morrison, 1963; Byrne et al., 2004). Nor may socio-economic status be problematic with regard to the healthy siblings. Nevertheless, we checked for the socio-economic status effect by grouping our parents by ethnic origin. In Israel this is a reasonable proxy measure for socio-economic status, since Israelis born in Asia or Africa generally have a lower socio-economic status than their counterparts born in Europe or America (Schwartz et al., 1991). We found that parents of patients with schizophrenia had a decreased cancer risk, regardless of their continent of origin.

Third, our index cases were diagnosed by clinicians who, in the nature of their work, do not attempt to achieve a research-standard diagnosis. To increase validity we extended the period of observation by using the discharge diagnosis from the last or only in-patient episode. However, if our final sample included people diagnosed as having schizophrenia but who in fact had other disorders, these false-positive cases would only buttress our results. Fourth, our finding that more mothers than fathers had schizophrenia might suggest a sampling bias, because in the psychiatric case register we found considerably more men than women hospitalised with schizophrenia, for both index cases and siblings. Conceivably, there is a greater likelihood that women marry and have children before they require hospital treatment, given that their mean age at disease onset is higher than that of men. Fifth, although we did not have access to lifestyle issues highly associated with cancer, such as smoking, we doubt that health-promoting behaviour is any more frequent among parents of offspring with
schizophrenia than among parents in the reference population. Finally, early death among the parents of people with schizophrenia cannot be ruled out as a confounding factor. However, there is no evidence to suggest that early mortality is linked to cancer risk.

Our register-based study had two particular methodological strengths. Both databases provided us with fairly complete and accurate data; furthermore, we repeated the database linkage procedures to make sure that the matches were correct. Additionally, the differential cancer risk among ethnic groups in Israel and the closeness to the rates found in the USA for frequent cancer sites (Freedman et al., 2006) reduce the possible group-specificity of our results.

In sum, if the strengths of our study outweigh its limitations, we are confident that we detected a consistently lower risk of cancer among the parents of people with schizophrenia, and a trend among index cases and siblings. Several hypotheses have been proposed to explain these findings (Mortensen, 1994), among them, the dual role of a tumour suppressor gene such as p53 (Catts & Catts, 2000; Ni et al., 2005). Tumour suppressor p53 has been identified as the most frequently mutated gene in human cancer. The gene is usually activated following DNA damage or other types of cellular insult. Activated p53 may block progression through the cell cycle or, alternatively, may lead to apoptosis, and in this way prevent the accumulation and transmission of genetic damage to daughter cells. In the specific context of neural development, it has been shown that p53 fulfils an important role in the normal apoptosis-driven neurogenesis of various brain structures. In accordance with this notion, it has been postulated that increased p53 levels in schizophrenia patients may increase cell death in potentially critical areas of the central nervous system. This hypothesis is consistent with the multiple structural defects in cerebral anatomy reported in schizophrenia. On the other hand, mutations in the p53 gene (the most common cause for p53 accumulation) may be associated with genomic instability in particular tissues and a greater probability of organ-specific neoplastic transformation. Taken together, these features suggest that p53 might constitute a dual effector with important roles in the aetiology and development of schizophrenia and, concomitantly, in cancer protection or promotion.

Although the genetic hypothesis has been challenged (Jablensky & Lawrence, 2001), based on our findings this hypothesis constitutes an attractive tentative explanation that deserves further research. An advantage of this line of research would be that by probing into the purported link between schizophrenia genes and cancer, we might learn something about schizophrenia by studying cancer (Kalkman, 2006).

ACKNOWLEDGEMENTS

The National Association for Research on Schizophrenia and Affective Disorders partly funded this study. A.P. was partially supported by the Ministry of Immigrant Absorption, Israel. The funding agencies had no role in the study’s design. Professor M. Shani, Professor A. Weizman and Dr A. Grinshpoon smoothed the way for this study and their help is gratefully acknowledged. Mr N. Steigman edited this paper.

REFERENCES


Neural correlates of the misattribution of speech in schizophrenia

PAUL ALLEN, EDSON AMARO, CYNTHIA H.Y. FU, STEVEN C.R. WILLIAMS, MICHAEL J. BRAMMER, LOUISE C. JOHNS and PHILIP K. McGUIRE

Background  The neurocognitive basis of auditory verbal hallucinations is unclear.

Aims  To investigate whether people with a history of such hallucinations would misattribute their own speech as external and show differential activation in brain areas implicated in hallucinations compared with people without such hallucinations.

Method  Participants underwent functional magnetic resonance imaging (fMRI) while listening to pre-recorded words. The source (self/non-self) and acoustic quality (undistorted/distorted) were varied across trials. Participants indicated whether the speech they heard was their own or that of another person. Twenty people with schizophrenia (auditory verbal hallucinations n=10, no hallucinations n=10) and healthy controls (n=11) were tested.

Results  The hallucinator group made more external misattributions and showed altered activation in the superior temporal gyrus and anterior cingulate compared with both other groups.

Conclusions  The misidentification of self-generated speech in patients with auditory verbal hallucinations is associated with functional abnormalities in the anterior cingulate and left temporal cortex. This may be related to impairment in the explicit evaluation of ambiguous auditory verbal stimuli.

Declaration of interest  None.

Audiory verbal hallucinations are a cardinal feature of schizophrenia but their neurocognitive basis is unclear. Theoretical accounts proposed that such hallucinations result from a breakdown in the monitoring of the intention to generate inner speech, through a loss of the ‘efference copy’ associated with the generation of verbal material. This efference copy serves to inform an internal monitor of forthcoming action and may thus help to distinguish self-generated from externally generated verbal material (Blakemore et al., 2002). In the absence of this signal, inner speech may thus be misidentified as ‘alien’ and perceived as externally generated voices (Feinberg, 1978; Frith & Done, 1988). Hallucinations have therefore been conceptualised as resulting from a breakdown in the systems monitoring the current intention to make actions (Frith & Done, 1988).

However, monitoring can also occur at the level of the conscious evaluation of the verbal output (Levelt, 1983) when speakers hear their own voice. Impairment at this level may also lead to the erroneous misattribution of self-generated speech. When patients with schizophrenia who are prone to auditory verbal hallucinations speak and hear an acoustically distorted version of their own voice they tend to misidentify their own speech as being that of somebody else (Johns & McGuire, 1999; Fu et al., 2001; Johns et al., 2001). Although this impairment is consistent with a loss of efference copy, it could equally result from a problem with the conscious evaluation of auditory verbal feedback (Allen et al., 2004).

The purpose of our study was to use functional magnetic resonance imaging (fMRI) to examine the brain regions involved in the conscious appraisal of speech in people with schizophrenia who were and were not prone to auditory verbal hallucinations. The subjective experience of these hallucinations in schizophrenia is associated with activation in the inferior frontal, anterior cingulate and temporal cortex (McGuire et al., 1993; Shergill et al., 2000b). Furthermore, the processing of verbal material in people who are prone to such hallucinations has been associated with differential engagement of these regions relative to people with schizophrenia who do not experience hallucinations and controls (McGuire et al., 1995; Shergill et al., 2003) particularly, in the temporal cortex (Fu et al., 2001). We tested the hypothesis that in people with auditory verbal hallucinations the appraisal of speech would be associated with the differential engagement of temporal, prefrontal and anterior cingulate cortices. More specifically, we tested the prediction that external misattributions in people with these hallucinations would be associated with altered activation of the temporal cortices.

METHOD

Participants  All participants were right-handed men who spoke English as their first language and had no history of hearing problems. The study had local research ethics committee approval and all participants gave informed consent.

Control group  A control group of 11 healthy volunteers was recruited from the local community through advertisements. Applicants with a history of medical or psychiatric disorder, a drug or alcohol use problem, a family history of psychiatric disorder, or who were receiving medication were excluded. Their mean age was 28 years and their mean IQ, estimated with the National Adult Reading Test (NART; Nelson & O’Connell, 1978), was 115 (see Table 1).

Patient groups  All patients met DSM–IV criteria for schizophrenia (American Psychiatric Association, 1994) and were recruited through the South London and Maudsley National Health Service Trust. Clinical teams were systematically contacted with a request to identify patients with schizophrenia who either had prominent and current auditory verbal hallucinations, or had no current or previous history of such hallucinations. This information was corroborated by careful review of the patients’ clinical records. Potentially eligible patients were then
approached by the investigators and assessed using the Scale for the Assessment of Positive Symptoms (SAPS; Andreasen, 1984a), the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1984b), the Calgary Depression Scale (Addington et al., 1990) and the NART.

The hallucinator group (n=10) comprised patients who scored ≥ 3 on the SAPS auditory hallucination item (clear evidence of voices and that they had occurred in the past week). All of these patients had a documented history of auditory verbal hallucinations. Patients in this group were also experiencing other positive symptoms, particularly delusions, and had low levels of negative symptoms (see Table 1). Nine of this group were in hospital at the time of testing and two were receiving out-patient treatment. None reported hallucinations during the fMRI scanning procedure.

The non-hallucinator group (n=10) was composed of patients who were not experiencing auditory verbal hallucinations at the time of testing and had no previous history of such hallucinations. This was assessed by detailed inspection of the patients’ notes, and consultation with clinical staff. Patients with any history of such hallucinations were excluded. Patients in this group had positive symptoms other than hallucinations – particularly delusions (see Table 1). Eight of these patients were in hospital at the time of testing and two were receiving out-patient treatment.

Exclusion criteria for both patient groups included the presence of an Axis II DSM-IV diagnosis or another Axis I diagnosis, a neurological disorder or a history of substance or alcohol misuse. Patients with an IQ below 80 were also excluded. All patients had been receiving regular doses of antipsychotic medication for at least 1 month prior to testing. Potential participants who reported a history of hearing problems were excluded. The healthy volunteers had a higher premorbid IQ than either patient group; the IQ score was therefore included as a covariate in the between-group analyses.

### Table 1: Group demographic and clinical characteristics.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Control group</th>
<th>Non-hallucinator group</th>
<th>Hallucinator group</th>
<th>Group comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(n=11)</td>
<td>(n=10)</td>
<td>(n=10)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mean (s.d.)</td>
<td>Mean (s.d.)</td>
<td>Mean (s.d.)</td>
<td></td>
</tr>
<tr>
<td>Age, years</td>
<td>29.21 (4.26)</td>
<td>34.78 (11.4)</td>
<td>34.83 (6.88)</td>
<td>NS</td>
</tr>
<tr>
<td>Education, years</td>
<td>14.34 (3.2)</td>
<td>12.3 (1.64)</td>
<td>11.7 (1.41)</td>
<td>NS</td>
</tr>
<tr>
<td>Premorbid IQ score</td>
<td>115 (5.78)</td>
<td>99 (8.56)</td>
<td>100 (7.42)</td>
<td>F=16.9, p&lt;0.001</td>
</tr>
<tr>
<td>Age at first onset, years</td>
<td>21.31 (5.63)</td>
<td>22.5 (5.13)</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Duration of illness, years</td>
<td>16.32 (12.42)</td>
<td>12.33 (9.35)</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>SAPS scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AVH</td>
<td>0</td>
<td>4.47 (0.74)</td>
<td>U=0, p&lt;0.001</td>
<td></td>
</tr>
<tr>
<td>Other hallucinations</td>
<td>0</td>
<td>0.82 (0.32)</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Delusions</td>
<td>4.15 (1.37)</td>
<td>4.41 (0.78)</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Formal thought disorder</td>
<td>1.57 (1.15)</td>
<td>0.95 (0.42)</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Bizarre behaviour</td>
<td>0.73</td>
<td>0.55</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Total score¹</td>
<td>6.38 (2.82)</td>
<td>10.21 (1.40)</td>
<td>U=10.5, p=0.004</td>
<td></td>
</tr>
<tr>
<td>SANS scores</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total score²</td>
<td>6.75 (5.51)</td>
<td>6.70 (3.82)</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Attentional problems</td>
<td>1.83 (1.25)</td>
<td>1.5 (1.05)</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Antipsychotic medication</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Typical:atypical, n:n</td>
<td>3:7</td>
<td>4:6</td>
<td>χ²=0.11, p=0.73</td>
<td></td>
</tr>
<tr>
<td>Depression (CDSS score)</td>
<td>5.51 (6.77)</td>
<td>8.00 (7.22)</td>
<td>NS</td>
<td></td>
</tr>
</tbody>
</table>

AVH, auditory verbal hallucinations; CDS, Calgary Depression Scale; SANS, Scale for the Assessment of Negative Symptoms; SAPS, Scale for the Assessment of Positive Symptoms.

1. Mean of global scores for hallucinations, delusions, bizarre behaviour and formal thought disorder.
2. Mean of global scores for alogia, anhedonia, inappropriate affect, avolition and affective flattening.

The participants’ speech was recorded on Cool Edit 2000 for Windows, which allowed the recordings to be normalised, pitch-shifted and edited into 80 individual wave files. A pitch shift of ~4 semitones was used because it made the speaker’s voice more difficult to recognise without making the speech incomprehensible. A male researcher who was unknown to the participants recorded the words for the non-self condition (40 words in total). A researcher was chosen who used English received pronunciation.

### Stimuli

#### Word lists

Eighty adjectives applicable to people were used (e.g. ‘perfect’, ‘tall’). All the words were monosyllabic or bisyllabic with a Thorndike–Lorge frequency greater than 50 (Gilhooly & Logie, 1980), and were selected from lists used in a previous study (McGuire et al., 1996). The emotional valence of these words had previously been rated by 40 healthy volunteers as either negative, positive or neutral (Johns et al., 2001). Thus the 80 words used consisted of 27 positive, 27 negative and 26 neutral words. The sets of words presented in each condition were balanced for the number of syllables (i.e. equal amounts of one and two syllable words), word frequency and valence (equal amounts of positive, negative and neutral words).

#### Auditory stimuli

The participants’ speech was recorded on Cool Edit 2000 for Windows, which allowed the recordings to be normalised, pitch-shifted and edited into 80 individual wave files. A pitch shift of ~4 semitones was used because it made the speaker’s voice more difficult to recognise without making the speech incomprehensible. A male researcher who was unknown to the participants recorded the words for the non-self condition (40 words in total). A researcher was chosen who used English received pronunciation.
Design
A factorial design was used, with two levels for sources of speech (self, alien) and two levels of distortion (0, −4 semitones). There were 20 words in each of four speech conditions presented in the fMRI experiment (20 self undistorted, 20 self distorted, 20 alien undistorted, 20 alien distorted). The experimental manipulations were source of speech (self, alien) and distortion level (0, −4 semitones). Words were presented in a non-self (alien) voice as well as in the participant’s voice, to test whether any response bias was specific to self-generated words.

Procedure
Patients underwent symptom assessment using the SAPS and SANS either the day before or on the day of the fMRI scan. Approximately 1 hour before scanning all participants were presented with a list of 80 words on a piece of paper and asked to read them aloud in a clear voice at a rate of approximately one word per second. Participants read all 80 words, even though half would subsequently be presented to them in another person’s voice; this was to ensure that participants could not make judgements based on source information during the task. They were not asked to remember the words. Their speech was recorded by a computer. The experimenter then edited the recordings so that 40 of the words were replaced by a recording of the same word spoken in another person’s voice, and 40 were pitch-shifted. The subsets of words that were replaced and pitch-shifted respectively were pre-designated (allocated so that the subsets were matched for word length, frequency and valence). The same subsets of words were used for all participants. Once participants had been placed in the scanner a standardised instruction script was read out to them. Participants were told to listen carefully to each word and make a decision regarding the source of the speech; they were able to register a response of either ‘self’, ‘unsure’ or ‘other’ by means of a button box. The option to register an unsure response was included to avoid participants having to make a forced choice between a self or alien source even when they were unsure.

Image acquisition
Images were acquired in a 1.5 T Magnet (Signa LX; GE, Milwaukee, Wisconsin, USA) using a compressed gradient echo (Edmister et al, 1999), echoplanar image acquisition (Hall et al, 1999), with a time to repetition (TR) of 1.2 s (0.8 s of silence), flip angle 80°, time to echo (TE) 40 ms, 64 × 64 pixels, field of view 200 mm, slice thickness 7 mm and interslice gap 0.7 mm (voxel size 3.125 mm × 3.125 mm × 7 mm); 482 image volumes were acquired in two runs of 6 min each. Of the 482 images 80 were experimental events (20 in each speech condition) and the remainder were rest (i.e. no auditory stimulus was presented). Each whole-brain volume consisted of 14 axial slices parallel to the anterior-posterior intercommissural line.

Stimuli were presented in random order in an event-related design, with a variable interstimulus interval (4–12 s) following a non-gaussian random distribution (Poisson function peaking at 7 s) individually set for each condition (Dale, 1999). Image acquisition and stimulus presentation were synchronised by a transistor–transistor logic (TTL) pulse from the scanner to the computer used to present the stimuli and record the behaviour. The compressed acquisition permitted presentation of each word in the absence of acoustic scanner noise. Each response time was locked to the beginning of the word presentation.

Image analysis
Data were analysed with software developed at the Institute of Psychiatry, using a non-parametric approach. Data were first processed (Bullmore et al, 1999a) to minimise motion-related artefacts. Responses to the experimental paradigms were then detected by first convolving each component of the experimental design with each of two gamma variate functions (peak responses at 4 s and 8 s respectively). The best fit between the weighted sum of these convolutions and the time series at each voxel was computed using the constrained blood oxygen level dependent (BOLD) effect model suggested by Friman et al (2003). Following computation of the model fit, a goodness-of-fit statistic was computed. This consisted of the ratio of the sum of squares of deviations from the mean image intensity (over the whole time series) due to the model to the sum of squares of deviations due to the residuals (SSQ ratio). Following computation of the observed SSQ ratio at each voxel, the data are permuted by the wavelet-based method described and extensively characterised by Bullmore et al (2001). Using this distribution it is possible to calculate the critical value of SSQ ratio needed to threshold the maps at any desired type I error rate. The detection of activated voxels is extended from voxel to cluster level using the method described in detail by Bullmore et al (1999b). Events in the four experimental conditions (self, self distorted, alien and alien distorted speech) were contrasted against rest volumes for all participants.

Group mapping
The observed and permuted SSQ ratio maps for each individual, as well as the BOLD effect size maps, were transformed into the standard space of Talairach & Tournoux (1988) using the two-stage warping procedure described in detail by Brammer et al (1997). Group activation maps were computed by determining the median SSQ ratio at each voxel (over all individuals) in the observed and permuted data maps (medians are used to minimise outlier effects). Cluster-level maps were thresholded at less than one expected type I error cluster per brain. The computation of a standardised measure of effect SSQ ratio at the individual level, followed by analysis of the median SSQ ratio maps over all individuals, treats intra- and inter-individual variations in effect separately, constituting a mixed-effect approach to analysis which is deemed desirable in fMRI.

Repeated-measures contrasts
The analysis was performed using the brain activation data from each participant under each condition. The permutation-based analysis was performed by first determining the median change across all participants and between participant treatments. The treatment labels were then permuted and the median change computed. The use of median statistics renders this analysis robust to outlier data in individual cases. The data were then analysed using a non-parametric repeated-measures analysis of covariance (Bullmore et al, 1999b). The experimental conditions were defined according to the source of the speech (self or alien) and the level of distortion (undistorted or distorted). The data were analysed using a series of non-parametric factorial analysis of variance (ANOVA). We examined the main effect of speech source, distortion and their interactions with group. The effect of the emotional valence of the words on the fMRI data was not examined because it had no significant effect on behavioural results. To test for the interaction
between the source of speech, level of distortion and group we examined the main effect of distortion on self speech and the interaction with group and the main effect of distortion on alien speech and its interaction with group. To examine the neural correlates of the misattribution of speech, we analysed the main effect of the accuracy of attribution (correct responses or misattributions errors). Events were categorised as correct or misattributions according to each participant’s behavioural response. Trials associated with unsure responses were excluded from this analysis. Maps of the difference in the effect size of the BOLD response associated with correct and incorrect attributions were generated. In this particular analysis the effect size statistic was used because the numbers of trials associated with correct and incorrect responses were not equal across conditions. The effect size statistic is relatively insensitive to differences in the number of responses per condition. Use of the effect size statistic also avoids the possibility that differences in BOLD response could reflect changes in the denominator of the statistic (noise) rather than signal, as can occur when using standardised statistics such as $t$, $F$ or SSQ ratio. All between-group contrasts were covaried for NART premorbid IQ scores (using XBAM version 3.4; http://www.brainmap.co.uk/xbam.htm).

**RESULTS**

The demographic and clinical characteristics of the participants are shown in Table 1.

**Behavioural data**

Analysis of variance was conducted for misattribution errors, defined as misidentifications of the source of the speech (i.e. an ‘other’ response when hearing their own speech or a ‘self’ response when hearing alien speech), excluding ‘unsure’ responses (Fig. 1). The data were analysed using an ANOVA for repeated measures.

**Analysis of variance**

For misattribution errors the main effects for source ($F=6.00$, $d.f.=1,28$, $P=0.02$), distortion ($F=12.36$, $d.f.=1,28$, $P=0.002$) and group ($F=6.18$, $d.f.=2,28$, $P=0.006$) were all significant. As there was a significant between-group variance in NART scores this variable was used as a covariate. After the inclusion of this covariate the between-subjects effect for group remained significant ($F=4.67$, $d.f.=2,28$, $P=0.02$). There was a significant interaction between the effects of source of speech and group ($F=3.50$, $d.f.=2,28$, $P=0.04$). A post hoc one-way ANOVA revealed a significant group difference in the self speech condition ($F=11.24$, $d.f.=2,30$, $P<0.001$). A Bonferroni $t$-test showed that those in the hallucinator group made significantly more misattribution errors than the participants in both the non-hallucinator ($P=0.001$) and control groups ($P=0.001$). There was no significant group difference in either of the alien speech conditions (for alien undistorted speech, $F=0.09$, $d.f.=2,29$, $P=0.91$; for alien distorted speech, $F=0.21$, $d.f.=2,29$, $P=0.13$). The interaction between source, distortion and group was non-significant ($F=1.16$, $d.f.=2,28$, $P=0.32$). All main effects and interactions involving valence were also non-significant.

**Imaging data: task-related activation independent of condition**

Performance of the task across all conditions and all groups (independent of condition) was associated with bilateral activation in the inferior frontal, anterior cingulate and superior temporal gyri, the brain-stem and the cerebellum.

**Source of speech and group interaction**

The main effect of source of speech is presented in Table 2. There was a significant interaction between the source of speech and group in the left superior temporal gyrus (Fig. 2(a,b)). Examination of the SSQ ratios from this region revealed that both the control group and the non-hallucinator group showed greater activation when processing distorted vs. undistorted self speech, whereas the opposite was true in the hallucinator group. In the right superior temporal gyrus the hallucinator group showed greater activation for distorted vs. undistorted self speech, the converse was evident in the non-hallucinator group, and distortion had little effect on activation in the control group. The group interaction for the effect of distortion on alien speech was restricted to the right anterior cingulate gyrus (Table 3). In this region both the control group and the non-hallucinator group showed greater activation when processing alien speech that was distorted as opposed to undistorted. However, in the hallucinator group distortion had no effect on the level of activation in this region.

**Distortion and group interaction**

The main effect of distortion is shown in Table 2. There was an interaction between the effects of distortion and group (Fig. 2a,c). In both the control group and the non-hallucinator group processing distorted relative to undistorted speech was associated with activation in the cingulate gyrus. In the hallucinator group the response in this region was unaffected by acoustic distortion (Table 2).

**Effects of distortion on self and alien speech and group interactions**

There were significant interactions between the effect of distortion on self speech and group in the left anterior cingulate and the right superior temporal gyrus (Fig. 3a,b; Table 3). In the cingulate gyrus both the control group and the non-hallucinator group showed greater activation when processing distorted vs. undistorted self speech, whereas the opposite was true in the hallucinator group. In the right superior temporal gyrus the hallucinator group showed greater activation for distorted vs. undistorted self speech, the converse was evident in the non-hallucinator group, and distortion had little effect on activation in the control group. The group interaction for the effect of distortion on alien speech was restricted to the right anterior cingulate gyrus (Table 3). In this region both the control group and the non-hallucinator group showed greater activation when processing alien speech that was distorted as opposed to undistorted. However, in the hallucinator group distortion had no effect on the level of activation in this region.

**Main effect and group interaction for correct v. misattributed responses**

For all participants correct responses (regardless of speech source or the level of distortion) were associated with greater activation in the middle temporal gyrus bilaterally relative to misattributions. No area was more activated in association with misattributions than with correct responses. There was an interaction between response accuracy (correct/misattribution) and group in the left middle temporal gyrus. In both the control and non-hallucinator groups there was greater activation for correct responses (correct identification of either self or alien speech) than for misattributions.
whereas there was no difference in the hallucinator group. In order to test our specific hypothesis about activation being associated with external (self to alien) misattributions, the analysis was then restricted to the self speech condition (i.e. the correct identification of self speech vs. its misattribution to an external source). Again there was an interaction with group in the left middle temporal gyrus, with the same patterns of activation as described above (Fig. 3c, Table 3). When the effect of response accuracy was examined in the alien speech condition alone there was no significant interaction with group.

### DISCUSSION

Our study used fMRI to study the neural correlates of making self/non-self judgments about the source of pre-recorded speech in the presence and absence of acoustic distortion. We examined the effects of speech source and of distortion in patients with auditory verbal hallucinations, patients without such hallucinations and controls. In addition, by using event-related fMRI we were able to categorise the neural response to each word according to the accuracy of the self/non-self attribution and thus examine the correlates of external misattributions.

A tendency for patients with hallucinations to misattribute their own distorted speech to an alien source was first demonstrated using a paradigm in which participants overtly articulated single words and heard what they said in real time (Johns & McGuire, 1999). We used the same paradigm, except that participants heard the words but did not speak. As in a recent study using this modified version of the task, we found that patients with auditory verbal hallucinations also made more external misattributions than both the non-hallucinator group and the control group (Allen et al, 2004), particularly when their speech was distorted (although this did not achieve statistical significance in our study). This may reflect a lack of power, as the number of trials per condition was limited by the practicalities of the fMRI experiment.

Overall, the task activated a network of inferior frontal, temporal and cingulate regions as well as areas in the brain-stem and cerebellum. This is consistent with data from previous studies of voice processing (Binder et al, 2000) and a study of the same task in healthy volunteers (Allen et al, 2005). Within this network, across all three groups there were regions that were more activated when participants processed self-generated speech compared with alien speech and vice versa. However, the hallucinator group differed from both controls and the non-hallucinator group in the effect of the source of the speech on activation in the left superior temporal gyrus. In this region both the reference groups showed increased activation when listening to alien speech compared with self speech, whereas the activation in the hallucinator group was relatively unaffected by the source of the speech. Activation during the task was also influenced by the acoustic distortion of the stimuli. Again, there were significant differences in the effects of distortion between the hallucinators and the other two groups. In the control and non-hallucinator groups distortion was associated with the engagement of the anterior cingulate gyrus, but this effect was absent in the hallucinator group.

The above data suggest that when patients who were prone to hallucinations evaluated speech, the left temporal cortex and the anterior cingulate were differentially responsive to its source and its acoustic quality respectively relative to the reference groups. These findings are consistent with our hypothesis and with data from previous studies that have implicated these regions in schizophrenia (Shapleske et al, 1999; Carter et al, 2001) and the pathophysiology of auditory verbal hallucinations (Suzuki et al, 1993; Shergill et al, 2000a).

The group differences in the effects of source on the left superior temporal activation suggest that this region is normally sensitive to whether speech has been self
Brain activation maps (a) and SSQ plots for (b) the interaction between the effects of source of speech and group in the left superior temporal gyrus and (c) the interaction between the effect of distortion and group in the left ACC (P < 0.01 < 1) false positive cluster. (ACC, anterior cingulate cortex; SSQ, sum of squares; STG, superior temporal gyrus).

Table 3 Main effects and group interactions for the effects of distortion on both self and alien speech and analysis of response accuracy; all contrasts are reported at a clusterwise threshold of P = 0.01 (less than one false positive cluster)

<table>
<thead>
<tr>
<th>Cerebral region</th>
<th>Side</th>
<th>Coordinates</th>
<th>Cluster size</th>
<th>BA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
<td>z</td>
</tr>
<tr>
<td>Effect of distortion on self speech × group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cingulate gyrus</td>
<td>L</td>
<td>-4</td>
<td>22</td>
<td>26</td>
</tr>
<tr>
<td>Superior temporal gyrus</td>
<td>R</td>
<td>51</td>
<td>-18</td>
<td>4</td>
</tr>
<tr>
<td>Effect of distortion on alien speech × group</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cingulate gyrus</td>
<td>R</td>
<td>4</td>
<td>30</td>
<td>26</td>
</tr>
<tr>
<td>Response analysis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correct &gt; misattribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>L</td>
<td>-50</td>
<td>-30</td>
<td>-7</td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>R</td>
<td>-51</td>
<td>-13</td>
<td>0</td>
</tr>
<tr>
<td>Misattributions &gt; correct</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group interaction (all speech)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>L</td>
<td>-50</td>
<td>-30</td>
<td>-2</td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>R</td>
<td>-51</td>
<td>-13</td>
<td>0</td>
</tr>
<tr>
<td>Group interaction in the self speech condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>L</td>
<td>-50</td>
<td>-30</td>
<td>-2</td>
</tr>
<tr>
<td>Group interaction in the alien speech condition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Middle temporal gyrus</td>
<td>R</td>
<td>-51</td>
<td>-13</td>
<td>0</td>
</tr>
</tbody>
</table>

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

Information on the neural correlates of misattributions themselves was obtained by comparing activity associated with misattributions and correct responses. When participants in the hallucinator group made external misattributions (when processing their own speech) these were associated with activation in the left middle temporal gyrus, whereas in the control and non-hallucinator groups there was a greater left temporal response when participants correctly identified their own speech. This distinction between the groups was specific to external misattributions, as there were no group difference in activation when participants misidentified alien speech as their own (internal misattributions).
Both the behavioural and neuroimaging results of our study are similar to those reported using a version of the task that involved participants articulating the words aloud (McGuire et al., 1996; Fu et al., 2001). Thus, in both cases, patients with hallucinations tended to make external misattributions when processing their own distorted speech, and this misattribution was associated with activation of the temporal cortex relative to the correct recognition of self-generated speech. The overall similarity of the results despite the absence of an efference copy component in this study suggests that the differences between the hallucinator groups and the other groups might be related to impairment with the evaluation of auditory verbal material, rather than defective corollary discharge. For example, patients with auditory verbal hallucinations usually have delusions, and delusions are associated with abnormalities of reasoning manifested as a tendency to 'jump to conclusions' (Garety et al., 1991). Indeed, recent behavioural work suggests that misattribution errors on verbal self-monitoring tasks may be related to delusions rather than to hallucinations (Johns et al., 2006). However, this finding was not replicated in our study.

The study has some limitations. Although it focused on how biased judgements might contribute to the experience of externality, it does not explain how the events that are being judged occur in the first place. Contemporary models of hallucinations propose that they arise through the combination of the generation of anomalous experiences and problems in the appraisal of these experiences (Seal et al., 2004; Ditman & Kuperberg, 2005) The biased judgement of sensory material could also contribute to other symptoms, such as delusions: in this case faulty judgements might lead to the misinterpretation of external events such as other people’s behaviour. The coincidence of auditory hallucinations and delusions in schizophrenia is consistent with these symptoms sharing cognitive mechanisms. Second, it is possible that attentional problems may contribute to the tendency to make misattribution errors. The patient groups did not differ on a measure of SANS attentional problems; however, a more rigorous assessment of attentional impairments would have helped to exclude this possibility. The attenuated anterior cingulate response observed in the hallucinator group may reflect problems in these domains. Furthermore, there are strong reciprocal connections between the anterior cingulate and temporal cortex (Petrides & Pandya, 1988). It is possible that the superior temporal gyrus response seen in the hallucinator group is associated with altered ‘top down’ modulation of this region by the anterior cingulate (Fletcher et al., 1999). Although the causation is speculative, it is possible that impaired anterior cingulate modulation of the temporal cortices is associated with making faulty source judgements about perceived speech. The functional integration between the cingulate and temporal cortices could be tested in future work examining the effective connectivity between regions and how this altered in patients with hallucinations.
In summary, external misattributions of speech in patients with hallucinations can occur independently of any self-monitoring deficit, suggesting that hallucinations may be related to problems with the conscious evaluation of verbal material rather than the breakdown of an ‘efficient copy’. This impairment was associated with the abnormal engagement of the temporal cortex along with the anterior cingulate. Although the study involved the evaluation of external rather than inner speech (which is more relevant to verbal hallucinations), it is possible that the same mechanisms are used to appraise internal and external speech.

REFERENCES


Correspondence: Paul Allen, Division of Psychological Medicine PO69, Institute of Psychiatry, De Crespigny Park, Denmark Hill, London. UK. Tel: +44(0)207 848 0514; Fax: +44(0)207 848 0287; email: p.allen@iop.kcl.ac.uk

(First received 25 April 2006, final revision 27 June 2006, accepted 25 September 2006)
Police contact within 3 months of suicide
and associated health service contact

KEITH R. LINSLEY, NEIL JOHNSON and JESSICA MARTIN

Summary  We evaluated police contact with individuals prior to suicide, using a systematic study of suicides within County Durham and Darlington over a 3-year period and analysis of police computer records covering the same area. A total of 205 cases of suicide were identified. A fifth of these (n=41) had a documented contact with police within 3 months prior to the suicide, there was an equal mixture of victims and alleged perpetrators of crime, and a significant number of those with police contact had also seen a health professional recently. As many people see a police officer in the 3 months prior to their suicide as see a mental health professional within 12 months prior to suicide. Our findings have implications for suicide prevention.

Declaration of interest  None.

In the setting up of a multi-agency suicide prevention task force in northern England (further details available from the authors) we wanted to know the likely incidence and type of contact with the police prior to suicide. In one of the few studies addressing this topic, Murphy et al (1971) acknowledged that roughly 80% of a police officer’s activity is devoted to social service, including dealing with suicidal individuals, but that few police officers receive training in dealing with such people, although this contact provides significant opportunities for intervention.

METHOD

One of us (N.J.) identified all residents of County Durham and Darlington who had died between 1 January 1999 and 31 December 2001, with a suicide, open or accidental death verdict. This information was collected from two sources and cross-referenced to reduce the possibility of cases being missed. First, a systematic search was made of the mortality register. Second, information was collected directly from the local coroner’s office. All cases in which a suicide verdict was recorded by the coroner were included in this study. Cases with open or accidental verdicts were scrutinised by two of us and a decision made on whether these deaths should be included in this study as ‘probable’ suicides; see Linsley et al (2001) for a discussion of this problem.

A long time can elapse between death and the issuing of a coroner’s verdict. Therefore records for the 24 months following December 2001 were scrutinised to ascertain ‘late’ verdicts where the death had occurred between 1 January 1999 and 31 December 2001. Coroner’s inquest reports were analysed and information collected in a semi-structured format following a tool used in previous research (Linsley et al, 2001; Schapira et al, 2001). Cases were then cross-referenced with the local police force computer records for the area, concentrating on individuals seen within 3 months of death (90 days or less), considered to be a realistic time frame for intervention. This police force covers the same geographical area as the coroner’s records. Evidence of contact was recorded in a semi-structured format. All data were entered into a Statistical Package for the Social Sciences (SPSS), version 14 for Windows database alongside existing data detailing contact with other agencies.

RESULTS

A total of 133 suicide verdicts were recorded in the 3-year period; 43 open and 29 accidental death verdicts were included as ‘probable’ suicides. This gave a total of 205 probable suicides within the 3-year period.

Twenty-four individuals (12%) had had contact with the police within 3 months of death as victims of crime and 24 individuals (12%) had been arrested as alleged perpetrators of crime (Table 1). Seven individuals had been both a victim of crime and an alleged perpetrator in the 3-month period, leaving an actual total of 41 people (20%) who had been in contact with the police. A wide range of violent and non-violent crime was evident (further information available from the authors).

Among the 41 cases with police contact 17 (41%) had also seen their general practitioner (GP) within the same period (although we were unable to access general practice records in 6 cases). Four cases had no diagnosis recorded at the last consultation, but all four patients had received either antidepressant or benzodiazepine medication. The main diagnoses in the remaining 13 cases were all related to physical health. No particular condition predominated.

Six (15%) individuals had attended a local accident and emergency department in the same period: 3 for self-harm and 3 for other reasons. Almost a third (32%; n=13) had a history of local mental health service contact in the year prior to suicide. Diagnoses included depression (n=6), personality disorder (n=3), alcohol problem, adjustment disorder and anxiety disorder (all n=1). One case had no diagnosis.

In addition to the main findings, 21 cases had impending court appearances (for criminal matters), making it likely that the individual had ongoing police contact. Of these, 14 had been arrested within the last 3 months, and of these 14 people, 6 had also reported crime within that period. Contact with health agencies is summarised in Table 1.

DISCUSSION

The results show a high rate of contact with police for either arrest or reporting crime. Seven cases additionally had impending court appearances. Thus, in nearly 25% of all cases of suicide the person had had a criminal justice contact within 3 months of their death. The National Confidential Inquiry into suicides in the UK found that in just 24% of suicides and open verdicts the person had been in contact with mental health services within 1 year of death (Appleby et al, 2001): thus, as many people see a police officer within 3 months of their
suicide as a mental health professional within 12 months. This makes it imperative that suicide prevention addresses this area.

We accept that the police see a huge number of people and therefore the rate of suicide per contact would be low. However, the same could be said of contact with accident and emergency, mental health and primary care services. This should not deter agencies from trying to prevent suicide.

The median time of contact with police (for those with contact in their last 3 months) was 17 days for alleged perpetrators and 30 days for victims of crime (Table 1). Further analysis of police contact up to 1 year before suicide identified 34 additional contacts (14 reporting crime, 20 arrested) – that is, over a 9-month period. This suggests that the frequency of police contact increases nearer to the time of suicide and warrants further research, as does the nature of the contact and how this relates to health service interaction.

Perhaps most surprising was the more or less equal distribution between victims of crime and alleged perpetrators. It is noteworthy that a comparatively greater proportion of females reported crime than were alleged perpetrators, but nevertheless two-thirds of those reporting crime were male. Services have been developed jointly by criminal justice agencies and mental health services but these have not focused on victims of crime. This study found that a greater percentage of victims than alleged perpetrators had been in contact with mental health services (46% vs. 29%), suggesting victims may warrant a greater degree of liaison between police and mental health services.

Of equal importance is the medical and psychiatric contact these individuals had. Cases of suicide with police contact recorded a higher rate of contact with mental health and accident and emergency services compared with cases of suicide without police contact, and in around 40% of cases the person had seen a GP in the same period. This suggests it is important for police to have some understanding about health contacts, and conversely for health services to be aware of police involvement.

Further thought and work are needed before these findings can be applied generally in practice. However, we believe there is a need to help the police to identify vulnerable individuals, ascertain level of risk and obtain guidance on whom they can contact if concerns are raised. Local protocols between health agencies and police might help. These should include a number of levels of action that apply to victims as well as alleged offenders. This has to be backed up by training programmes for police on recognising suicide risk factors and dealing with suicidal individuals effectively. Furthermore, it should be acknowledged that contact with the criminal justice system itself can have a negative impact on already vulnerable individuals. Whether this contact is the final stressor for some or part of being under existing stress, it is important that all criminal justice agencies – including court diversion schemes, prison authorities and agencies dealing with victims of crime – are aware of this negative impact and the increased risk of suicide.

Finally, we should remember that police have contact with individuals at risk of suicide not included in our research: for instance, individuals who have self-harmed, those expressing suicidal thoughts (e.g. people saved from jumping off bridges) and those detained via police powers under section 136 of the Mental Health Act 1983. Policies should cover these eventualities as well, in particular to ensure that such individuals receive appropriate assessments and do not ‘fall through the net’.

### Table 1

<table>
<thead>
<tr>
<th>All suicides</th>
<th>Reported victim of crime within 3 months of suicide</th>
<th>Arrested for alleged offence within 3 months of suicide</th>
<th>Impending court appearance</th>
</tr>
</thead>
<tbody>
<tr>
<td>n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>205 (100)</td>
<td>24 (12)</td>
<td>24 (12)</td>
</tr>
<tr>
<td>Mean age, years</td>
<td>42.1</td>
<td>36.0</td>
<td>35.9</td>
</tr>
<tr>
<td>Median contact time within the 3 months before suicide, days</td>
<td>NA</td>
<td>30</td>
<td>17</td>
</tr>
<tr>
<td>GP contact within last 3 months, n (%)</td>
<td>98 (100)</td>
<td>11 (46)</td>
<td>9 (38)</td>
</tr>
<tr>
<td>Contact with mental health services within last year, n (%)</td>
<td>61 (30)</td>
<td>11 (46)</td>
<td>11 (52)</td>
</tr>
<tr>
<td>Contact with A&amp;E department within last 3 months, n (%)</td>
<td>24 (12)</td>
<td>3 (12)</td>
<td>5 (21)</td>
</tr>
</tbody>
</table>

A&E, accident and emergency; GP, general practitioner; NA, not applicable.

1. Three cases of GP contact had both crime victim and alleged perpetrator, making 17 ‘actual’ cases in total.
2. Five cases of mental health contact had been both crime victim and alleged perpetrator, making 13 ‘actual’ cases in total.

### REFERENCES


Treatment adequacy for anxiety and depressive disorders in six European countries


Summary The aims of this study are to describe the adequacy of treatment for anxiety and depressive disorders in Europe and how it differs between providers, using data from the ESEMeD study. The overall proportion of adequate treatment was 45.8% (57.4% in the specialised sector and 23.3% in the general medical care sector). Between-country differences were found in treatment adequacy in the specialised setting. Organisational and political aspects may explain these findings.

Declaration of interest Partial funding from several drug companies involved in the manufacture of antidepressant medication; full acknowledgements in a data supplement to the online version of this paper.

Research on quality of care for mental disorders has systematically reported low rates of treatment guideline adherence (Ramana et al, 1999; McConnell et al, 2002; Oquendo et al, 2002; Kessler et al, 2003; Wang et al, 2005). This has significant health consequences, since treatments meeting clinical guidelines are cost-effective and decrease years lived with disability (Andrews et al, 2004). The majority of previous studies have been conducted in the USA, and little is known about treatment adequacy in Europe.

This study is based on a European epidemiological study of the prevalence and treatment of mental disorders. Our aims are to describe treatment adequacy for anxiety and depressive disorders in Europe, how it differs between countries and providers, and which factors are associated with appropriate care.

METHOD

The European Study of the Epidemiology of Mental Disorders (ESEMeD) project is a cross-sectional household survey representative of the non-institutionalised adults of Belgium, France, Germany, Italy, The Netherlands and Spain. A stratified, multi-stage, clustered area, probability sample without replacement design was used. Data for the project were provided by 21,425 respondents. A description of the ESEMeD methodology has been provided by Alonso et al (2004). Response rates ranged from 45.9% in France to 78.6% in Spain.

Mental health status was assessed with the Composite International Diagnostic Interview 3.0 (Kessler & Ustun, 2004). The diagnoses included in this paper were DSM-IV major depressive episode and anxiety disorders (social phobia, generalised anxiety disorder and panic disorder) (American Psychiatric Association, 1994). Individuals reporting any use of health services as a result of their ‘emotions or mental health problems’ in the 12 months before the interview were asked to select whom they visited from a list including psychiatrist, psychologist, general practitioner (GP) or any other medical doctor. Psychiatrists and psychologists constituted the specialised mental health category; GPs and other doctors formed the general medical care category.

Criteria for minimally adequate treatment were receiving antidepressant pharmacotherapy (for depression) or antidepressant or anxiolytic pharmacotherapy (for anxiety) for at least 2 months plus at least four visits with a psychiatrist, a GP or any other doctor; or at least eight sessions with a psychologist or a psychiatrist lasting an average of 30 min (American Psychiatric Association, 1998, 2000; Guidelines Advisory Committee, 2001; Kessler et al, 2003; Royal Australian and New Zealand College of Psychiatrists, 2003; National Institute for Clinical Excellence, 2004; Wang et al, 2005).

Data were weighted to adjust for the multistage probability sampling. Population projection weights were used to restore the representativeness of the sample regarding age and gender distribution in each country. A logistic model was used to analyse factors associated with treatment adequacy. Since the same individual could have received treatment in both the specialised and general medical sectors, a generalised estimating equation model was used, including two observations for those treated in both sectors (Zeger & Liang, 1986). Statistical analyses were carried out using Stata version 8.0 and SAS version 9.1 for Windows.

RESULTS

An average of 29.5% (429 individuals) of those with a diagnosis of major depressive episode or anxiety disorder in the past 12 months had consulted any health service during that period. Of these individuals, 59 lived in Belgium, 89 in France, 49 in Germany, 36 in Italy, 62 in The Netherlands and 134 in Spain. The overall proportion of treatment adequacy for any disorder was 45.8% (95% CI 39.2–52.4), ranging between 45.8% (95% CI 38.47–53.05) for major depressive episode and 54.5% (95% CI 44.78–64.19) for anxiety disorder. By setting, rate of treatment adequacy for any disorder was 57.4% (95% CI 49.74–65.1) in the specialised care category and 23.3% (95% CI 16.7–29.8) in the general medical care category (specialised care as reference, OR=0.25, 95% CI 0.16–0.38). The same pattern was observed for both types of disorder.

By country, overall proportions of adequacy varied from 32.5% (95% CI 21.5–43.2) in Spain to 55.4% (95% CI 40.3–70.5) in The Netherlands (P=0.11). The proportion of individuals receiving minimally adequate treatment in the specialised care varied widely, from 29.2% (95% CI 17.4–41.0) in Spain to 78.2% (95% CI 65.4–91.0) in France (P<0.001). In the general medical setting, proportions varied between 14.9% (95% CI 1.0–28.7) in Belgium and 33.6% (95% CI 14.4–52.9) in Italy (P=0.54).

Being treated by a general medical provider was associated with a lower probability of receiving adequate treatment in Belgium (OR=0.24, 95% CI 0.19–0.64), France (OR=0.09, 95% CI 0.04–0.23), Germany (OR=0.16, 95% CI 0.05–0.56) and The Netherlands (OR=0.35, 95% CI 0.18–0.69). Provider differences in each country according to disorder were similar to the overall differences.

Two different models were run in order to ascertain the factors associated to treatment adequacy. After adjusting by gender, age (centralised around median value, 42 years
old), urbanicity (living in a city with >100,000 inhabitants v. smaller), presence or absence of chronic illness, and health state assessed using the EuroQol, only type of provider and country were related to treatment adequacy. As some interaction between provider and country was detected, we adjusted a second model. In this model, provider by itself was not significant (taking specialised care as reference, OR=0.76, 95% CI 0.34–1.71). Using Spain as reference, living in France (OR=8.91, 95% CI 3.37–23.55), Germany (OR=5.16, 95% CI 1.81–14.18) and The Netherlands (OR=5.14, 95% CI 1.94–13.62) was related to increased probability of receiving adequate treatment. Only the interactions between provider (generalised care) and France (OR=0.10, 95% CI 0.03–0.35) or Germany (OR=0.20, 95% CI 0.05–0.84) were statistically significant. (The results are summarised in a data supplement to the online version of this paper.)

DISCUSSION

Results should be interpreted considering the following limitations. First, information about treatment was self-reported. Second, the final sample considered was small and data should be interpreted with caution. Third, we have not been able to analyse how national differences in response rate affect the results. Finally, we might have underestimated treatment inadequacy owing to the loose criteria used.

In spite of the limitations, our results suggest that treatment adequacy rates for anxiety disorders and major depressive episodes in Belgium, France, Germany, Italy, The Netherlands, and Spain are similar to those found by Wang et al (2005) in the USA. Rates of minimal adequate treatment in the USA were 52.0% in the specialised setting and 14.9% in the general medical setting; in Europe the rates were 57.4% and 23% respectively. However, Wang’s study included all DSM-IV diagnoses, whereas we focused on only two types of disorder.

Although overall rates of adequacy were similar across Europe, the differences between providers varied. In the northern countries (Belgium, France, Germany and The Netherlands) treatment adequacy was higher in the specialised sector, whereas in the southern countries (Italy and Spain) there was no difference. This result was not anticipated, since published studies systematically report that those treated in a specialised setting are more likely to receive adequate treatment (Kniesser et al, 2005; Wang et al, 2005).

Differences in European healthcare systems might explain these variations. Spain and Italy have a national health service financed by general taxation; the other countries have a system of compulsory social health insurance. In Spain and Italy a GP referral is usually needed to access specialised care. Practice guidelines could also explain differences. Practice guidelines have, at least theoretically, an important role in France, Germany and The Netherlands. In France, the National Agency for Accreditation and Evaluation of Health Care has published a depression guideline; Germany has an Institute for Quality and Efficiency that promotes evidence-based treatments; and in The Netherlands both GPs and psychiatrists publish guidelines for depression (more information on the healthcare systems of these countries can be obtained from the European Observatory, http://www.euro.who.int/observatory). However, the role of practice guidelines has been questioned by Gilbody et al (2003), who highlight the point that simple guideline creation is ineffective. The finding that France and Germany have a high overall adequacy rate but low adequacy in the general medical setting, whereas The Netherlands has one of the highest rates of treatment adequacy in the general medical setting, could be explained by the fact that guidelines in The Netherlands were developed by both primary care physicians and specialists, supporting the hypothesis that collaborative care improves quality of care.

REFERENCES


American Psychiatric Association (2000) Practice Guideline for the Treatment of Patients with Major Depressive Disorder. APA.


Localisation of increased prefrontal white matter in pathological liars

YALING YANG, ADRIAN RAINÉ, KATHERINE L. NARR, TODD LENCZ, LORI LaCASSE, PATRICK COLLETTI and ARTHUR W. TOGA

Summary  We examined white matter volumes in four prefrontal subregions using structural magnetic resonance imaging in 10 pathological liars, 14 antisocial controls, and 20 normal controls. Liars showed a relatively widespread increase in white matter (23–36%) in orbitofrontal, middle and inferior, but not superior, frontal gyri compared with antisocial and normal controls. This white matter increase may predispose some individuals to pathological lying.

Declaration of interest  None.

METHOD

Participants were taken from a total sample of 108 community volunteers drawn from five temporary employment agencies in Los Angeles (Rainé et al, 2000). The three groups consisted of 10 people with a history of repeated lying (‘liars’), 20 normal controls who had neither antisocial personality disorder nor a history of pathological lying, and 14 ‘antisocial’ controls matched for antisocial behaviours but with no history of pathological lying. Participants were defined as pathological liars if they fulfilled criteria for: pathological lying on the Psychopathy Checklist–Revised (PCL–R; Hare, 1991); or conning/manipulative behaviour on the PCL–R; or deceitfulness for DSM–IV antisocial personality disorder (American Psychiatric Association, 1994), or malingering (for details see Yang et al, 2005). Of the 10 liars in this study, 5 were classified as malingers. Full informed, written consent was obtained from all participants in accordance with institutional review board procedures. Five brain volumes from the original sample (Yang et al, 2005) could not be segmented owing to irretrievable corruption on data storage. Missing data were relatively evenly distributed across groups, with two from the liar group, two from the antisocial control group and one from the normal control group.

Structural MRI was carried out on a 1.5-Telsa Philips 515/ACS (Selton, Connecticut, USA) scanner using three-dimensional T1-weighted gradient-echo scans (for details see Yang et al, 2005). All image data-sets were processed with a series of preparatory steps before manual delineation of prefrontal subregions (Sowell et al, 1999, 2002). First, all images were anonymised to exclude personal information. Second, non-brain tissue and the cerebellum were removed from the brain volume, and signal intensity inhomogeneities were corrected (Sled & Pike, 1998). Third, fully automated tissue segmentation was applied and brain voxels were automatically classified as gray matter, white matter, or cerebrospinal fluid using a validated partial volume correction method (Shattuck et al, 2001). Finally, a spherical mesh surface was created using a three-dimensional active surface algorithm to facilitate identification of anatomical boundaries (MacDonald et al, 1994).

The parcellation of the prefrontal lobe into four subregions for each hemisphere followed the methods of Ballmaier et al (2004). A three-dimensional shape representation and coronal two-dimensional MRI scan of the segmentation of the prefrontal cortex of one of the participants are shown in the data supplement to the online version of this paper. All anatomical delineations were conducted by two research assistants trained by Y.Y. Unlike gray matter subregions, which are clearly defined by sulcal landmarks, white matter delineations are arbitrary and the segmentation results should be viewed as estimated volumes. To assess interrater reliability, all anatomical regions were delineated on ten randomly chosen image data-sets; intraclass correlation coefficients ranged between 0.90 and 0.97 for gray matter and white matter in all four frontal subregions. Each of the eight subregional volumes was divided by total intracranial volume to account for potential differences in individual brain size. Since there was a lack of hemisphere effect, white matter volumes from two hemispheres were averaged to create a mean regional volume.

RESULTS

A multivariate analysis of variance (ANOVA) showed a main group effect for whole-brain-corrected white matter volume in prefrontal regions (i.e. inferior frontal, middle frontal, orbitofrontal and superior frontal cortices); F(8, 78) = 4.19, P = 0.001, r² = 0.29. Groups differed in the volume of white matter in the inferior (F(2, 41) = 11.09, P = 0.001), middle (F(2, 41) = 7.05, P = 0.003) and orbitofrontal cortex (F(2, 41) = 6.87, P = 0.001), with increased white matter in liars. However, a trend towards lower white matter volume was found in the superior frontal cortices for liars (F(2, 41) = 0.42, P = 0.66). Liars showed significantly increased white matter in inferior, middle and orbitofrontal cortex compared with both antisocial controls (P = 0.001, P = 0.004, and P = 0.006, respectively) and normal controls (P = 0.001, P = 0.005, and P = 0.001 respectively; Fig. I). No difference was found for gray matter volume in the four subregions (F(8, 78) = 0.54, P = 0.82).

DISCUSSION

Following our previous finding of a prefrontal white matter increase in people who lie, cheat and manipulate others (Yang et al, 2005), this study found pathological liars to have increased white matter volumes in some prefrontal subregions, particularly orbitofrontal cortex (22–26% increase), inferior frontal cortex (32–36% increase) and middle frontal cortex (28–32% increase) compared with both antisocial and normal controls. An important exception was that no white matter increase was found for the superior frontal cortices. Such an increase might be expected based on findings of an fMRI study in which activation of superior frontal cortices was found during a deception task
invoking motor responses (Langleben et al., 2002). In contrast, one study using a potentially more realistic lying task involving a verbal response found prefrontal activation specifically in ventrolateral and orbitofrontal cortex, but not superior frontal cortices (Spence et al., 2004). Moreover, these nonsuperior frontal regions are most frequently shown to be activated by deception tasks (e.g. Spence et al., 2004; Langleben et al., 2005; Phan et al., 2005). This may in part explain why we observed white matter increases in the ventral (orbitofrontal cortex), ventrolateral (inferior frontal cortex), and inferior aspect of dorsolateral (middle frontal cortex), but not superior dorsolateral (superior frontal cortices), frontal regions in the liar group.

One interpretation of the white matter increases in the ventral and lateral nonsuperior frontal regions could be that a pre-existing variation in prefrontal structure may predispose individuals to engage in pathological lying. Alternatively, several studies have argued that long-term training may induce regional increases in white matter volume (Schmithorst & Wilke, 2002; Bengtsson et al., 2005). In the case of lying, it is conceivable that excessive lying repeatedly activates the prefrontal circuit underlying lying, resulting in permanent changes in brain morphology. This ‘Pinocchio’s nose’ hypothesis of pathological lying could be compared with the competing predispositional hypothesis using a prospective longitudinal study assessing both white matter volume and degree of lying from childhood to adulthood.

The engagement of ventral and lateral prefrontal regions in lying may be anticipated from fMRI studies, several of which have associated these regions with executive functions crucial to successful deception, including decision-making, moral reasoning, rule maintenance/retention and response inhibition (Bunge, 2004). Although some studies showed partial activation in the superior frontal cortex when lying involved a non-vocal motor response (Langleben et al., 2002, 2005), this region is more associated with functions less directly linked to deception, such as spatial information processing, attention reorientation and novelty detection (Gomot et al., 2006). Conversely, gains in white matter volume in these prefrontal regions (in the absence of gray matter reduction) may lead to faster sharing of information within fronto-cortical circuits in pathological liars. Thus, increased white matter in these subregions of the prefrontal cortex in liars may predispose to maintaining a life-style of pathological lying and malingeredness. The use of advanced imaging techniques such as diffusion tensor imaging to assess neural connectivity (Nakamura et al, 2005) may allow more thorough investigation of the subtle abnormalities responsible for pathological lying.

ACKNOWLEDGEMENTS

We thank Samantha Henry, Elizabeth Culley, Donna Kha, Reimar Macaranas, Henry Wu, Lyda Lee, Sum-yen Ng and Sridhar Chadalavada for data collection and scoring. This study was supported by grants to A.R. from NIMH and the Wacker Foundation, grants from the National Center for Research Resources the NIH Roadmap for Medical Research and a NARSAD Young Investigator Award for K.L.R. and A.W.T.

REFERENCES


Correspondence

EDITED BY KIRIAKOS XENITIDIS and COLIN CAMPBELL

Contents ■ Mental health and incapacity legislation ■ Violence and psychiatric morbidity ■ Assessment of manic symptoms in different cultures

Mental health and incapacity legislation

I enjoyed the article by Dawson & Szumukler (2006) because I like to keep up to date with legal and ethical issues in mental health. However, their claim for equivalence between mental and physical diseases sits uneasily with scientific papers published in the Journal. Shaw et al (2006) found that schizophrenia had a prevalence of 5% in perpetrators of homicide, compared with 1% in the general population. I would love to see comparable figures for the prevalence of hypertension, multiple sclerosis, leprosy etc., but meanwhile we have a problem. The Ritchie report on the inquiry into the care of Christopher Clunis reveals capacity’s dark side by showing how psychiatrists repeatedly brought a patient to the point at which he could make his own decisions, then left him to fend for himself (Ritchie et al, 1994). Perhaps the best way for services to reduce the stigma and discrimination associated with psychiatric illness is to reduce the 5% figure? Somehow, I cannot see capacity-based legislation playing a lead role in achieving that objective.


A. Maden Department of Forensic Psychiatry, Academic Centre, West London Mental Health NHS Trust, Southall, Middlesex. Email: a.maden@ic.ac.uk

doi: 10.1192/bjp.190.2.176

Psychiatric patients can be treated involuntarily even if they possess the mental capacity that would render the involuntary treatment of a medical patient illegal. Dawson & Szumukler (2006) describe this as a form of discrimination and propose that the relevant legislation be ‘fused’ so that, like medical patients, most psychiatric patients could be treated involuntarily only if they lacked mental capacity. I see a number of advantages to using mental capacity as a legal criterion (Buchanan, 2002, 2005). I suspect, however, that Dawson & Szumukler’s solution encourages its own form of discrimination. Under the proposals, ‘non-forensic’ patients could be treated involuntarily only if they lacked mental capacity. However, ‘forensic’ patients would be liable to a different, and easier-to-meet, set of criteria.

Underlying the distinction seems to be an assumption that the duties of doctors are different in respect of mentally disordered offenders. Some of the patients that forensic psychiatrists treat, Dawson & Szumukler write, are ‘not . . . under treatment primarily for their own benefit, but for the protection of others’ (p. 508). This seems to mistake a difference in emphasis for something more significant. First, benefiting patients and protecting others are not mutually exclusive. Second, treatment directed to both of these ends is not limited to forensic psychiatry. Third, where a tension does exist the position is straightforward. Exceptional cases notwithstanding, a doctor’s primary responsibility is his patient’s well-being. Ethical guidelines make no distinction in this regard between ‘forensic’ and other patients (Gunn & Taylor, 1993; Bloch & Green, 2006).

If capacity principles are to govern the coercion of psychiatric patients, I am not convinced that any ‘forensic exception’ is necessary. In England and Wales the important area is the hospital order under section 37 of the Mental Health Act 1983 (945 cases in 2004, 288 with restrictions). Here Dawson & Szumukler have two suggestions. The first would replace the hospital order with something like the present ‘hospital direction’ under section 45A of the Act. The second would sanction the involuntary treatment of a patient with mental capacity for a period ‘proportionate to the seriousness of the offence’ if a court thought that this would reduce reoffending. Presumably, the same treatment would be clinically indicated in many cases but the suggested criteria do not require this. Psychiatrists have complained that the hospital direction requires them to declare patients ‘fit for punishment’ (Mullen et al, 2000). The second suggestion implies the use of compulsory psychiatric treatment to achieve a legal end.

Instead, if capacity is to govern involuntary psychiatric treatment, why not make the passing of a hospital order, with or without restrictions, dependent on the patient consenting (or, if the patient lacks capacity, dependent on treatment being in their best interests)? The law could then permit re-sentencing if the convicted defendant changed his mind (or regained capacity and refused treatment), when the situation would be similar to the breaching of a probation order with a condition of treatment. The initial decision to give consent would often be difficult especially where the offence was serious and the choice lay between a substantial prison term and indeterminate detention in hospital. However, I am not clear that a competent defendant should be prevented from making it, particularly if the interim hospital order under section 38 of the Act remained available for cases where the psychiatrist was unsure whether to offer treatment or the patient was unsure whether to accept.

Because adherence is often partial there would still be cases where the doctor’s subsequent decision that a failure to participate in treatment amounted to withdrawal of consent could be seen as declaring the patient ‘fit for punishment’. Such a scheme would also have to overcome objections that section 37 of the Act already provides an efficient way of getting treatment to people who need it, resources permitting. However, by making court-ordered treatment dependent on consent, it would bring the management of those with psychiatric illness more into line with that of patients elsewhere in medicine. Moreover, it would do so without replacing one form of discrimination with another.


A. Buchanan Department of Psychiatry, Yale University, 34 Park Street, New Haven, CT 06519, USA. Email: alec.buchanan@yale.edu doi: 10.1192/bjp.190.2.176a

Authors’ reply: Dr Buchanan notes that we allow an exception to ‘pure’ incapacity principles where a serious offence has been committed by a person with a mental disorder. We would allow the involuntary treatment of this narrowly defined subgroup of forensic patients under certain conditions, even if they regained capacity, to prevent harm to others. Buchanan believes this would discriminate unfairly between these patients and non-forensic patients. We are not convinced, however, that this would involve unjustified discrimination, because the commission of a serious offence constitutes a significant difference between their positions.

Nevertheless, Dr Buchanan’s suggestion that convicted offenders might be given a choice, on disposition from the court, of accepting imprisonment or consenting to treatment in hospital deserves serious consideration. However, we think a time limit should still be placed on the period during which a patient could be treated in hospital on this basis. That time would be proportionate to the seriousness of their offence. Otherwise, the patient who accepts hospitalisation and treatment initially, but later refuses treatment when they regain capacity, would face return to court for resentencing for an indeterminate period. Or, if the patient were to make a rapid recovery with treatment, would discharge very soon after a serious offence be politically acceptable?

Professor Maden, as we understand it, fears that the legislation we propose would not reduce homicides by people with mental illness, but we have little knowledge of the effect of mental health laws on rates of serious offending. What is most likely to reduce rates of violence is early access to effective treatment. Our proposal would allow involuntary treatment for the right reasons at the right time, and it may permit intervention sooner than under the 1983 Act. Some people with personality disorders who pose a risk of harm to others may not meet our incapacity test, and the transitional position of such persons who are already detained in our mental health facilities would have to be addressed. However, on balance, we think our proposals are likely to reduce violence overall, by allowing earlier access to effective treatment for persons who are incapacitated, regardless of the cause.

J. Dawson Faculty of Law, University of Otago, PO Box 56, Dunedin, New Zealand. Email: john.dawson@stonebow.otago.ac.nz doi: 10.1192/bjp.190.2.177

G. Szmukler Institute of Psychiatry, King’s College, London, UK

Violence and psychiatric morbidity

Coid et al (2006) reported an important cross-sectional survey of 8397 persons in UK households and found that psychosis was independently associated with a sixfold increase in the reporting of five or more violent incidents. Given the controversy and sensitivity over the stigma associated with psychiatric illness, particularly concerning public perceptions of links between psychosis and violence, this kind of result is prone to generate misleading impressions.

In a recent comprehensive review Hiday (2006) points out that surveys of this type are prone to exaggerate the contribution of mental illness and other diagnostic labels to violence as a result of several methodological weaknesses. The first is associated with the issue of comorbidity. It was not clear from the presentation of their data whether Coid et al were able to investigate the comorbidity of psychosis and other diagnostic categories and violence. It is possible that once comorbid substance misuse, personality disorder or other issues were taken into account, the unique contribution of psychosis to violence might have diminished dramatically (Hiday, 2006).

There is an even more fundamental problem that underpins violence surveys of this type: a neglect of the confounding factor that those with mental illness are more likely to reside in violent neighbourhoods and this could be the key predictive variable, not the illness itself. The term now used to describe the places where most people with severe mental illness live is ‘socially disorganised communities’, and these combine a multiplicity of factors that promote violence completely independently of psychiatric dysfunction (Silver et al, 2001). Features of these environments include chronic disabling poverty, few employment prospects or educational opportunities, decaying buildings and few amenities. In these neighbourhoods families and similar social institutions have broken down, leaving most individuals devoid of traditional social guidance and control (Swanson et al, 2002).

Living and growing up in such environments is possibly the key variable that predicts violence, not the mental illness of the individual (Hiday, 2006). Community household surveys such as that reported by Coid et al (2006) represent a unique opportunity to explicate the contribution of ecological factors when violence appears to be linked to mental illness. It would therefore be useful in terms of advancing the debate over the link between violence and mental illness if a wider theoretical background to such analyses could be encouraged in the future.


R. Persaud Westways, The Maudsley Hospital, 49 St James Road, West Croydon CR0 2UR, UK. Email: R.Persaud@slp.ki.ac.uk doi: 10.1192/bjp.190.2.177a

Authors’ reply: We do not want our finding of a sixfold increase in reporting five or more violent incidents in persons with psychosis to give a misleading impression regarding the association of violence with mental illness. This was the only finding suggesting increased risk and means that there is a small subgroup of people with psychosis who are repeatedly violent. The real message of our paper should have
been that the true risks of violence from people with psychosis, at the population level, are exceedingly small.

Professor Persaud’s impression might be owing to the space in our paper devoted to discussing the public health impact of alcohol misuse and antisocial personality disorder on violence. In an additional paper published recently in the *American Journal of Epidemiology* we make the point about psychosis more strongly (Coid et al., 2006). Researchers with an interest in violence and psychosis often emphasise that relative risks of violence are greater for individuals with psychosis but they ignore the fact that illnesses such as schizophrenia are rare and that persons with psychosis account for an exceptionally small number of violent incidents at the population level. Detaining more persons with psychosis in hospital would have a very small effect in reducing violent crime (Fazel & Grann, 2006).

Misleading impressions based on relative risks are typical for homicides perpetrated by people with psychosis. These are often based on Scandinavian countries where the base rate is exceptionally low (Hodgins & Janson, 2002). In locations where the base rate is very high, for example, certain areas in the USA and South American countries, people with psychosis hardly feature in criminal statistics.

Careful reading of our paper will reveal how we dealt with confounding from comorbid conditions. We agree with Professor Persaud’s point about residents in violent neighbourhoods entirely, but the sampling frame was intended to exclude bias from factors such as socioeconomic deprivation. We used two-level hierarchical models throughout the analysis to take account of clustering from these areas. We would concede, however, that our study did not adequately explore the important issue of neighbourhood effects.


J. Coid Forensic Psychiatry Research Unit, St Bartholomew’s Hospital, William Harvey House, 61 Bartholomew Close, London EC1A 7BE, UK. Email: j.w.coid@qmul.ac.uk

Assessment of manic symptoms in different cultures

Mackin et al (2006) make a laudable attempt to evaluate cultural differences in the perception of psychiatric symptoms. Unfortunately, aspects of their methodology make it difficult to draw definitive conclusions. I will leave it for the statisticians to decide whether the sample sizes for the English and Indian groups (n=20 and 24 respectively) are large enough to allow the findings to be generalised. Given the authors’ concerns about the influence of confounding variables on the findings, however, the disparity between the size of these groups and that of the American clinicians (n=82) is striking. A demographic breakdown of the various groups might have been useful in allaying these concerns.

A further source of potential bias is introduced by asking the participants to complete rating scales for only two patients of a single nationality. There is a risk that cultural differences between nationalities might influence attitudes as to what can be considered ‘normal’ behaviour for people of other nationalities. Certainly, an English psychiatrist whose expectations of a ‘typical’ American have been shaped by stereotyped media images might not be expected to register certain aspects of the patients’ behaviour as pathological on the Young Mania Rating Scale. The threshold for recognition of manic symptoms might well have been different had they been asked to rate their own compatriots. More revealing conclusions could perhaps have been drawn had all participants been asked to complete rating scales for patients of a variety of nationalities, including their own.

The authors make a compelling argument about the potential consequences of cultural differences in the recognition of symptoms of mental illness, and have provided a useful starting point for future discussion and research. Unfortunately, they fall short of proving these differences exist with their preliminary data.


M. Sanderson Department of General Adult Psychiatry, Bushley Fields Hospital, Bushley Fields Road, Dudley, West Midlands, UK. Email: matthewsandel@doctors.org.uk
implications of these difficulties by stating that ‘similar variability is likely to be present when ranking patients in routine clinical practice’. Few would debate the existence of inter-observer variability, but the core issue here is whether the authors’ data support culture as being a central factor in this phenomenon. The design of the study simply does not permit this conclusion.


R. Reed Barnet, Enfield and Haringey Mental Health NHS Trust, 58–60 Silver Street, Enfield EN1 3EP UK. Email: ruthreed@gmail.com doi: 10.1192/bjp.190.2.178a

Author’s reply: We wholeheartedly agree with Dr Sanderson’s conclusion that this study provides ‘a useful starting point for future discussion and research’. Clearly, the number of assessed patients was small as was the number of clinician-raters. We acknowledge these points in our discussion and conclude by recommending other large studies using patients from real-life clinical settings. We also agree that perception of ‘normal’ behaviour would vary according to nationality and this might have very real significance when assessing the mental state of an individual. This warrants further research.

Drs Sanderson and Reed both comment on the lack of socio-demographic data on the rating clinicians but unfortunately these data are not available. We disagree with Dr Reed’s assertion that we are required to make the assumption that the groups are similar in all respects except culture. We state clearly that ‘we cannot exclude the possibility that other factors, in addition to cultural background, may have influenced these results’, and we go on to prescribe potential confounding influences, including age, gender, psychiatric training, years of experience, etc. Similarly, Dr Reed’s suggestion that we minimised the implications of these difficulties is unfounded; in fact, we highlight the possibility that multiple factors, including cultural biases, might affect the accuracy of scores on the Young Mania Rating Scale between clinicians from different countries. It is highly probable that similar variability will be present when this rating scale is used in routine clinical practice by clinicians from diverse cultural backgrounds.

Notwithstanding the preliminary nature of our study and the methodological considerations discussed above, we believe our data support the suggestion that cultural background influences the interpretation of manic symptoms when using the Young Mania Rating Scale.

P. Mackin University of Newcastle upon Tyne, Department of Psychiatry, Leazes Wing, Royal Victoria Infirmary, Queen Victoria Road, Newcastle upon Tyne NE1 4LP UK. Email: paul.mackin@ncl.ac.uk doi: 10.1192/bjp.190.2.179

One hundred years ago


Falcìola

Dr. Falcìola has collected a large number of papers upon the growth of nails and the changes noted after disease, which he has tested by his own observation. He is not disposed to agree to the assertion of Parisot and Paget that the state of the nails is an index of trophic alterations in the body, although he admits that their growth is affected by a general disturbance in the economy of the organism. He found that in melancholy the growth of the nails is slower. The increase of the nails is somewhat irregular, being greater at one time than another, and differing in each finger, although there is a general equality in growth, which is more marked in the three middle fingers. The nails of one hand do not grow at exactly the same rate as in the other. He fails to find either marked acceleration or slowness of growth in states of mental depression or mania. In general he finds that the study of the growth of the nails in insane patients appears to support the views of Kraepelin on the clinical unity of all those types of mental disease which writers generally wish to treat as distinct, but which, in truth, only represent different episodes of one fundamental malady.

REFERENCE

Journal of Mental Science, January 1907, 185. Researched by Henry Rollin, Emeritus Consultant Psychiatrist, Horton Hospital, Epsom, Surrey doi: 10.1192/bjp.190.2.179a
Book reviews

EDITED BY SIDNEY CROWN, FEMI OYEBODE AND ROSALIND RAMSAY

Dementia – Mind, Meaning and the Person

Wonder ... is the first principle which prompts mankind to the study of philosophy
(Adam Smith, 1795)

The best College meeting I have attended was the joint conference between the Philosophy Special Interest Group and the Faculty of Old Age Psychiatry in Newcastle, in the autumn of 2002. I am not a philosopher but the ‘wonder’ generated by this meeting has remained with me. I was a naive but entirely interested reader of Dementia – Mind, Meaning and the Person, pleased that contributors to the conference, clinicians and philosophers, had written chapters.

Inevitably in a multi-author book some chapters are more appealing and accessible than others. Being an orderly person I began at the beginning, but chapter 1 is by far the most difficult for the philosophy initiate – it does introduce the feast to come but as a hard-going menu in a foreign language, not as an aperitif. I returned to the first chapter having finished the book and at that point found it more digestible.

The book covers not only the philosophical but also social, spiritual, ethical and practical perspectives and the negative, soul-destroying attitudes about dementia in modern society. Dementia, with its progressive inevitable deterioration in memory, language and skills, causes us to consider the meaning of personhood and identity. Hughes had previously enlarged on the Locke–Hume reductive view that a person is no more than connected mental states, seeing the patient as a ‘situated-embodied-agent’. The person with dementia has to be understood in terms of relationships, not because that is all that is left to them but because that is characteristic of all our lives. ‘He is not the man I married, Doctor’, is commonly heard and needs to be taken seriously. In his quiddity he is the same, not a vegetable nor in second childhood, but the same man with a dementing illness. However, from a social constructionist viewpoint the usual mutual task of holding and preserving identity is now a solitary and heartbreaking one for the partner. There is a need to guard against using a purely social constructivist approach in order not to deny the patient’s personal human agency and meaningful intersubjectivity. To be semiotic requires some capacity to express genuine intentions and meanings. We need to search for rational patterns which reveal the mind at work. Dementia threatens the process of meaning-making, the hermeneutics of life.

The book only occasionally entertains Doctor Scholasticus with angels dancing on pinheads. Ordinary clinicians need to suspend criticism of the way some philosophical arguments are constructed or supposed syntactical sentences put into the mouth of someone with severe dementia.

The chapter for all in old age psychiatry entitled ‘Respectare: moral respect for the lives of the deeply forgetful’ looks again and more carefully at the experience of persons with dementia. ‘Hypercognitive’ snobbery is moral blindness, an elitism which asserts that some are less worthy of moral concern than others, ‘them’ and ‘us’. The book affirms a common humanity. Our job as staff is to preserve identity.

This is a good book. It will not change base metal into gold but via a mosaic of ideas introduces a way of thinking. Osten-sibly it is about dementia, actually it is about what it is to be human. The view that the person may survive into severe dementi-a is now also receiving attention from psychodynamic psychotherapists who recog-nise that even into the late stages of the disease the ability to forge a relationship is retained. The only way to come anywhere near an understanding of what it may feel like to have a dementia is by close and empathic listening, fusing the horizons of physician and patient even when speech is failing. We underestimate the complexity of the inner life of the patient with dementia. Language is not the whole of the emotional experience. Personhood is retained and to this we should relate.

Jane Garner Chase Farm Hospital Mental Health Unit, The Ridgeway, Barnet, Enfield and Haringey Mental Health Trust, Enfield EN2 8JL, UK Email: janet.jagot@beh-mht.nhs.uk
doi: 10.1192/bjp.bp.106.028878

Handbook of Medicine in Psychiatry

Psychiatrists often find that their knowledge and skills in physical healthcare have slowly diminished over time but are still needed in practice. The Handbook of Medicine in Psychiatry has been written specifically for psychiatrists and provides evidence-based information on the causes, diagnosis and management of many medical disorders.

The American authors tackle the most common medical conditions, which they found in a retrospective case review of over 1000 psychiatric in-patients in the USA who had received a medical assessment. The chapters cover symptoms and signs, such as chest pain and red eye, as well as conditions such as obesity. Each chapter has the same format of clinical presentation, differential diagnosis, risk stratification,
assessment and management in the psychiatric unit, and most have an assessment and treatment algorithm.

There are differences between the UK and USA in guidelines for the management and treatment of certain conditions, and in this book treatments recommended for hypertension differ from those in the current guidelines from the National Institute for Health and Clinical Excellence. The algorithm for the assessment and treatment of chest pain indicates electrocardiography (ECG) only for patients with suspected cardiac ischaemia. In the UK, ECG would also be undertaken for the investigation of other causes of chest pain such as panic attacks and serious conditions such as pulmonary embolism and aortic dissection. If followed exactly, this algorithm might lead to problems with diagnosis.

The chapter on cardiac arrest does not present an algorithm for advanced life support but treatments are shown in tables instead. There are differences from UK practice in recommended medication; for example, the initial dose of aspirin recommended for the treatment of myocardial infarction is given as 325 mg, whereas the recommended dose is 300 mg in the UK. This might not be clinically significant but could lead to confusion. Mannitol is listed as a treatment for constipation and enemas with tap water are recommended for the prevention of faecal impaction in the bedridden; both would be regarded as unusual treatments in the UK.

Psychiatrists who are unfamiliar with UK guidelines and standards of medical practice may not wish to rely solely on this book for medical information. In the UK, this book faces strong competition from the Oxford Handbook of Medicine and the Oxford Handbook of General Practice.

Mark Tuthill  Royal Marsden Hospital, Chelsea, London
Irene Cormac  Rampton Hospital, Retford, Nottinghamshire DN22 OPD, UK
Email: irene.cormac@notts.hct.nhs.uk
doi: 10.1192/bjp.bp.106.026500

R. D. Laing: Contemporary Perspectives

As a sixth-former I was an avid listener to Anthony Clare’s Radio 4 programme ‘In the Psychiatrist’s Chair’. Of the interviews Clare conducted, two continue to stick in my mind: one with Jimmy Savile, the other with R. D. Laing. The image with which Clare left me was that of Laing as a rather romantic, yet tragic individual. Raschid’s edited volume on Laing both reinforced and modified this impression. The volume is divided into three sections: introduction, ideas and therapy, with the second section the longest and, for me, the most worthwhile. The papers are largely either original contributions or derive from the R. D. Laing conferences, organised under the auspices of the Philosophy Special Interest Group of the Royal College of Psychiatrists. There are some papers that deal less directly with the work of Laing: the contributions of Fuchs, Sass and Matthews in particular serve, from different perspectives, as lucid and clear introductions to phenomenology and the philosophy of psychiatry. What is refreshing is that the contributors who engage directly with Laing’s thought and influence are not unquestioning, bedazzled disciples. Many of the papers are critical of Laing’s views on mental illness, psychotherapy, politics and his use and understanding of philosophy. There was a tendency, particularly in Laing post-Divided Self, to romanticise mental illness. This was combined with an aspiration towards transcendence and otherwordliness. The book achieved one very important thing for me personally: it shifted the image of the tormented Laing as interviewed by Clare and replaced it with that of the young army psychiatrist spending hours trying to interview and understand the distressed soldiers under his care. This is the Laing I am left in full admiration of. The tragedy is not so much that of Laing’s own personal life but rather his own seeming loss of this immediate pre-reflective ability to be with and understand people in distress. One could interpret his later work as an attempt to reify, in an increasingly esoteric fashion, that which once came so easily.

There are a few pedantic criticisms of the book. The same point is repeated by different authors in different papers leading to some degree of repetition, not all works cited in the text are referenced and there is no index. Many contributors to the volume also offer an incorrect or simplistic interpretation of Jaspers’ views on understanding those with mental illness. I would still recommend to medical students and trainee mental health clinicians The Divided Self as an account of engaging with those with mental illness and am grateful to R. D. Laing: Contemporary Perspectives for reminding us of the passion of Captain Laing.

Matthew Broome  Section of Neuroimaging, Division of Psychological Medicine, Institute of Psychiatry, De Crespigny Park, London SE5 8AF, UK
Email: m.broome@icp.kcl.ac.uk
doi: 10.1192/bjp.bp.106.024232

Gambling as an Addictive Behaviour: Impaired Control, Harm Minimisation, Treatment and Prevention
By Mark Dickerson & John O’Connor. Cambridge University Press. 2006. 204 pp. £55.00 (hb). ISBN 052184701X

Gambling, albeit a leisure activity for most, can in a significant minority progress to problem gambling or pathological gambling, with wide-ranging adverse interpersonal, financial and social consequences. There is mounting evidence that increased availability and easy accessibility to gambling opportunities can result in increased incidence of problem gambling – a point of particular relevance to the UK, at the present time, given the impending deregulation of gambling legislation as proposed in the Gambling Bill. So too, over recent years, gambling research has emerged to occupy an important place within the field...
of addiction studies. It is against this background that I read this book and I found it to be timely and highly readable.

This monograph is part of the International Research Monographs in the Addictions series (the series editor is Griffith Edwards) and is authored by two experts in the field of gambling research. Although the book has an Australian bias (as the authors, the research described and the policy issues discussed are Australia-based), the theoretical constructs/models discussed and the implications for policy makers are generalisable. This book takes a look at gambling as an addictive behaviour – more specifically at the dimension of self-control over gambling behaviour and the various psychological variables that influence it. The core theme of this book is ‘impaired self-control’ and the authors eloquently summarise key findings from their 5-year gambling research programme. This book consists of eight chapters – all self-contained and well-organised. The authors provide an excellent overview of the key psychological variables that determine self-control over gambling: emotional factors, individual differences, cognitive variables and coping, and they set this in the context of different models of impaired control and two studies that explored this subject. In addition the following topics are also well covered: implications for psychological treatment of pathological gamblers, and the concept of harm minimisation or ‘responsible gambling’. I found the chapter presenting a case study of the implementation of harm minimisation strategies in Victoria, Australia to be particularly fascinating, because of its relevance to clinicians and policy makers in the UK. This book provides interesting insights into the dimension of impaired self-control, and succeeds in highlighting its key role in the psychological conceptualisation of gambling and addictive behaviours in general.

All in all, a good read, although some of the theoretical debate presented could be intellectually taxing. At a price of £55, it may be a bit over-priced, and it may also not appeal to the non-specialist. Despite the above-noted criticisms, this book is likely to be of value to those with an interest in gambling research and policy.

Sanju George  Locum consultant in addiction psychiatry, Birmingham and Solihull Mental Health Trust, B37 7UR. Email: sanju.george@talk21.com
doi: 10.1192/bjp.bp.106.026187

The Psychiatry of Intellectual Disability

Few books have been published on the psychiatry of intellectual disabilities. This book is more succinct, better presented and more consistent in the quality of writing and information than the corresponding title in the College Seminars series.

Despite this overall recommendation, the book has several flaws. It is presumably aimed primarily at trainees in the subspecialty. Clearly not a reference book, it should have been a more practical manual. Luty & Cooper’s chapter on older people with intellectual disabilities gives useful guidelines for assessment and management but other chapters are much less practice-orientated. Although certainly better than many previous texts in limiting the content regarding general ‘handicap’, there are still too many references to the primary health-care of people with intellectual disabilities. The era is long gone where psychiatrists in intellectual disability act as pseudo-general practitioners. It was also unnecessary to have paragraphs on such obscure conditions as Coffin-Siris syndrome. The reality is that most referrals to psychiatrists in intellectual disabilities are for problem behaviours. The trainee must learn that it is not their responsibility to solve these problems with medication alone but they should act as the only professional who has the training and expertise to take the holistic overview of the patient in biopsychosocial terms. There is also uncritical acceptance of the vague, catch-all term of ‘challenging behaviour’, which hampers rather than helps approaches to problem behaviours. It was a mistake therefore to include a chapter on medication without one on basic psychological assessments and interventions that a trainee needs to understand and implement.

Roy’s chapter on multidisciplinary working gives an unjustifiably rosy view of the current state of (dis)organisation of services. There are undoubtedly good working relationships between hard-working and committed professionals in learning disability services but good intentions do not compensate for lack of focused working. It is scandalous that, 15 years after its introduction, the care programme approach (CPA) has not been implemented nationwide for people with intellectual disability and mental health problems. For those who seek to improve the rights of people with intellectual disability one powerful starting point would be to demand that such people with mental health problems should have their care coordinated through the recognised national standard of the CPA. It is not good enough therefore that CPA is described in this book as ‘useful’ rather than ‘mandatory’. Community learning disability teams, which vary in focus and make-up throughout the UK, are also blithely described as ‘useful’ without any recourse to evidence of service delivery models.
In summary, although this is perhaps the best introductory short text available in this sub-speciality, it will be unlikely to improve mental healthcare services and service delivery for people with intellectual disability.


Attachment from Infancy to Adulthood: The Major Longitudinal Studies

Tracing the historical roots of attachment theory though its evolutionary stages, Attachment from Infancy to Adulthood brings the reader up to date with recent developments in the area as well as providing thoughts for the way forward into the future. It benefits greatly from contributions by people whose names are instantly recognisable for their seminal work within the area of attachment theory and it is interesting to learn about their diverse backgrounds and what motivated their interest in the subject. It is perhaps this eclectic mix of experience, and the synthesis of key disciplines including ethology, behavioural psychology and linguistics, that have enabled attachment theory to progress and develop to the extent it has over the years.

In addition to reporting findings from major longitudinal studies carried out in America and Europe, the authors discuss the many and complex methodological issues inherent in this type of research, particularly in relation to studying development across the lifespan. They examine the effects of mediating variables and their influence on the relationship between infant attachment and adult outcomes. The studies also show that it is not only mothers but fathers and, as discussed in chapter 7, multiple caregivers in the Kibbutz setting in Israel who influence and contribute to attachment outcomes in later life. Another important dimension, as outlined in chapter 11, is a focus on children who experience disruptions as a result of their placement in foster care.

The book is accessible and is written in an approachable manner that will appeal to students, researchers and others at various stages in their careers. One criticism I have is that the extremely naive reader has to wait until chapter 10 before a full description of the ‘strange situation’ experiment, referred to throughout the book, is given.

As a proponent of longitudinal research and a user of the recently introduced Northern Ireland Household Panel survey, I found this book interesting and informative about the relationship between early attachment processes and outcomes in later life. As a parent it has caused me to study, somewhat warily, my own adult children’s attachment behaviours!

Krisztina L. Lloyd Queen’s University Belfast, Institute of Governance, Belfast BT7 1NN, UK
Email: k.lloyd@qub.ac.uk
doi: 10.1192/bjp.bp.106.026203

Prevention and Treatment of Suicidal Behaviour: From Science to Practice

Prevention and Treatment of Suicidal Behaviour is an accessible book that will appeal to those working in the field and those with a more casual interest. From the perspective of a trainee in psychiatry, it also clarifies the rationale behind the themes of the National Suicide Prevention Strategy for England (Department of Health, 2002). The editor, Keith Hawton, has drawn on the expertise of international authors; chapter by chapter they describe and critically appraise the evidence base, offer practical clinical guidance and identify areas in need of research.

Strategies for dealing with high-risk groups are addressed through the identification of risk factors for suicidal behaviour and chapters focusing on mental health service users, prisoners, and those who misuse substances. Qualitative research is used to identify population-based strategies for reducing suicide, which might also represent logical steps towards tackling mental illness in general. Other chapters explore the population-targeted interventions of restricted access to the means of suicide, the influence of the media’s reporting, and controversies surrounding antidepressant use and suicide rates.

In addition to the two-pronged epidemiological approach to the level of intervention, the authors also present a biopsychosocial exploration of suicide prevention strategies. Psychologically, the ‘entrapment model’ and the role that traumatic stress has in suicidal behaviour are emphasised. Biologically, the focus shifts to descriptions of neurobiological and genetic aspects of the predisposition to suicidal behaviour. The growing concerns about self-harm and psychosocial interventions intended to reduce repetition
are explored from adolescent, working-age and older-aged adult perspectives. To complete this 360-degree analysis, voluntary services and those bereaved by suicide are emotively discussed.

Keith Hawton has thoughtfully structured the book, allowing its themes to be developed in subsequent chapters and the authors to present differing critical appraisals. Midway through, I did become a little pessimistic about research into initiatives to prevent suicide, with the authors’ repeated criticism of the lack of randomised controlled trials and the unacceptably low power of the existing trials. However, some optimism is introduced with the stance that preventive initiatives that are not based on evidence from clinical trials can contribute to our understanding of this area.

Inevitably, those familiar with the *International Handbook of Suicide and Attempted Suicide* (Hawton & Van Heeringen, 2000) will find some themes repeated, but *Prevention and Treatment of Suicidal Behaviour* is by no means a concise or rehashed version of this earlier book: it is a useful text with important ethical, societal and psychiatric messages.

---


**Clare Nolan** Leeds Mental Health Teaching NHS Trust, Leeds, UK. Email: clarenolan@hotmail.com doi: 10.1192/bjp.bp.105.019976

---

**Recovery Beyond Psychiatry**

By David Whitwell.


David Whitwell and his publishers are to be congratulated for producing a book of rare and companionable honesty which, in being personal and specific, offers insight into the experience of every thoughtful clinician. To my knowledge this book is unique in being the reflections of a self-critical and highly experienced practitioner refracted through an understanding of the recovery movement and leading to a personal revaluation of practice.

It is also a paradoxical text that resists many of our scientific conventions. It is written simply and with restraint, lacking the dense referencing and citation that scholarly works depend upon for their credibility — as such it is radically unimpressive and befriending of the reader. The dedication to his family is a reminder that psychiatrists are people too, and more than a few have been touched deeply by the same issues as those they work with.

David has long been troubled by the gap between our apparent knowledge (what he calls ‘naïve psychiatry’) and our ineffectiveness in producing recovery through conventional psychiatric treatment, but found it ‘easier to help people in distress once I had acknowledged my doubts’. His experience was that in learning how to be less knowledgeable he became better connected to the reality of people’s lives and struggles. In turn he describes discovering that a recovery-based approach, focusing on people’s aspirations, hopes and needs, and supporting the active role of the individual in their own recovery, on their own terms, was a better way to work.

In many ways this is a companion text to *Postpsychiatry: Mental Health in a Postmodern World* (Thomas & Bracken, 2005) which sees us as being caught up in and confined by science-based approaches that focus on the deficits of individuals and resort to technical solutions which relegate meanings, values and the social context to secondary consideration.

The Royal College of Psychiatrists has set ‘recovery’ as the theme for its annual meeting in 2007 and this will offer ample opportunity for both positive testimony and critical evaluation. *Recovery Beyond Psychiatry* is an unusual, welcome and timely publication, which is a stimulus to this developing discourse and deserves wide readership and reaction.


---

**Glenn Roberts** Winford House Hospital, Dryden Road, Exeter EX2 5AF, UK. Email: glenroberts@doctors.org.uk doi: 10.1192/bjp.bp.106.021675

---

**Psychiatric Disorders and Pregnancy**


In recent decades the relationship between childbirth and psychiatric disorders has come to greater prominence with postnatal (post-partum) depression becoming a generally recognised, if clinically imprecise, term. This has brought undoubted benefits in the fight for services and in reducing stigma for women with mood disorders at this time. A focus on postnatal depression, however, has not been without problems. First, the range of important disorders occurring in relationship to childbirth is
far greater than non-psychotic episodes of depression – the weight of evidence suggests that bipolarity has a specific relationship to the post-partum period for example. Second, a focus on the post-partum period has diverted attention from the importance of psychiatric disorders that occur in pregnancy, an issue which is addressed in this book.

*Psychiatric Disorders and Pregnancy* sets out to redress the balance and covers a wide range of areas – from screening for risk of severe post-partum illness, through specific conditions (including mood disorders, psychosis, eating disorders, personality disorders, post-traumatic stress disorder and substance use disorders) to specific modes of treatment and models of healthcare. Although focused on pregnancy, many chapters also deal with the postnatal period. In addressing these issues the book is to be commended and there is much of interest here for both the specialist and the casual reader. In this respect I feel it is a real shame that the book has been priced at £75, putting it beyond the reach of anybody but the most committed perinatal specialist and libraries.

The predominantly British authors include many well-recognised experts in their fields and a number of chapters are of particular interest and would reward revisiting. As with any multi-author book covering a specific area there is some repetition, and there is always the danger that advice regarding the safety of medications during pregnancy and breast-feeding is quickly out of date – recent data on paroxetine, for example, were clearly only available after publication deadlines.

Although I agree with the editors that mood disorders during pregnancy have been neglected in clinical practice and research, I believe it remains to be demonstrated that this is a period of higher risk for episodes of major depression compared with the post-partum period (as opposed to depressive symptoms more generally). In this area as in others, more research is clearly needed. In fact, the abiding impression I have from reading this book is the exciting opportunities that exist to further research on psychiatric disorders in relation to pregnancy. If this book goes someway to highlight the areas where work needs to be done, then it has served an important function. I just wish it were cheaper.

Ian Jones  Neuropsychiatric Genetics Unit, Department of Psychological Medicine, University of Wales College of Medicine, Cardiff CF14 4XN, Wales, UK. Email: jonesir1@cf.ac.uk
doi: 10.1192/bjp.bp.106.026195
From the Editor’s desk

PETER TYRER

LEAVING THE TRAMLINES

All good research challenges existing orthodoxy and there are several papers in this issue that illustrate this well. Many years ago I had a patient under my care who had been diagnosed with breast cancer and who also developed schizophrenia round about the same time. Whenever I saw her I was concerned about her apparent indifference to her cancer, while she was equally concerned about my indifference to her fixed delusion that she was about to be murdered. I took my concerns to her surgeon; he was openly flabbergasted that she was still alive as her cancer was inoperable when he first saw her six years earlier. ‘You and your colleagues must do some research on this; there must be something in schizophrenia that protects people against cancer’, he told me with some feeling. The neurodevelopmental genetic hypothesis suggested by Levav et al (pp. 156–161) may have been close to the sort of thing he had in mind. The association between diet, serum lipids and self-harm is even more counter-intuitive but the studies by Garland et al (pp. 112–117) and Hallahan et al (pp. 118–122) suggest that it should be taken very seriously. In an era in which everyone is being exhorted to lower their cholesterol intake to protect against cardiovascular disease it is salutary to be reminded that there may be problems in having lower levels of both cholesterol and essential fatty acids – and it also helps to justify my penchant for chocolate eclairs.

The finding of reduced self-harm in those treated with omega-3 fatty acid supplements (we do need a shorter identifying name for these additives – why not fish?) is a striking one but the sample is small and replication is needed (Hallahan et al., pp. 118–122). Ireland is known for both its export of fish and low domestic fish consumption (Fleming et al, 1997) and it would be interesting to know whether there is, as Peet (2004) alleges for depression, variation in figures for self-harm across countries with different fish consumptions.

Diet also features in the recommendations in the paper from Howard et al (pp. 129–134) showing that hip fractures are more common in those with schizophrenia and those taking prolactin-raising antipsychotic drugs. The authors’ suggestion for screening for osteoporosis and other health checks in schizophrenia echoes the need to keep both the regulatory policies for treatment (Barbui & Garattini, pp. 91–93) and the medical needs of those with severe mental illness (Osborn et al, 2006) under constant review. The tramlines imposed by specialist requirements can be a handicap for those with mental illness in medical settings too; the low rate of treatment adequacy for common mental disorders is a disturbing European statistic (Fernández et al, pp. 172–173).

So do not too get stuck in those familiar grooves. Break out into novelty. Perhaps it is no accident that those great tramline leavers, the Irish, have three papers in this issue – this illustrates the lofty position Ireland holds in the international psychiatry stakes (Marusic, 2004). How do they manage to do this while eating hardly any fish?

CONSEQUENCES OF PUBLICATION IN THE BRITISH JOURNAL OF PSYCHIATRY

It is often assumed that the authors of original research papers attain their peak self-esteem when their work is recognised by the world. This was put to me most vividly by a colleague who told me the nearest he had ever come to ecstasy was when he received a letter from the Archives of General Psychiatry saying that his paper had been accepted for publication. I gave him a moment or two to allow a change of name to the British Journal of Psychiatry but he stayed silent – perhaps this was the next challenge. To determine the high points of publication, we have been asking our authors to let us have details of the consequences of appearance of their papers in the Journal; we thank them for their continuing feedback. What is extremely encouraging is the international impact of publication, with so many papers leading to new links between researchers, including a very interesting one between Africa and Iran (Assadi et al, 2006). The publication by Dean et al (2006) led to a conference on female offenders, the replication of the gene–environment interaction with the serotonin transporter gene in depression led to reports in Time magazine and a nationwide shortage of reprints (Wilhelm et al, 2006), and the paper by Smit et al (2006) was a core one in helping to develop a Policy White Paper in The Netherlands, where both the evidence of health and economic gain of preventing depression are being embraced by the government in a forthcoming health initiative. There was so much media interest in the paper on month of birth in suicide (Salib & Cortina-Borja, 2006) that the senior author became quite overwhelmed, ending his letter ‘I have published a number of papers in BJP over that last 20 years, the last one is very special’. Now this is getting quite close to ecstasy too, and I suspect fish fats are not the reason.


