The pragmatic randomised controlled trial

Matthew Hotopf

In recent years there has been much debate regarding the evaluation of treatments in medicine. The evidence-based medicine (EBM) movement has formed partly out of the realisation that clinical practice is often poorly informed by the best available evidence, and that many widely used treatments are either completely untested, or tested and proven to be ineffective or even harmful. EBM has been characterised as a stick by which policy-makers and academics beat clinicians (Williams & Garner, 2002). However, another side to EBM has been the realisation that research performed to test new treatments has often been of poor quality, or has asked the wrong questions (Hotopf *et al*, 1997; Thornley & Adams, 1998; Barbui & Hotopf, 2001). We have previously argued that clinicians could justifiably criticise the research establishment for failing to provide answers to relevant clinical problems of everyday practice (Hotopf *et al*, 1999).

The main tool to answer such problems is the randomised controlled trial (RCT). However, the majority of RCTs in psychiatry have been designed to answer a relatively narrow set of questions, predominantly relating to pharmacological treatments. As clinicians, we do much more than prescribe drugs – for example, we admit people to hospital, discharge them, use differing levels of supervision under the Care Programme Approach (CPA), use assertive outreach teams and refer patients to other professionals. There have been very few RCTs to assess these complex aspects of health care, which may be no less important than the drugs we prescribe.

Two main approaches can be taken in order to address these complex problems in health care provision (as well as some simpler ones about the delivery of pharmacological and psychotherapeutic interventions). One is to attempt to extend RCT methodology to incorporate more complex interventions, another is to abandon RCTs and to rely instead upon observational data. In this article, I will make the case for the former approach, using as an example the problem of the recognition and management of common mental disorders in primary care.

The problem

Most depression and anxiety is treated in primary care. These disorders frequently go undetected, and when they are recognised treatment is often haphazard, with inappropriate use of antidepressants and poorly coordinated psychological services. Given the high prevalence and costs of common mental disorders it is clearly important to improve existing services.

Much of our knowledge on the effective treatment of common mental disorders is based on RCTs, comparing one or more antidepressant or other drug in secondary care. Box 1 describes some of the differences between this body of research and actual clinical practice in primary care. While policies such as the National Service Framework (Department of Health, 1999) suggest that common mental disorders should be managed in primary care, most of the evidence regarding effective treatments comes from secondary care. Typically, patients are highly selected and so may differ from those seen in primary care in many respects, and they go through a range of time-consuming procedures that are quite unlike the normal experience in clinical practice. These trials (which give essential evidence on efficacy) do

Matthew Hotopf is Reader in Psychological Medicine at Guy's, King's and St Thomas' School of Medicine (103 Denmark Hill, London SE5 8AZ, UK) and the South London and Maudsley Trust. His research interests include medically unexplained symptoms, depression and physical illness, and 'Gulf War illness'.

Box 1 Some of the differences between routine clinical practice and traditional randomised controlled trial (RCT) design

Events in a typical RCT	*Events in the real world*
Patients are recruited from specialist centres, or by advertising	Patients are mainly treated in primary care
Patients with comorbid medical or psychiatric disorders are excluded	Patients are probably treated whatever comorbid disorders are present
Patients are carefully selected to generate homogeneous diagnostic groups according to DSM and ICD	Patients with heterogeneous diagnoses according to DSM or ICD are 'lumped' together
Patients are allocated the treatment at random	Treatment is allocated via a complex process of explanation and negotiation
Patients are given detailed information (which may be overinclusive) for informed consent	Patients provided brief information (which may be underinclusive) for informed consent
Patients are given a 1-week placebo run-in period to remove placebo responders	All patients are given active treatment from the start
Placebo is used to compare active treatment	No placebo is used: choice is between active treatment and no treatment
Patients are followed at frequent intervals and given detailed checklists of side-effects	Patients are followed at very varying intervals according to haphazard practice
Assessment end-point is typically 4–6 weeks after treatment begins	Patients continue on treatment for 6 months, and patient and clinician are interested in much longer end-points
Assessment of outcome is based on depressive symptoms and side-effects	To patient and doctor, functional outcomes (e.g. return to work) may be more important
Patient and clinician are blind to treatment group	Both (usually) are aware of the drug the patient is given

---

not tell us whether a treatment is effective in wider clinical practice. Thus, it could be argued that such trials are so far removed from clinical practice that they are, at best, of dubious value. Despite having good internal validity (in other words, a good design that aims to minimise bias) such studies lack external validity to the extent that their usefulness in routine practice is compromised.

Not only is the traditional RCT limited in terms of its generalisability, it may also be limited to answer a relatively narrow set of questions relating to efficacy of well-defined treatments. However, 'real life' questions for those commissioning services are often complex. For example, general practitioners (GPs) might wish to employ additional staff to help manage patients with common mental disorders. Should they employ counsellors, community psychiatric nurses (CPNs), clinical psychologists or psychiatrists? Does intensive follow-up after anti-depressant prescription make a difference? Are there algorithms of interventions for common mental disorders that could be implemented and tested? Does providing in-depth education about the recognition and management of depression improve patient outcomes? Some of these questions have been partially answered by RCTs, but there are still large gaps in our knowledge.

Why randomise?

The choice faced by those designing studies to address 'real world' problems is either to describe outcomes in patients treated in completely natural-istic settings, without using randomisation (observational studies), or to try to adapt the traditional RCT to retain some of its key advantages, but emphasise external validity.

The problems inherent in unrandomised studies are demonstrated below. Brugha *et al* (1992) assessed the outcome of patients with depression who were prescribed antidepressants in psychiatric out-patient departments and compared this with the outcome of patients who had depression but did not receive a prescription. This observational study found that there was no difference between the two

groups. Receiving a prescription for an anti-depressant did not seem to make patients better. The authors suggested that this showed that in real life practice, antidepressants were less effective than would have been expected on the basis of RCTs. However, an alternative view is that the two groups being compared were different. The authors were able to compare, in terms of the severity and duration of symptoms, those who received antidepressants with those who did not, and they found no differences. However, it is unlikely that the doctor's decision to prescribe an antidepressant was completely arbitrary. There are many subtle (and not-so-subtle) factors that determine whether a patient with depression is given a prescription. The prescribers were probably picking up on subtle factors that might have played a major role in determining prognosis. If the group who received an antidepressant had a worse prognosis, this could have disguised the benefits of antidepressant treatment.

Another observational study sought to determine whether patients with depression treated in the USA, who were first seen by psychiatrists, had outcomes different from those seen by primary care physicians (Simon *et al*, 2001). The study was set up in a health maintenance organisation and patients whose depression was first managed by primary care physicians (GPs) were compared with another group who were first treated by psychiatrists. Patients were followed up and outcomes, both in terms of process of care and clinical state, were compared between the two groups. The study found that some aspects of the process of care were very similar – both GPs and psychiatrists often failed to use adequate dosages of antidepressants. Clinical outcomes in terms of remission from symptoms of depression were also similar.

Again, interpreting this study is difficult, for similar reasons. It could be argued that, because there were no differences in outcome between patients seen by their GPs and those seen by psychiatrists, either professional would be as good as the other in managing depression (in effect, this would be to argue that GPs were better, because they would also presumably have been the cheaper option). However, it seems probable that the patients who saw a psychiatrist had more complex or severe problems. The study did, indeed, find that members of this group were more disabled by their depressive symptoms than the patients visiting GPs, even though levels of symptoms were similar. In other words, there may be real advantages to being seen by a psychiatrist but these were hidden because psychiatrists saw a group with more severe illness.

These examples (and many others) demonstrate the problem of non-randomised evaluations of health care interventions – you are never sure whether you are comparing like with like. In each case, randomisation would have overcome these difficulties. The key advantage of randomisation, which makes it such a powerful tool, is that it rules out confounding (including unknown confounders that had not even occurred to the investigators). Provided that enough patients are randomised, the two groups will be similar in most respects. The rationale for randomisation is, therefore, to balance the two groups receiving treatment as perfectly as possible. It is for this reason that the RCT has achieved its status as the gold standard for assessing treatments.

Generalisability and randomisation: getting the best of both worlds

So far, I have suggested that the traditional RCT is sometimes so far from routine clinical practice that its external validity (or generalisability) is called into question, whereas observational studies that simply assess outcomes under service conditions lack internal validity. The pragmatic RCT aims to bridge the gap between these methodologies. Although there are good reasons for many of the features of typical RCTs (Box 1) they are not necessarily synonymous with good trial design if interventions are too narrowly defined, patient selection is too restrictive or outcomes are of limited interest to clinicians and their patients. Pragmatic RCTs often make alterations to such features (Roland & Torgensen, 1998), and some of these adaptations are described below and in the case studies.

Design features of pragmatic RCTs

Participants

A key aim for pragmatic RCTs is to reflect the heterogeneity of patients encountered in clinical practice. They aim to keep exclusion criteria to a minimum – comorbid medical conditions are a common feature of common mental disorders in primary care – and researchers may therefore choose to include medically ill patients who might have been excluded in conventional trials. Suicidal ideas are common in depression, so the aim would be to keep patients in the trial, even if these are expressed.

Another aspect relating to the selection of participants for pragmatic RCTs is that they may not be so preoccupied with narrow diagnostic labels.

Most GPs are uninterested in the arcane distinctions between depressive episodes; mixed anxiety and depression; and adjustment disorder with depressive features. Such disorders are probably treated in a similar way in clinical practice and they frequently overlap. Thus, pragmatic RCTs may 'lump' patients together more than traditional RCTs.

Pragmatic RCTs may also focus on specific clinical groups or presentations that include a wide range of diagnoses. There have been pragmatic RCTs on 'high utilisers' of primary care who are distressed (Katon *et al*, 1992); these patients form a recognisable group for GPs, but do not fit neatly into diagnostic categories. A similar approach might define the population to be studied as patients presenting with early-onset psychosis or deliberate self-harm. The presentation, not the diagnosis, is what defines these groups of patients and it is the presentation that leads to specific challenges for service providers.

Treatment and control groups

Pragmatic RCTs are often concerned with complex interventions. These may include approaches to screening to detect common mental disorders, or referring patients to different health professionals. Whereas traditional trials tend to use rigid dosing regimens for drugs, or very specific therapeutic models for psychotherapy, this tends not to be the case in routine clinical practice, and pragmatic RCTs are therefore more flexible in defining the intervention (Box 2). For example, a trial might assess the use of social workers for patients with depression but allow the nature of the social work intervention to vary according to what the social worker saw as the key issues for an individual client. The social work client might receive non-directive counselling, problem-solving or more pragmatic social care (e.g., help with housing problems).

Box 2 The design of pragmatic RCTs

Pragmatic RCTs:

- reflect the heterogeneity of patients in general practice
- minimise exclusion criteria
- focus on groups with a wide range of diagnoses
- define patient groups by presentation rather than diagnosis
- may not employ placebos
- may not be blinded
- must carefully conceal allocation during randomisation

The classic inactive comparison for traditional pharmacological studies is the placebo. This is not explicitly used in clinical practice and under some circumstances pragmatic RCTs may evaluate drug treatment without using placebos. In other words, if the choice in reality is to have an active treatment, or no treatment at all, this should be reflected in pragmatic RCTs. Similarly, classic psychotherapy trials have attempted to control for the non-specific effects of the therapist's time and attention in order to detect specific treatment effects, by giving some form of 'placebo' psychotherapy. In the pragmatic RCT there is a tendency to deal with each treatment as a black box – usually, the concern is not to try to understand specific active ingredients within the box.

Another approach for control groups is to compare a new treatment with 'usual care' (an approach used in case studies 1 and 2 below). 'Usual care' is a difficult term to define because it will depend heavily on the knowledge, skills and resources of the health care professionals delivering it. If there is evidence that usual care, as experienced by most patients in the health service, is in fact sub-optimal care, the trialist may attempt to provide guidance to participating doctors on what usual care should involve.

Case study 1: Does referral to psychiatry improve outcome in patients with refractory depression? (After Katon *et al*, 1999)

This study aimed to determine whether referral to a psychiatrist working in close collaboration with a GP would improve outcomes for patients who had been started on an antidepressant but had not made a satisfactory recovery.

Patients (aged 18–80) of several health centres were contacted. They had been prescribed antidepressants for depression or anxiety. Those who still had significant symptoms at 6–8 weeks after the prescription were invited to participate and randomised to either 'usual care' or 'stepped collaborative care'. Usual care meant that their GP continued to manage their symptoms. Stepped collaborative care involved a complex intervention delivered by a psychiatrist based in primary care reviewing the patients, giving them information on the treatment of depression and monitoring progress closely. Follow-up was performed by telephone interviewers who were blind to the randomisation status of the patients. The main outcomes assessed included process outcomes, such as compliance with medication, and clinical improvement measured by the Hopkins symptom checklist (Derogatis *et al*, 1971) and health care visits.

The main result was that the intervention was associated with greater adherence to medication, greater satisfaction with care and lower depression scores at 6 months. The intervention did not influence the frequency with which the patient visited the GP.

Case study 2: Which is the most effective treatment of depression: counselling, CBT or usual care by the GP? (After Ward *et al*, 2000)

This study aimed to compare the effectiveness of non-directive counselling with cognitive–behavioural therapy (CBT) or 'usual care' for patients presenting with depression to their GP.

It used broad inclusion criteria – anyone aged over 18 who had a Beck Depression Inventory (Beck *et al*, 1961) score >13 was eligible, provided that they did not have serious suicidal intent, recent exposure to the therapies being compared or an organic brain syndrome. Of those referred to the study, 74% were included.

Patients with a strong preference for one treatment were allowed to opt for that, and the majority chose CBT (81) as opposed to counselling (54) or usual care (2). The investigators changed the protocol during the study because many patients who wanted a psychological treatment, but did not have a strong view about which one they received, opted for the patient-preference arm in order to avoid usual care. A three-way randomisation was made on 197 patients and 130 were randomised in a modified randomisation between CBT and counselling. Patients could not be blinded to the treatments that they received.

Outcomes were based on questionnaire ratings of symptoms (hence avoiding observer bias) at 4 and 12 months. A further outcome, reported elsewhere (Bower *et al*, 2000), was cost effectiveness.

The main result was that at 4 months the two psychological treatment groups were more effective than usual care, but there were no differences between CBT and counselling. By 12 months there were no differences in depression scores between the three groups.

Tackling patient preference

In clinical practice, the choice of treatment is not simply determined by the doctor. It is usually a two-way process in which the doctor gives advice, but the patient is – with some notable exceptions in psychiatry – expected to exercise judgement to arrive at a preference. The ethical basis of RCTs is that there is a state of equipoise regarding which treatment is effective. If one treatment is proven to be more effective than another, there would be no ethical justification in comparing them. Patients, however, may make decisions about treatments on completely different grounds. Many patients with depression prefer counselling to anti-depressants, despite rather sparse data in favour of the former and much more robust evidence in favour of the latter. Some patients would never accept randomisation in the first place because they would not countenance being allocated to a treatment of which they disapproved. Under certain circumstances, this may be a significant proportion of patients.

One way to include data from such patients, and to achieve better external validity, is to perform a patient-preference trial. In this design, all patients are asked to contribute outcome data for the trial, but before allocation they are given a choice to either be randomised or receive a specific treatment. The approach, although attractive, is not without complications. One example (described in more depth in case study 2) is a trial comparing usual care with referral to a counsellor or referral to a clinical psychologist in the management of depression in primary care (Ward *et al*, 2000). Many patients opted for referral and very few opted for usual care. The problem caused by the patient preference arm of the trial was that patients who would have been happy to have been randomised to receive counselling or CBT opted for the patient-preference arm to prevent the risk of being allocated to usual care.

Blinding

Traditional RCTs are typically run double-blind. This means that neither the patient nor the clinician knows to which treatment the patient has been assigned. There are two main reasons for blinding: the first is to prevent the patient's knowledge that he or she has been allocated to a new treatment influence his or her assessment of any clinical improvement. The second reason is to prevent observer bias. This occurs when the researcher's knowledge of the treatment allocated influences (either consciously or unconsciously) his or her appraisal of the patient.

Pragmatic RCTs may not be able to use blinding. For example, when a comparison is being made between receiving treatment for depression from a GP or a psychiatrist, once randomised, it is impossible to blind the patient to his or her allocated group. Thus, the patient cannot be blind to the treatment received. Under such circumstances, it is theoretically possible for the patient's outcome to be assessed by a researcher who remains blind to treatment allocation, but if the researcher has face-to-face contact with the patient, it is likely that he or she will become unblinded (see case study 1 below for an example of when this was possible).

Particular attention, therefore, needs to be paid to two aspects of study design in pragmatic RCTs. First, concealment of allocation during randomisation must be carefully maintained. This refers to how predictable it is that a patient will be allocated to one or other treatment. The amount of attention researchers pay to concealment of allocation strongly

indicates the overall quality of RCTs (Schulz *et al*, 1995). Some randomisation techniques allow the researcher to predict accurately the treatment group for patients. For example, where randomisation takes place in 'random permuted blocks', batches of patients are randomised in blocks of four or six. If a researcher knows what treatment the last patient was allocated, he or she could predict the treatment to which the next patients would be allocated. Having predicted which treatment the patient will receive, the researcher may decide not to randomise the patient if a poor outcome is expected on this treatment. Everything possible must be done to ensure that allocation is concealed and there is no degree of predictability. One approach is to use an independent researcher to perform the randomisation.

The second aspect of the design, which may require special attention in unblinded studies, is the outcome measures used. In other branches of medicine, 'hard' outcomes such as death or cardiovascular events are often used. In psychiatry, outcomes are frequently much more subjective and include symptoms or self-reported quality of life. The problem with such outcomes is that they are particularly open to observer bias. This occurs when the knowledge that a patient has received a specific treatment influences the way the investigator assesses the outcome. For example, when rating symptoms of depression, an assessor who knows that a patient has received treatment in which he or she has considerable faith, may consciously or unconsciously ignore symptoms at follow-up.

The most commonly used outcome scales in traditional RCTs for depression are the Hamilton Rating Scale for Depression (HRSD; Hamilton, 1960) and the Montgomery-Åsberg Depression Rating Scale (MADRS; Montgomery & Åsberg, 1979). These rating scales are particularly subject to observer bias because they are administered after an unstructured clinical interview. Alternative approaches are structured interviews or questionnaires – these reduce the possibility of observer bias because they are presented to each participant in an identical manner.

Outcomes

Outcome measures should be chosen not just to prevent observer bias, but also to reflect the 'real world' concerns of patients, clinicians and policy-makers (Box 3). It is difficult to interpret a three-point reduction in the Hamilton Depression Rating Scale. It is easier to interpret binary outcomes such as recovery or remission, but the definitions of these are frequently arbitrary, such as a 50% reduction in

---

> **Box 3 Outcomes in a pragmatic randomised controlled trials**
>
> Outcome measures should reflect 'real world' concerns:
> - return to work
> - readmission to hospital
> - reduction in visits to the GP
> - cost-effectiveness
> - suicide attempts
> - death from suicide
> - acts of violence
>
> Functional outcomes should be emphasised
> Outcomes must be measured over a sufficient time period

---

symptoms, or a reduction in symptoms below a threshold.

In pragmatic RCTs, more emphasis is placed on functional outcomes that indicate more than symptomatic improvements. Some examples of relevant outcomes are shown in case study 3. Another consideration is the time spans over which outcomes are measured. Many traditional RCTs assess outcome over a 4- to 8-week period. Given the relapsing and remitting nature of many psychiatric disorders, this is an insufficient time period for many interventions.

The unit of randomisation

The traditional RCT almost always randomises individual patients. However, some interventions in health services research are aimed not at patients but at units within the health service (Gilbody & Whitty, 2002). For example, one approach to the problem of poor recognition and management of common mental disorder in primary care has been to educate GPs. If the intervention (education) is to be aimed at GPs or GP practices, these should be the unit of randomisation (see case study 3 below). Outcomes are then measured on clusters of patients seen by each GP or practice.

The advantage of cluster randomisation is that it allows study of the relevant level of health service. In the example of improving GP knowledge about common mental disorders, it might be feasible to randomise patients to see either a GP trained in the detection of common mental disorders or one who was not trained, but this would not reflect the clinical reality. The intervention is not given at the level of the individual patient but at the level of the doctor.

The main problem with cluster randomisation is that it is cumbersome to perform and analyse. This is mainly because patients within each cluster are not statistically independent of one another. For example, if general practices were randomised some would contain predominantly affluent patients and others predominantly impoverished ones. This intra-cluster correlation reduces the statistical power of the study and means that many clusters often have to be randomised in order to achieve a good balance.

Case study 3: Does education for GPs in the management of depression improve process and clinical outcomes? (After Thompson *et al*, 2000)

This study was designed to determine whether giving GPs clinical practice guidelines and education about depression improved their detection of cases or the outcome of depression in patients they were treating.

A cluster randomisation was made on 59 practices. Those assigned to the intervention group were given an education programme and practice guidelines on the management of depression. The control group received the intervention at a later date, after the assessments of outcome had been performed.

The study used two main outcomes. The first was the recognition of depression by the GPs. The GPs were each assessed on 30–40 of their patients. These patients were screened for depression using the Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983). The GPs were blind to the patients' HADS scores, but rated whether they thought that patients had depression. The second outcome was the patients' recovery. Those who scored positive on the HADS were followed up at 6 weeks and 6 months to determine whether they had recovered.

A complex power calculation was carried out in the study that included an adjustment for the effect of clustering within practices.

Of the 21 409 patients screened for depression, 4192 scored above the threshold on the HADS. The main result was that there was no change in the ability of GPs to detect depression following the education programme, nor any improvement in the outcome of depression in patients who were managed by GPs who had received the intervention.

Conclusions

Pragmatic RCTs are becoming a major tool in the evaluation of complex interventions and services. They provide a realistic compromise between observational studies, which have good external validity at the expense of internal validity, and conventional RCTs, which have good internal validity at the expense of external validity. It is hoped that their use will bridge gaps in understanding

between clinicians and researchers and ultimately lead to improvements in services.

References

Barbui, C. & Hotopf, M. (2001) Forty years of antidepressant drug trials. *Acta Psychiatrica Scandinavica*, **104**, 92–95.

Beck, A. T., Ward, C. H., Mendelson, M., *et al* (1961) An inventory for measuring depression. *Archives of General Psychiatry*, **4**, 561–571.

Bower, P., Byford, S., Sibbald, B., *et al* (2000) Randomised controlled trial of non-directive counselling, cognitive–behaviour therapy, and usual general practitioner care for patients with depression. II: Cost effectiveness. *BMJ*, **321**, 1389–1392.

Brugha, T., Bebbington, P., MacCarthy, B., *et al* (1992) Antidepressants may not assist recovery in practice: a naturalistic prospective cohort. *Acta Psychiatrica Scandinavica*, **86**, 5–11.

Department of Health (1999) *National Service Framework for Adult Mental Health.* London: Department of Health.

Derogatis, L. R., Lipman, R. S., Covi, L., *et al* (1971) Neurotic symptom dimensions (SCL–58). *Archives of General Psychiatry*, **24**, 454–464.

Gilbody, S. & Whitty, P. (2002) Improving the delivery and organisation of mental health services: beyond the conventional randomised controlled trial. *British Journal of Psychiatry*, **180**, 13–18.

Hamilton, M. (1960) A rating scale for depression. *Journal of Neurology, Neurosurgery and Psychiatry*, **23**, 56–62.

Hotopf, M., Lewis, G. & Normand, C. (1997) Putting trials on trial: the costs and consequences of small trials in depression: a systematic review of methodology. *Journal of Epidemiology and Community Health*, **51**, 354–358.

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

Katon, W., von Korff, M. & Lin, E. H. (1992) A randomized controlled trial of psychiatric consultation with distressed high-utilizers. *General Hospital Psychiatry*, **14**, 86–98.

—, —, —, *et al* (1999) Stepped collaborative care for primary care patients with persistent symptoms of depression. *Archives of General Psychiatry*, **56**, 1109-1115.

Montgomery, S. A. & Åsberg, M. (1979) A new depression scale designed to be sensitive to change. *British Journal of Psychiatry*, **134**, 382–389.

Roland, M. & Torgensen, S. (1998) Understanding controlled trials: what are pragmatic trials? *BMJ*, **316**, 285.

Schulz, K. F., Chalmers, I., Hayes, R. J., *et al* (1995) Empirical evidence of bias. *JAMA*, **273**, 408–412.

Simon, G. E., von Korff, M., Rutter, C. M., *et al* (2001) Treatment process and outcomes for managed care patients receiving new antidepressant prescriptions from psychiatrists and primary care physicians. *Archives of General Psychiatry*, **58**, 395–401.

Thompson, C., Kinmonth, A. L., Stevens, L., *et al* (2000) Effects of a clinical-practice guideline and practice based education on detection and outcome of depression in primary care: Hampshire Depression Project randomised controlled trial. *Lancet*, **355**, 185–191.

Thornley, B. & Adams, C. (1998) Content and quality of 2000 controlled trials in schizophrenia over 50 years. *BMJ*, **317**, 1181–1184.

Ward, E., King, M., Lloyd, M., *et al* (2000) Randomised controlled trial of non-directive counselling, cognitive–behaviour therapy, and usual general practitioner care for patients with depression I: clinical effectiveness. *BMJ*, **321**, 1393–1399.

Williams, D. D. R. & Garner, J. (2002) The case against 'the evidence': a different perspective on evidence-based medicine. *British Journal of Psychiatry*, **180**, 8–12.

Zigmond, A. S. & Snaith, R. P. (1983) The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, **67**, 361–370.

Multiple choice questions

1. As regards randomisation:
   a it can only be performed at the level of individual patients
   b it guarantees that the study will be free of confounders
   c it automatically leads to poor external validity (or generalisability)
   d concealment of allocation of randomisation refers to whether the outcome of randomisation can be predicted
   e non-randomised studies are likely to be more affected by confounding than randomised studies.

2. In pragmatic RCTs:
   a narrow inclusion criteria are used to select patients
   b steps are taken to ensure that patients are representative of those seen in general practice
   c patients are treated in a setting they are likely to experience in everyday clinical practice
   d patients with the same presenting problem may be treated together even if they have different diagnoses
   e patients are not told that they are taking part in the trial.

3. As regards treatments compared in pragmatic RCTs:
   a pragmatic RCTs are usually limited to pharmacological interventions
   b pragmatic RCTs may assess models for the delivery of services

   c patient-preference arms allow patients to influence the result of randomisation
   d patient preference arms may interfere with recruitment to the trial
   e in the 'usual care' arm the clinician should be encouraged to use ineffective treatments, because this is what usually happens.

4. The outcomes used in pragmatic RCTs:
   a can only be assessed blind
   b are assessed blind, to prevent observer bias
   c might include cost-effectiveness of interventions
   d aim to be short-term outcomes
   e may include 'process outcomes'.

5. Complex treatments (such as psychotherapies):
   a cannot be properly tested in randomised trials
   b are difficult to give blind
   c may have differential drop-outs between groups because of the patients' preferences
   d cannot provide realistic control conditions for such treatments
   e may have limited generalisability of results because of patient selection.

MCQ answers

1		2		3		4		5	
a	F	a	F	a	F	a	F	a	F
b	F	b	F	b	T	b	T	b	T
c	F	c	T	c	F	c	T	c	T
d	T	d	T	d	T	d	F	d	F
e	T	e	F	e	F	e	T	e	T