Schizophrenia trials: past, present and future

CLIVE E. ADAMS

*Co-ordinating Editor, Cochrane Schizophrenia Group, Chair Adult Psychiatry,
University of Leeds, Leeds, UK*

THE PAST

Richard Doll, a pioneer of the randomised controlled trial, recently described the situation before random allocation became widely adopted.

"...new treatments were almost always introduced on the grounds in the hands of Professor A or in the hands of a consultant at one of the leading teaching hospitals, the results of a small series of patients (seldom more than 50) had been superior to those recorded by Professor B (or some other consultant) or by the same consultant previously. Under these conditions variability of outcome, chance and the unconscious (let alone the conscious) bias in the selection of patients brought about apparently important differences in the results obtained." (Doll, 1998).

In order to evaluate the effect of a specific treatment, the consequence of using this treatment in a group of people should be compared with not using it in a control group. The only difference in the two groups should be that one receives the treatment under evaluation, and the only way to achieve this is through random allocation. It is now over 50 years since the value of streptomycin in tuberculosis was demonstrated in the first randomised trial (Medical Research Council Streptomycin in Tuberculosis Trials Committee, 1948) and this methodology is now seen as the 'gold standard' means by which different forms of mental health care are evaluated (WHO Scientific Group on Treatment of Psychiatric Disorders, 1991).

The introduction and uptake of randomised controlled trial methodology in medicine coincided with a revolution in the care of people with schizophrenia. Drug treatments began to appear which seemed to dramatically improve the mental state of those for whom little hope had previously existed. (Freeman, 1958) Psychiatrists welcomed the randomised trial and a tradition of evaluative research was greatly strengthened.

The creation of registers of randomised controlled trials to assist the preparation and maintenance of systematic reviews affords an opportunity to assess the quality and content of evaluative research within well-defined sampling frames (Adams & Gelder, 1994) When this large survey was planned there were examples of surveys of randomised trials within specific journals (Ahmed *et al.*, 1998; Lent & Langenbach, 1996; Fahey *et al.*, 1995; Doll, 1998) but research into the quality and content of trials within specific health care specialities was, and is, less common (Chalmers, 1986; Vandekerckhove *et al.*, 1993; Silagy, 1993; Chalmers *et al.*, 1986; Nicolucci *et al.*, 1989; Cheng *et al.*, 2000). Full details of this particular survey are published elsewhere (Thornley & Adams, 1998a).

All reports on the Cochrane Schizophrenia Group's Register were eligible. The Register contains reports of published and unpublished randomised trials and controlled clinical trials (parallel group comparative studies where treatment allocation is not explicitly stated to be random). These studies relate to the care of people with schizophrenia and other non-affective psychoses. Key journals were identified and hand-searched from 1948 to the present. Conference proceedings were also hand-searched. Comprehensive searches of Biological Abstracts, CINAHL, *The Cochrane Library*, EMBASE, Lilacs, PsycLIT, Psyndex, Medline and Sociofile were undertaken. The resulting 30,000 electronic records were

Indirizzo per la corrispondenza: Professor C.E. Adams, Co-ordinating Editor, Cochrane Schizophrenia Group, 15 Hyde Terrace, Leeds LS2 9LT, (UK).

Fax: +44-(0)113-233.2723
E-mail: ceadams@cochrane-sz.org
Paper presented at the SIEP Congress, October 2001, Rome.

checked for duplicates before two trained selectors highlighted studies that were clearly or possibly relevant. Full copies of these reports were then acquired (6000), re-inspected and, if meeting the inclusion criteria, added to the final register. Time constraints forced an arbitrary limit of 2,000 trials (in 2275 reports) for this survey. Papers were read and details of type and date of publication, country of origin and language were reliably extracted. A measure of methodological quality was used, based on each trial's description of randomisation, blinding and treatment withdrawals (Jadad *et al.*, 1996). This measure was chosen for its validity (Jadad *et al.*, 1996) ease of use, and because low scores, indicating poor quality of reporting, have been shown to be associated with an increased estimate of benefit (Moher *et al.*, 1998). Study size, details of treatment setting, participants, interventions and outcomes were also recorded.

Coding of variables was reliable. There was over 90% agreement in all but 'numbers completing the study' (70%) and listing of outcome instruments, where, in 10% of reports, the principal rater failed to identify one of the scales, often amongst several used.

Over 95% of the sample was published in English and the vast majority of reports relate to research undertaken in North America. Trials from outside of the USA and the EU account for less than 10% of the sample. The majority (85%, 1940/2,275) of reports were fully published in journals, whilst the remainders were conference abstracts (253, 11%), or published as letters, books, chapters, or product monographs (82, 4%). On average, there were just fewer than three reports of trials per year in the early 1950's, and just over 90 new trials annually by the late 1990's.

After determining, to the best of our ability, which reports were multiple publications of the same trials, a new sample of 2,000 *studies* was created. When multiple publications of the same trial reported different aspects of the study, variables were combined in order to make a composite whole.

Unsurprisingly (given the focus of the group's interest) most of the trials involved people with schizophrenia. Within these 2,000 trials, 1,326 (66.3%) were carried out in hospitals or institutions, 272 (13.6%) in the community, and 11(0.6%) in both. Study location was unclear in 391 (19.6%) trials. Trialists mentioned a statistical power calculation in only twenty trials (1.0%).

Generally, the quality of reporting in schizophrenia trials, as rated by the Jadad scale is poor (mean 2.2) and there is little change over time. Quality of reporting is associated with direction of effect; poorly reported trials

are more likely to describe positive effects than studies that achieve a score of five (Moher *et al.*, 1998).

The average number of people in the 1,941 trials that reported sample size, was 62 (median 60, mode 48). Over 50% of schizophrenia trials randomised less than 60 people. For a large difference between experimental and control groups, for example 20% in any binary outcome, a standard power calculation (power of 0.85, p=0.05) (Clark *et al.*, 1970) recommends the inclusion of 150 people per arm. Just 58 trials (3%) had sample sizes of above 300. Study size, is not increasing to an important extent across time. The duration of treatment for people with schizophrenia is obviously important. Over half of schizophrenia trials, however, have follow up of six weeks or less.

Many factors determine the proportion of participants randomised who drop out or withdraw early from
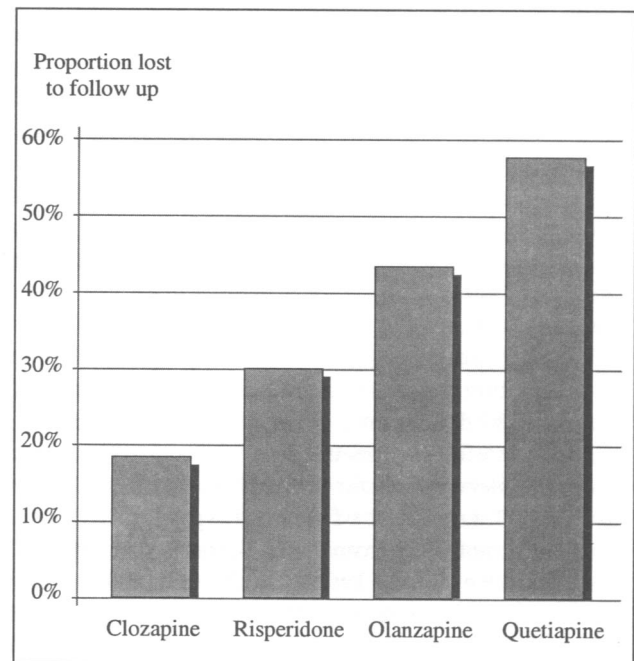

Figure 1. - *Proportion of trial participants leaving study by 6-12 weeks ('atypical' vs typical antipsychotics).*

treatment. We took a sub-sample of trials dealing with novel antipsychotic drugs. These trials are all of short duration and the numbers of participants leaving early, large. It is difficult not to assume that this dramatic loss to follow up is a result of trial design, rather than a reflection of the acute disturbance of those who have given informed consent to participate.

145

Interventions were classified into 'drug' (86.3% of total sample), 'therapy' (any 'talk-based' treatment) (8.2%), 'physical' (for example ECT and psycho-surgery) (3.9%), 'policy/care packages' (for example case management and assertive community team treatment) (8.6%) and 'other' (1.9%). 1725 (86%) of the 2000 trials evaluated the effects of 437 different drugs. The use of haloperidol in trials over time and is increasing. The use such a potent cause of adverse effects (Joy *et al.*, 2001) may be fostered by the stipulations of some drug regulatory bodies. It is of increasing concern that the regulatory authorities have "become servants of industry" (Horton, 2001) despite not entirely reassuring protestations to the contrary (Galson *et al.*, 2001).

Outcomes are difficult to rate reliably. For example, inter-rater agreement for non-scale outcomes was only 40%. Reliability of listing of outcome instruments, however, was better. In only one in ten reports did the principal rater fail to identify a single scale. 510 (25%) studies did not use rating scales to measure outcomes, but the remaining 1490 trials used 640 different instruments. Lack of statistical power was reflected in the use of an extraordinarily large number of rating scales. It is often possible to 'achieve' statistical significance on these fine-grain measures with small numbers. These devices by researchers leave unaddressed the clinical interpretation of these measures and the fact that scales are rarely used in clinical practice.

Further work by this group attempted to determine whether there was an association between using an unpublished outcome rating scale and finding a significant effect of treatment in randomised controlled trials of interventions for schizophrenia (Marshall *et al.*, 2000a). Two teams of raters inspected 300 trials randomly selected from the Cochrane Schizophrenia Group's Register. Trials were excluded if non-therapeutic or unavailable in English. The teams identified all comparisons between treatment and control groups rated by scales. The publication status of each scale was determined, blind to the results of any comparisons. Claims of a significant treatment effect were reliably recorded. The results were that 456 comparisons were identified in the 193 trials meeting inclusion criteria. Trialists were more likely to claim that a treatment was superior to control when the comparison was made using an unpublished scale (Relative Risk 1.37, 95% CI 1.03-1.83). This effect increased when claims of significance were restricted to comparisons based on summary scores at the end point of the trial (Relative Risk 1.94, 95% CI 1.45-2.60). The effect was

particularly marked in trials of non-pharmacological treatments where 56% of significant results were obtained using unpublished scales. Our conclusions were that unpublished rating scales are associated with more positive estimates of effect than those derived from published scales (Marshall *et al.*, 2000a).

Overall, our sample of 2000 trials is likely to be biased in some respects. Searching was largely, but not exclusively, in English, and our ability to code articles in other languages, limited. However, it is unlikely that there are enough undiscovered large, high quality trials published in other languages to substantially change the results of this survey. The sample studied is probably biased less by language and publication status than the trials readily available through Medline. The first 2000 trials entered on the Cochrane Schizophrenia Group's register were also a sub-sample of around 6000 currently identified (Anonymous, 2001). High availability of a report (indexing of a journal on the databases searched, publishing in the journals hand-searched, positive results (Easterbrook *et al.*, 1991; Stern & Simes, 1997)) would have increased the chance of early entry on to the register and therefore of inclusion in the survey. Smaller, more recent surveys based on samples from all 6000 studies (Wahlbeck *et al.*, 2000) however, suggest that the original sample of 2000 remains representative of all trials currently identified on the Cochrane Schizophrenia Group's Register.

On average, trials relevant to schizophrenia are conducted in the USA, involve less than 50 people, fail to report a power calculation, are most relevant to the pharmacological care of people in hospital, lasts six weeks and yields inadequate reports on scale-derived outcomes of limited clinical utility on between 50-70% of randomised participants.

Schizophrenia trials are difficult to conduct. The illness, affecting only one percent of people at some point in their lives, may lead to chaotic behaviour and disordered thinking, erosion of insight and, often, considerable mistrust of health care professionals. These factors, along with the relatively weak tradition of multi-centre, multi-national trials, may have promoted limited study size and duration. In addition, the pursuit of reliable outcome measurement has spawned an extraordinary number of rating scales, one third of which were unpublished at time of use, and therefore of unknown validity. This endeavour probably reflects not only the researchers' will to quantify objectively, but also the lack of statistical power of small trials to assess effects on more clinically relevant outcomes. It is easier to achieve *statistical* significance on these measures of

doubtful *clinical* relevance with small numbers of participants. Underlining their clinical irrelevance, scales are rarely used in clinical practice. In any event, the low quality of reporting in these trials seems likely to result in an overestimate of benefit (Moher *et al.*, 1998; Schulz *et al.*, 1994; Gotzsche, 1989). Hopefully, this mediocre reporting will change with wider adoption of the CONSORT recommendations (Moher *et al.*, 2001). Although quality of reporting has been a proxy measure of the methodological quality of a trial, it is important to be alive to the possibility that cosmetic adherence to CONSORT requirements will mask unreliable studies.

Increasing complexity of trial design must contribute, at least in part, to attrition far greater than would be seen in routine clinical care. As a result, 30-60% of data on the effects of the novel antipsychotic drugs are often based on untested assumptions about the fate of people once they left the study, applied by researchers employed by organisations with substantial pecuniary interests in the trials' results. Also, as trial complexity increases, so does the expense of carrying out randomised trials relevant to the care of people with schizophrenia. US-based public and private institutions are increasingly taking on the financial burden of supporting schizophrenia trials. Strict entry criteria to these trials, however, makes the applicability of results problematic, even within the USA (Wells, 1999), and only 2-4% of the world's population of people with schizophrenia live in North America.

The findings of this survey are as worrying, if not more worrying, than those for other specialties within health care (Chalmers *et al.*, 1986; Cheng *et al.*, 2000; Silagy, 1993; Silagy *et al.*, 1996b; Thornley & Adams, 1998b) As the numbers of trials relevant to schizophrenia rises, the need for up-to-date systematic reviews of these studies increases.

THE PRESENT

High quality systematic reviews can never fully compensate for limited, poorly designed, poorly conducted and poorly reported studies – but they may be able to go part of the way.

The Cochrane Collaboration was founded in 1993. It is an international group producing, maintaining and disseminating systematic reviews of the best evidence of the effects of health care (Chalmers *et al.*, 1992) It has been likened in importance to the human genome project (Naylor, 1995). In 1994, Cochrane Schizophrenia Group was registered with the Collaboration (Cochrane

Collaboration Schizophrenia Review Group, 1994) and since has been collecting relevant randomised control trials and producing and disseminating these within the Cochrane Library (www.cochrane.de).

A systematic review is a review of a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyse data from the studies that are included in the review. Statistical methods (meta-analysis) may or may not be used to analyse and summarise the results of the included studies. A Cochrane Review is a systematic, up-to-date summary of reliable evidence of the benefits and risks of healthcare. Cochrane Reviews are intended to help people make practical decisions and are disseminated within the Cochrane Database of Systematic Reviews (Update Software Ltd, 2002). Every attempt is made to select, extract, present and report on the data in as unbiased a fashion as is possible. As these reviews are published in electronic form, they can be maintained, in the light of new evidence. Cochrane systematic reviews are more comprehensive and objective than those seen in standard journals (Jadad *et al.*, 1998).

These systematic reviews, however, are based on the trials described above. The question remains whether they can be of value in the light of the limited quality and utility of the 'building blocks' of current randomised trials. In this section I argue that learning from the totality of evidence of past trials informs practice today, and helps study design of the future. Some examples may be helpful.

A person with schizophrenia may have been started on an antipsychotic drug such as chlorpromazine. The patient, knowing that chlorpromazine has been used for 50 years, could have asked the attending clinician what was the predicted value of this drug in terms such as "what are my chances of getting better?". It could help the clinician and patient to know that systematic reviews of the best available evidence, albeit limited evidence, suggest that chlorpromazine reduces relapse over six months to two years (Relative Risk (RR) 0.65 CI 0.5-0.9, Number Needed to Treat (NNT) 3 CI 2.5-4) (Thornley *et al.*, 2001). As patient has lost their job, some form of work rehabilitation may also be indicated. It would probably be helpful to know that supported employment in normal work setting is significantly more effective than pre-vocational 'sheltered' training (RR unemployed at 18 months 0.76 95% CI 0.64 to 0.89, NNT 4.5) (Crowther *et al.*, 2001). Six months later chlorpromazine has not been as helpful as was anticipated. The doctor thinks a switch to an atypical

drug, such as risperidone, is indicated. Recognising that responses are always idiosyncratic, it could help the doctor to know that, on average, risperidone seems to be statistically significantly better than rather large doses of haloperidol for its population group, but for a clinically irrelevant outcome (20% shift on the PANNS scale) and on a biased data set of poorly reported studies (Kennedy *et al.*, 2001) This humbling knowledge, clearly stated in the review, may help everyone set realistic expectations, although real-world clinical life can always spring pleasant surprises. Some form of case management may be indicated at this stage. Knowing that the low grade approach to case management nearly doubles the rate of hospital admission (Marshall *et al.*, 2000b) the team may wish to opt for a more assertive approach. After all, people receiving assertive care are more likely to remain in contact with services than people receiving standard community care (RR 0.51, 99%CI 0.37-0.70) and are less likely to be admitted to hospital (RR 0.59, 99%CI 0.41-0.85) (Marshall & Lockwood, 2000).

The evidence from randomised controlled trials relevant to people with schizophrenia is limited and rife with bias. Nevertheless, systematic reviews can go some way to cut through that bias and present useful information to help inform treatment today. As the same data may mean different things to different people, Cochrane reviews endeavor to present information as clearly as possible in order to allow the reader to make an informed decision. Cochrane systematic reviews tend not to 'guide', merely inform. Although the concept of evidence based medicine has been partially hi-jacked by those with a propensity to a more traditional approach to guidance of clinicians (Adams & Gilbody, 2001) those working within the Cochrane Schizophrenia Group still feel strongly that provision of the best available evidence in order for the reader to make up their own mind remains best policy.

All Cochrane reviews point to the future. Many summarize the best evidence, but need more to draw firm conclusions, often on outcomes of real relevance to patients and carers. Others do not include data from any trials at all. For example, a recent review on use of "as required" medication for people coming into hospital, failed to identify any relevant randomised studies (Whicher *et al.*, 2002). This very common clinical practice is unsupported by good evidence. It would seem reasonable that the specialty of mental health care, with its strong tradition of evaluation and basis in good science should treat the most vulnerable patients on the basis of not only judgment and wisdom, but also excellent science. More trials are needed, but not just

like what has gone before. We need, and should expect, randomised trials that follow the 'real world' needs of patents and carers, and not of researchers and industry.

THE FUTURE

With notable exceptions, the quality of trial evidence relevant to the care of schizophrenia could be better and systematic reviews can only go part way to address the inadequacies of trials. Often further randomised trials are needed. The great majority of randomised research in mental health relates to the 'efficacy' of drug treatments. These 'explanatory trials' are required for drug licensing and test the potential for a drug to work under ideal conditions. They seek to maximize internal validity so that any difference can be attributed to the treatment and not some external (confounding) factor (Schwartz & Lellouch, 1967). Whilst efficacy studies help establish if an intervention can work under ideal conditions, they may tell us little about how the results relate to real world practice (Hotopf *et al.*, 1999; Simon *et al.*, 1995; Gilbody *et al.*, 2002).

People agreeing to be randomised may well be different from the average person seen in day-to-day care (Bowen & Barnes, 1994). Employing additional, rigid entry criteria compounds this bias. Restrictive entry criteria may ensure relative homogeneity within the group of people recruited, but the resulting observations may have limited applicability (Britton *et al.*, 1999). Highly restrictive entry criteria helps ensure that studies are generally under-powered and therefore stand little or no chance of providing answers to clinical questions. Partly to compensate, rating scales are used as measures of treatment effect. These may be able to detect real, but subtle, differences that are the result of the treatments being compared. It is, however, unclear whether subtle but statistically significant changes on complex rating scales are the most important criterion of success for clinicians, patients and their carers (see below). Complex psychopathology scales are rarely employed in day-to-day care. Even researchers are unlikely to know whether certain changes or scores are worth achieving. Simple outcomes relating to real-life events, for example, hospital discharge, relapse, employment status and ability to live independently, are ignored in most studies despite the fact that these will be of interest to clinicians, people with schizophrenia, economists and policymakers.

Randomised studies that employ rigid inclusion and exclusion criteria, rigid treatment protocols and multiple,

complex outcome measures are difficult and expensive to undertake. Such explanatory trials probably also prove difficult for participants. A device often used by trialists to counter the expected attrition, is the use of short-term outcomes but nothing is entirely effective at arresting the haemorrhage of participants from recent drug studies. It is chastening to think that the licensing and subsequent use of drugs, for example olanzapine, are based on largely short term trial data from which 42% of people decided to leave or were withdrawn before three months treatment was complete (Duggan *et al.*, 2001). Then those undertaking the analysis, supported by the industry and therefore working under an impossible conflict of interest, would reconstruct the 'intention-to-treat-analysis', incorporating untested assumptions on about half the data.

The recognition of the shortcomings of efficacy trials has led some to conclude that randomised trials have no place in the evaluation of effectiveness in the real world.(Vandekerckhove *et al.*, 1993). Instead, observational data (often from large, insurance claims databases) have been used in 'outcomes research'. This work hopes to establish the effectiveness of treatment in the context of routine care (Egger *et al.*, 1997) and has been used to generate influential guidelines for treatment.(Silagy *et al.*, 1996a). There have been calls for such research designs to be used in mental health,(Wells, 1999a) however, such observational data are prone to numerous biases, not least the inability to extinguish confounding and selection biases (Sheldon, 1994).

As an alternative to regressing to the observational study, researchers have proposed the preservation of randomisation and the design of research that more closely replicates routine care – the *pragmatic or real world trial* (Hotopf *et al.*, 1999; Roland & Torgerson, 1998; Gilbody *et al.*, 2002). Efforts have been made to make trials elegantly simple so that recruitment is easy, involving minimal effort on the part of participants and requiring minimal deviation from routine care (Dickersin *et al.*, 1985). Such trials have not been widely adopted in psychiatry, although rare examples do exist (Simon *et al.*, 1996; Huf *et al.*, 2002). Unlike the research outlined above, pragmatic trials in schizophrenia should answer questions of direct relevance to clinicians, patients and wider society. Pragmatic trials should be conducted within existing health services. The patients who participate in pragmatic trials should reflect the heterogeneity of people with schizophrenia in routine practice that will ultimately receive the treatment. Exclusion criteria should only be specified if absolutely

necessary and should ideally be no more than an unwillingness or inability to offer informed consent or a specific contraindication to one of the trial interventions. Age, sex or the presence of comorbid disorders should not be considered exclusion criteria, unless they represent some treatment contraindication.

Both the intervention and the comparison treatment should be possible within routine practice. Thus, a new treatment might be compared with the standard care that it would replace if adopted into routine practice. When the prospect of randomisation to a treatment is unattractive or unacceptable to patients or clinicians, or where there are strong prior beliefs about one treatment being superior, then few will agree to participate. Treatment alternatives, therefore, require careful consideration and should reflect genuine 'equipoise' on the part of clinicians and patients, i.e. a genuine uncertainty about which treatment option is better.

Abstract and complex questionnaires administered by interview have little role in pragmatic trials and should not be the primary outcome by which success or failure of a treatment is judged. Instead, important outcomes relating to global well being, relapse, hospitalisation and social functioning should be recorded. These are often dichotomous outcomes, which can usually be recorded from information included in case notes or hospital administrative records. Costs should be collected using the minimum data needed to provide valid and representative estimates. A pragmatic trial should impose minimal burden on all concerned and outcomes should be of interest to patients, clinicians and decision-makers; all of who may need to be consulted in advance for advice.

Interference in routine care because of participation in a trial should be minimal. The only active decision which a patient or clinician should make is whether to participate and undergo randomisation. After that, the trial protocol should impose minimal restriction upon subsequent management. Those designing trials must balance the possible benefits of double or single blindness with the practical advantages of simplicity of design. With outcomes that may be less 'soft' than those usually recorded (e.g. self-harm, hospital admission, legal problems, living independently); blindness, which is likely to be imperfect anyway, may offer little advantage as regards objectivity of outcome assessment. Double-blinded assessment imposes unrealistic impositions on routine care, through, for example, the issue of coded bottles and pharmacist participation. The deviation from normal patterns of care which result from the imposition of attempts to adhere to double blindness

also threatens the validity and generalisability of any cost data used in the estimation of cost effectiveness.

The high losses to follow-up seen in schizophrenia trials are in part a consequence of the burdens of participation in explanatory studies. If the trial involves onerous time commitments, suspiciously and often incompletely disguised treatments, and long and abstract questionnaires administered by strangers, it should not be surprising if many patients withdraw during the course of the trial. Pragmatic study designs should lead to a control group who experience no more than the expected loss to follow-up in routine care. The simplicity of data acquisition and the non-intrusive nature of follow-up should allow medium and long-term trials to be designed. Additionally, the high attrition rates seen in explanatory trials are in part made up of those patients who discontinue or switch from treatment as randomised. Rather than being excluded by reason of 'protocol violation', every effort should be made to follow up these patients and include them in any analysis of a pragmatic trial. Discontinuation of randomised treatment and its consequences should be seen as an outcome worthy of study, rather than a reason for exclusion from a trial.

Other medical subspecialties now provide examples of collaborative trial work of such simplicity that the findings have direct repercussions on the care of millions of people (ISIS-4, 1995; Collaborative Eclampsia Trial Group, 1995). Specialists in the care of those with schizophrenia should consider and learn from these initiatives. There will remain a role for explanatory trials, as these are necessary steps in examining whether a treatment can work, albeit under ideal conditions. They are not, however, sufficient to judge whether a treatment can work in the real world. Pragmatic trials have represented a major advance in other areas of health care. Not last, and by no means least, it is now the turn of people with schizophrenia to have their care based on truly relevant randomised trials.

REFERENCES

Adams C. & Gelder M. (1994). The case for establishing a register of randomised controlled trials of mental health care. A widely accessible register will minimise bias for those reviewing care. *British Journal of Psychiatry* 164, 433-436.

Adams C. & Gilbody S. (2001). Nobody ever expects the Spanish Inquisition (Python, 1991). *Psychological Bulletin* 25, 291-292.

Ahmed I., Soares K., Seifas R. & Adams C. (1998). Randomized controlled trials in Archives of General Psychiatry (1959-1995): a prevalence study. *Archives of General Psychiatry* 55, 754-755.

Bowen J. & Barnes T.R.E. (1994). The clinical characteristics of schizophrenic patients consenting or not consenting to a placebo controlled trial of anti-psychotic medication. *Human Psychopharmacology* 9, 423-433.

Britton A., McKee M., Black N., McPherson K., Sanderson C. & Bain C. (1999). Threats to applicability of randomised trials: exclusions and selective participation. *Journal of Health Service Research Policy* 4, 112-121.

Chalmers I. (1986). A register of controlled trials in perinatal medicine. *WHO Chronicle* 40, 61-65.

Chalmers I., Hetherington J., Newdick M., Mutch L., Grant A., Enkin M., Enkin E. & Dickersin K. (1986). The Oxford Database of Perinatal Trials: developing a register of published reports of controlled trials. *Controlled Clinical Trials* 7, 306-324.

Chalmers I., Dickersin K. & Chalmers T. (1992). Getting to grips with Archie Cochrane's agenda. *British Medical Journal* 305, 786-788.

Cheng K., Smyth R., Motley J., O'Hea U. & Ashby D. (2000). Randomized controlled trials in cystic fibrosis (1966-1997) categorized by time, design, and intervention. *Pediatric pulmonology* 29, 1-7.

Clark M., Dubowski K. & Colmore J. (1970). The effect of chlorpromazine on serum cholesterol in chronic schizophrenic patients. *Clinical Pharmacology and Therapeutics* 16, 883-889.

Cochrane Collaboration Schizophrenia Review Group (1994). Schizophrenia and the Cochrane Collaboration. *Schizophrenia Research* 13, 185-188.

Collaborative Eclampsia Trial Group (1995). Which anticonvulsant for women with eclampsia? Evidence from the Collaborative Eclampsia Trial. *Lancet* 345, 1455-1463.

Crowther R., Marshall M., Bond G. & Huxley P. (2001). Vocational rehabilitation for people with severe mental illness. *Cochrane Database Systematic Reviews* CD003080.

Dickersin K., Hewitt P., Mutch L., Chalmers I. & Chalmers T. (1985). Perusing the literature: comparison of MEDLINE searching with a perinatal trials database. *Controlled Clinical Trials* 6, 306-317.

Doll R. (1998). Controlled trials: the 1948 watershed. *British Medical Journal* 317, 1217-1220.

Duggan L., Fenton M., Dardennes R.M., El-Dosoky A. & Indran S. (2001). Olanzapine for schizophrenia (Cochrane Review). *Cochrane Database Systematic Reviews* 2.

Easterbrook P., Berlin J., Gopalan R. & Matthews D. (1991). Publication bias in clinical research. *Lancet* 337, 867-872.

Egger M., Zellweger-Zahner T., Schneider M., Junker C., Lengeler C. & Antes, G. (1997). Language bias in randomised controlled trials published in English and German. *Lancet* 350, 326-329.

Fahey T., Hyde C., Milne R. & Thorogood M. (1995). The type and quality of randomized controlled trials (RCTs) published in UK public health journals. *Journal of Public Health Medicine* 17, 469-474.

Freeman H. (1958). The tranquilising drugs. In *Schizophrenia: a Review of the Syndrome* (ed. L. Bellak). Logos Press: New York.

Galson S., Kweder S., Houn F., Raczkowski V., Honig P., Hirshfield S., Kieffer L., McGuinn W., Rothmann M., Timmer W., Malozowski S., Christensen J., Walker A. & Horton R. (2001). The FDA and the Lancet: an exchange. *Lancet* 358, 415-418.

Gilbody S., Wahlbeck K. & Adams C. (2002). Randomized controlled trials in schizophrenia: a critical perspective on the literature. *Acta Psychiatrica Scandinavica* 105, 243-251.

Gotzsche P. (1989). Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Controlled Clinical Trials* 10, 31-56.

Horton R. (2001). Lotronex and the FDA: a fatal erosion of integrity. *Lancet* 357, 1544-1545.

Hotopf M., Churchill R. & Lewis G. (1999). Pragmatic randomised controlled trials in psychiatry. *British Journal of Psychiatry* 175, 217-223.

Huf G., Coutinho E.S.F. & Adams C.E. (2002). TREC III. The protocol and progress of TREC. *Schizophrenia Research* (Abstracts of the 11th Biennial Winter Workshop on Schizophrenia) 53, 187.

ISIS-4 (Fourth International Study of Infarct Survival) Collaborative Group (1995). ISIS-4: a randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. *Lancet* 345, 669-685.

Jadad A., Moore R., Carroll D., Jenkinson C., Reynolds D., Gavaghan D. & McQuay H. (1996). Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Controlled Clinical Trials* 17, 1-12.

Jadad A.R., Cook D.J., Jones A., Klassen T.P., Tugwell P., Moher M. & Moher D. (1998). Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. *Journal of American Medical Association* 280, 278-280.

Joy C.B., Lawrie S. & Adams C.E. (2001). Haloperidol versus placebo for schizophrenia (Cochrane Review). *Cochrane Database of Systematic Reviews 2*, Update Software.

Kennedy E., Song F., Hunter R., Clarke A. & Gilbody S. (2001). Risperidone versus typical antipsychotic medication for schizophrenia (Cohrane Review). *Cochrane Database of Systematic Reviews 2*, Update Software.

Lent V. & Langenbach A. (1996). A retrospective quality analysis of 102 randomized trials in four leading urological journals from 1984-1989. *Urological Research* 24, 119-22.

Marshall M. & Lockwood A. (2000). Assertive community treatment for people with severe mental disorders (Cochrane Review). *Cochrane Database of Systematic Reviews 2*, CD001089, Update Software.

Marshall M., Lockwood A., Adams C., Bradley C., Joy C. & Fenton, M. (2000a). Unpublished rating scales - a major source of bias in randomised controlled trials of treatments for schizophrenia? *British Journal of Psychiatry* 176, 249-252.

Marshall M., Gray A., Lockwood A. & Green R. (2000b). Case management for people with severe mental disorders (Cochrane Review). *Cochrane Database of Systematic Reviews 2*, CD000050, Update Software.

Medical Research Council Streptomycin in Tuberculosis Trials Committee (1948). Streptomycin treatment for pulmonary tuberculosis. *British Medical Journal 2*, 769-779.

Moher D., Pham,B., Jones A., Cook D., Jadad A., Moher M., Tugwell P. & Klassen T. (1998). Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? *Lancet* 352, 609-613.

Moher D., Schulz K.F. & Altman D.G. (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 357, 1191-1194.

Naylor D. (1995). Grey zones of clinical practices: some limits to evidence based medicine. *Lancet* 345, 840-842.

Nicolucci A., Grilli R., Alexanian A., Apolone G., Torri V. & Liberati A. (1989). Quality, evolution, and clinical implications of randomized, controlled trials on the treatment of lung cancer. A lost opportunity for meta-analysis. *Journal of American Medical Association* 262, 2101-2107.

Roland M. & Torgerson D. (1998). What are pragmatic trials? *British Medical Journal* 316, 285.

Schulz K., Chalmers I., Grimes D. & Altman D. (1994). Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *Journal of American Medical Association* 272, 125-128.

Schwartz D. & Lellouch J. (1967). Explanatory and pragmatic attitudes in therapeutical trials. *Journal of Chronic Diseases* 20, 637-648.

Sheldon T.A. (1994). Please bypass the PORT. *British Medical Journal* 309, 142-143.

Silagy C. (1993). Developing a register of randomised controlled trials in primary care. *British Medical Journal* 306, 897-900.

Silagy C., Gray S., Fowler G. & Lancaster T. (1996a). Development of a prospective register of smoking cessation trials. *Controlled Clinical Trials* 17, 235-41.

Silagy C., Gray S., Fowler G. & Lancaster T. (1996b). Development of a prospective register of smoking cessation trials. *Controlled Clinical Trials* 17, 235-41.

Simon G., Wagner E. & Vonkorff M. (1995). Cost-effectiveness comparisons using "real world" randomized trials: the case of new antidepressant drugs. *Journal of Clinical Epidemiology* 48, 363-373.

Simon G.E., VonKorff M., Heiligenstein J.H., Revicki D.A., Grothaus L., Katon W. & Wagner,E.H. (1996). Initial antidepressant choice in primary care. Effectiveness and cost of fluoxetine vs tricyclic antidepressants. *Journal of American Medical Association* 275, 1897-902.

Stern J. & Simes R. (1997). Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *British Medical Journal* 315, 640-645.

Thornley B. & Adams C. (1998a). Content and quality of 2000 controlled trials in schizophrenia over 50 years. *British Medical Journal* 317, 1181-1184.

Thornley B. & Adams C. (1998b). Content and quality of 2000 controlled trials in schizophrenia over 50 years. *British Medical Journal* 317, 1181-1184.

Thornley B., Adams C.E. & Awad G. (2001). Chlorpromazine versus placebo for schizophrenia (Cochrane Review). *Cochrane Database of Systematic Reviews 2*, Update Software.

Update Software Ltd (2002). Abstracts of Cochrane reviews.

Vandekerckhove P., O'Donovan P., Lilford R. & Harada T. (1993). Infertility treatment: from cookery to science. The epidemiology of randomised controlled trials. *British Journal of Obstetrics and Gynaecology* 100, 1005-1036.

Wahlbeck K., Adams C. & Thornley B. (2000). Much to improve: a survey of controlled Nordic schizophrenia trials. *Nordic Journal of Psychiatry* 54, 105-108.

Wells K. (1999). Treatment research at the crossroads: the scientific interface of clinical trials and effectiveness research. *American Journal of Psychiatry* 156, 5-10.

Whicher E., Morrison M. & Douglas-Hall P. (2002). 'As required' medication regimens for seriously mentally ill people in hospital. *Cochrane Database of Systematic Reviews 2*, Update Software.

WHO Scientific Group on Treatment of Psychiatric Disorders (1991). *Evaluation of Methods for the Treatment of Mental Disorders*. WHO: Geneva.