Editorial

Large simple trials in psychiatry: providing reliable answers to important clinical questions

JOHN R. GEDDES

THE EMERGENCE OF A GOLD STANDARD

The need for reliable evidence for medical interventions has long been recognised and the history of attempts to produce this evidence is thoroughly documented in the James Lind library (http://www.jameslindlibrary.org/). The status of randomised controlled trials as the most reliable method for evaluating the effects of treatments is now well established. Although the basic principles of the randomised trial have a long developmental history, it is often considered that the first modern medical randomised controlled trial came was the UK Medical Research Council trial of streptomycin in the treatment of tuberculosis (Medical Research Council, 1948). In the years following this trial, many randomised trials were conducted on a number of key questions in a number of disease areas including psychiatry and mental health. Many of these were funded by national public bodies, particularly National Institute for Mental Health in the United States and the Medical Research Council in the UK. These trials were often concerned with major questions regarding classes of treatments, rather than specific compounds. Notable among these early trials were RCTs investigating the efficacy of antipsychotic medication (National Institute of Mental Health Psychopharmacology Service Center Collaborative Study Group, 1964) and antidepressant drugs and electroconvulsive therapy (Medical Research Council, 1965). During the next decades until the early 1980s, a number of other trials were conducted investigating key issues about the newly introduced drug treatments in psychiatry, notably long-term treatment with antipsychotics (Leff & Wing, 1971), continuation treatment with antidepressants (Mindham et al., 1972), treatment of acute mania with lithium and chlorpromazine and prevention of relapse with lithium in both bipolar and unipolar disorders (Prien et al., 1973a, b).

From the 1980’s the situation changed considerably and in the area of pharmaceutical treatments, public funding declined and the vast majority of clinical trials of drug treatments in psychiatry are now sponsored and conducted by the pharmaceutical industry. This change had a number of effects. Firstly, the number of trials has increased dramatically. Secondly, trials were increasingly conducted by industry on single patented compounds and are designed to meet the requirements of the regulatory bodies such as the Food and Drug Administration (FDA) in the US and the Medicines and Healthcare Products Regulatory Agency (MHRA) in the UK. Although the criteria for the regulatory bodies is that the drugs should be safe and effective, the design of trials that are conducted for regulatory purposes often involves clinical compromises in order not to delay the introduction of drugs onto the market. The majority of such trials are short-term, highly controlled.
highly monitored typically designed to separate
between new drug and placebo as efficiently as possible. This approach to clinical trial design has been
described as explanatory (Schwartz & Lellouch, 1967) in
which the aim is to determine if the treatment can work. Patients in these trials are characteristically
highly selected, follow up is short and outcomes are essen-
tially sensitive proxies of clinically relevant outcomes.
For example, the prototypical trial of a new antidepres-
sant drug targeted at gaining regulatory approval is
placebo controlled, with or without an active compara-
tor (which is usually a currently marketed reference
antidepressant drug), of six weeks duration, and
excludes patients with psychiatric or physical co-mor-
bidity or suicidal ideation. The primary outcome mea-
sure is the change on a rating scale of depressive symp-
toms (usually the Hamilton depression rating scale) at
the end of six weeks. Loss to follow up is usually sub-
stantial (around 30%) because of the stringent protocol,
and intention-to-treat analysis is accomplished by using
the last observation carried forward.

THE LIMITATIONS OF REGULATORY TRIALS
FOR CLINICAL DECISION-MAKING

Because the design of these trials is based on a negoti-
ation between regulators and industry, these trials often
have substantial limitations for answering clinical ques-
tions in the real world and, hence, limited clinical credi-
bility. Nonetheless, when they are the only randomised
evidence available, they will often be used to answer clin-
cial questions for which they were not designed. For
example, evidence-based clinical practice guidelines and
health technology assessments (which characteristically
seek to provide recommendations about the optimal aver-
age treatment policy for unselected patient in several
clinical subgroups) are often ill served by the results of
these commercial trials. While the limitations of the
available trials are often emphasised in the systematic
reviews that are undertaken to inform the policy state-
ment, the guideline developers understandably wish to
base their recommendations on the best available evi-
dence and so they will often choose to use the existing
randomised evidence rather than no evidence at all. The
results of the available evidence often need to be extrap-
olated - to an unknown degree - when they are applied to
the target clinical question. Two main problems emerge
when policymakers go beyond the available evidence.
First of all, such recommendations may be premature.
The drugs may not be cost-effective and may even turn
out to be less effective - or more harmful - than existing
treatments (Geddes, 2003; Geddes et al., in press). Secondly, the recommendations may effectively and
inappropriately end the condition of equipoise, making
any further trials extremely difficult to conduct. Three
UK examples show some of the problems that can arise
when short-term commercial trials are used to produce
policy:

Atypical antipsychotics

In the UK, the systematic review (http://www.nice.
org.uk/pdf/AssessmentReport-v2-221001.pdf) undertak-
en for the National Institute of Clinical Excellence
appraisal of the new atypical antipsychotics reported:
Evidence for the effectiveness of the new ‘atypical’
antipsychotics compared to older drugs is, in general,
of poor quality, based on short term trials and difficult
to generalise to the whole population of people with
schizophrenia. Evidence for the effectiveness of the
new ‘atypical’ antipsychotics compared to each other
is still limited. Evidence for the cost-effectiveness of
the new ‘atypical’ antipsychotics in the UK compared
to each other and to older drugs is also limited. All
conclusions are therefore based on limited evidence
and should be treated with caution.

The NICE Appraisal Committee considered this report
but essentially ignored these reservations and recom-

dended that atypical drugs should be a first line therapy
option for patients with schizophrenia. These recommen-
dations then became official NHS policy – and made it
very difficult to conduct the trials that had been also iden-
tified as necessary within the same report.

Anticholinesterase inhibitors

The anticholinesterase inhibitors (including rivastig-
mine, donepezil and galantamine) were introduced in the
1990’s to treat Alzheimer’s Disease. As in the case of the
atypicals the UK National Institute of Clinical Excellence
considered the evidence review (http://www.nice.org.uk
/pdf/Alzheimer_hfa.pdf). Although the review found that
the randomised evidence was of reasonable quality and
confirmed improvement in cognitive impairment, the tri-
als were of unknown applicability and the cost-effective-
ness of the drugs was very uncertain. The review also
noted that a large simple trial – AD2000 (commissioned
and funded by the National Health Service) – was in
progress. Rather than defer a conclusion until AD2000
reported, the NICE Appraisals Committee recommended
that anticholinesterase inhibitors should be made avail-
able to patients on the NHS. This recommendation or ded

Epidemiologia e Psichiatria Sociale, 14, 3, 2005

123

Downloaded from https://www.cambridge.org/core. IP address: 54.191.40.80, on 29 May 2017 at 20:58:58, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1121189X00006357
Selective serotonin reuptake inhibitors (SSRIs)

A recent example of important clinical question being ill-served by phase III explanatory randomised controlled trials is in the controversy over the association between SSRIs and suicide. In this example, the main problem arose because the existing trials were not designed to have sufficient power to investigate a potential preventive effect on suicide or even on the prevention of the emergence of serious suicidal ideas. The absolute risk of suicide is so low that it is unlikely that it could ever be used as a primary outcome in a clinical trial (Gunnell & Ashby, 2004). However, that is perhaps not the main issue that limits the usefulness of the existing randomised evidence. In the SSRI trials it is known that regulatory requirement for placebo control means that only patients with relatively mild disorders will be recruited because it is unlikely that a clinician would choose to withhold a treatment of known efficacy from a patient with severe symptoms (Geddes & Cipriani, 2004). Furthermore, the exclusion of patients with suicidal ideation means that any improvement or reduction of suicide risk in this group of patients would not be observed. Therefore, rather than providing information about what might happen to real-world patients with depressive disorders, these trials provide information about what may happen in a subgroup of patients with relatively mild uncomplicated disorders with no existing suicidal ideation. In this group, there is emerging evidence that there may be a small-risk of increased suicidal ideation following initiation of antidepressant therapy (tricyclics as well as SSRIs) although there is no good evidence that this is translated to an increase in rates of suicide (Cipriani et al., 2005; Fergusson et al., 2005; Gunnell et al., 2005; Martinez et al., 2005). In the presence of the limitations of the available evidence, the difficulty facing clinical decision makers – for example those developing the NICE clinical practice guideline on the management of depressive disorders (http://www.nice.org.uk/page.asp x?o=235213) – is how to balance the possible benefits and risks of treatments when making recommendations for patients with different severity of disorder.

DANGERS OF APPLYING REGULATORY STANDARDS OF CONDUCT TO NON-COMMERCIAL TRIALS

The dominance of highly controlled efficacy trials has clearly had a substantial negative effect on the amount of clinically applicable randomised evidence. There are also concerns that the extensive monitoring of trial conduct required under the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) guidelines for Good Clinical Practice (http://www.ich.org/UrlGrpServer.jsr?@_ID=276&@_TEMPLATE=254) makes the cost of clinical trials prohibitive to all but the very largest drug companies. Even for pharmaceutical companies the risk of mounting a very large expensive clinical trial on a promising compound has to be offset against the risk that such a trial will demonstrate that the investigational compound is either of limited or no efficacy, or causes serious adverse effects. Pharmaceutical companies are therefore under pressure to do everything they can to design the trial to show their product in the optimal light and to ensure that any positive results have maximum impact, leading to as rapid uptake of the new drug as possible following market authorisation.

Providing evidence for clinical decision making: the emergence of large simple randomised trials

It has been recognised for at least two decades that the effects of many treatments although clinically worthwhile, are still only moderately sized and that much larger trials are required to provide precise estimates of their risk and benefits. (Yusuf, et al., 1984). In some fields – for example, cardiology – preventing even a few per cent of negative outcomes per year is worthwhile when the negative outcome in question is death or permanent disability. Very large trials – sometimes called megatrials – are required to detect this size of benefit. The ISIS series of trials (ISIS, 1985; 1988; 1992; 1995) and others such as GISSI (Gruppo Italiano per lo Studio della Sopravvivenza nell’Infarto Miocardico, 1990), Heart Protection Study (2002) and others have become almost commonplace in some clinical areas.

Although similarly very large trials would be needed to demonstrate reductions in mortality (including suicide) in patients with psychiatric disorders, the principles that have been developed to make very large trials practicable can be applied generally to trials in psychiatry that need to recruit hundreds or thousands, rather than tens of thou-
Large simple trials in psychiatry: providing reliable answers to important clinical questions

sands of participants. It is probably true to say that trials in mental health need to be much larger than they currently are (Hotopf et al., 1997; Thornley & Adams, 1998).

Simply increasing the scale of trials as they are currently conducted in psychiatry would lead to a very high cost per trial. The key to achieving large numbers and reasonable costs is to making the trials procedures very simple and efficient and allowing widespread recruitment by practising clinicians. Facilitating widespread recruitment has a further benefit: representative samples of patients can be enrolled in the trial which means that the results will be of widespread applicability to future patients.

The key design aspects of large simple trials focus on maximising internal validity and avoiding bias while at the same time making the trial procedures as simple as possible.

The main principles that inform the design of such trials are:

- The benefits of medical treatments on important outcomes are usually only moderately sized, although still may be clinically worthwhile. This means that the trial needs to be very large to be able to detect the effect reliably and with sufficient precision.
- Inclusion criteria are broad and as unrestrictive as possible. Recruitment has been shown to be adversely affected by the use of strict eligibility criteria (Prescott et al., 1999). Unrestrictive entry criteria make it more likely that the required sample size will be achieved of a broad, heterogeneous range of patients can be recruited thereby increasing the general applicability of the trial results (Peto et al., 1993). The key entry criterion is that both the patient and investigator are substantially uncertain which of the trial treatments would be most appropriate. This “uncertainty principle” has been successfully used in other large-scale trials. It is ethical because it effectively excludes patients for whom a specific treatment is known to be most appropriate. It maximizes recruitment by allowing the widest possible eligibility for trial entry. Crucially, the uncertainty principle also makes clinical sense. We believe that many clinicians in effect randomise privately when making treatment choices: a formal trial can convert personal uncertainty into reliable knowledge.
- All trial procedures are radically simplified. One of the key barriers to participation in trials by clinicians is time availability. Removing this barrier by keeping the trial procedures and data collection to an absolute minimum is essential to achieve the widespread participation and, hence, required sample size for realising the key objectives of the trial. (Gunnell et al., 2005; Peto et al., 1993). The trial should have adequate support materials to make sure that patients understand the reasons for the trial and have additional verbal explanation from their physician.

The latter two principles have led to such trials being termed pragmatic (Schwartz & Lellouch, 1967). In contrast to the question of can the treatment work answered by the explanatory trial, pragmatic – or practical – trials ask does the treatment work? (Haynes, 1999)

The concept of the large simple trial is beginning to be of interest in psychiatry however, with the encouraging signs that there is once again an interest in independent clinical trials being run on important clinical questions (Geddes, 2003). Funding is likely to remain an issue. Public funding for clinical trials in the UK appears to be declining (Chalmers et al., 2003). Trial expertise in the academic sector is also limited. To date - because there is more familiarity with the explanatory trial design - many trials continue to use complex procedures that are at least as complex as most industry phase III trials. This has made these trials feasible only in specialist centres with extra staff and there has not been the widespread participation or scalability that is the hallmark of the true large simple trial. There remains reluctance to focus on one simple hard end point, many trials continue to involve a huge number of rating scales that require independent raters. These add enormous complexity to the cost and ease of the trial and makes such trials as costly as industry trials but without the efficiency that often results from the firm management of commercial organisations. Too often in trials run by academic institutions, once the grant is obtained, the incentive for recruitment disappears and trials commonly fail to meet their recruitment targets.

EXAMPLE: BALANCE – A LARGE SIMPLE TRIAL OF MAINTENANCE THERAPY IN BIPOLAR DISORDER

BALANCE was designed as a large, simple trial and is being conducted to compare the efficacy of the combination of lithium and valproate compared with lithium monotherapy and with valproate monotherapy in bipolar disorder. Drawing on experience of conducting large simple trials in other fields of medicine, the design of BALANCE was highly simplified to allow large scale recruitment of heterogeneous patients with bipolar disorder. BALANCE has simple primary outcomes: time to
hospital admission and use of adjunctive medication for emergent mood episodes.

BALANCE is an international trial, currently recruiting patients from more than 50 sites in UK and one in each of Italy and USA. Recruitment is shortly due to start in several sites in France.

Although the trial was generously funded by the Stanley Medical Research Institute and supported by Sanofi-Aventis, the manufacturer of the trial investigational treatments, sufficient funding to achieve the planned sample size was not forthcoming. Nevertheless, BALANCE is already one of the largest independent drug trials undertaken in bipolar disorder and should provide important and clinically relevant information to guide the treatment of future patients. Further details of BALANCE, including the protocol, are available at www.psychiatry.ox.ac.uk/balance.

REFERENCES


Epidemiologia e Psichiatria Sociale, 14, 3, 2005