Viewpoint

Looking at effectiveness: ideas from the couch

Jenny Firth-Cozens

Introduction

Throughout the world, issues surrounding the clinical effectiveness of treatments are high on health service agendas. These include the fact that so little of our health care has been shown by research methods to be effective; the problems of getting research findings into clinical practice; ways to measure outcomes in day to day clinical practice; and how and to what extent the patient should be involved. There is an urgent need to develop new ways to tackle these issues – ones which can be integrated alongside clinical care, conducted without great extra cost, and which require patients’ input into both what they want to be changed and whether change has actually happened.

This paper considers ways that research in clinical psychology and psychotherapy have tackled these questions, not necessarily with a view to extolling the use of these models within other forms of clinical care, but rather to urge that the options for looking at effectiveness of care are kept open and that ways to involve patients’ views are central to this process.

Considering the effects of treatments

Within medicine, the effectiveness of treatments has almost always been evaluated with randomised controlled trials (RCTs), in which a randomly selected group of patients is given the treatment in question and the outcome is compared with that of a randomly selected group of controls. This method has served medicine since the 1940s, and well designed randomised controlled trials remain the only reliable way we have to estimate relatively small effects of treatment. Nevertheless, limitations have been recognised from very early in the use of randomised controlled trials. These include the fact that many procedures predate randomised controlled trials and are already embedded in clinical practice in ways which make unethical the denial of treatment to patients in the control group. Beyond this there are questions on the validity of generalising from the results: a clinician is entitled to ask, “Are research patients really just like patients in my clinic, so can I apply the findings with confidence to my own group?”

Randomised controlled trials are, after all, about populations and not essentially about individual people. In fact, it is apparent that certain groups, such as women and elderly people, for various research reasons, have been systematically excluded from many randomised controlled trials. This may be good science but it means that the clinician faced with an individual patient has reason to question whether practice should be changed in this particular case. Practising clinicians have to face these sources of variance and others every day, and the differences between research and clinic populations may well be one of the reasons why research findings are decidedly slow to be adopted.

Equally important for the individual patient is that, where health care has been evaluated and randomised controlled trials have been used, account may not be taken of many important aspects of the patient’s individuality. Although consideration is given to the various well recognised attributes for outcomes – such as the duration and severity of the disease and differences between clinicians and settings – it is difficult to take account of all relevant characteristics of the patients, such as life events, social support, expectations, or personality. Brewin and Bradley showed that more can be done about this by raising the importance of patients’ preferences and motivations for some treatments when assessing effects on outcomes, and this then was used as an integral part of a trial of abortions, looking at acceptability of treatments. These individual patient characteristics may have significant effects on outcomes. For example, the personality type usually called “negative affectivity” may seriously influence measurements of quality of life.

A very large proportion of our health care is simply not evaluated at all. This may apply to as much as 80% of our procedures and treatments within medicine, and a much larger proportion within most of the other healthcare disciplines. There are many reasons for this, including ethical issues and the enormity of the task of constantly looking back over what has seemed to work. However, the emphasis on the randomised controlled trials methodology itself also plays a part, possibly through reluctance of reviewers to allow other methodologies either at the level of grant applications or publication. If true, this would dramatically limit the amount of evaluation possible and, perhaps worse, act as a rein on developing new ways to consider effectiveness, particularly those which hold the individual patients as central.

The scientist-practitioner

Because of these important issues, and because of the expense and time entailed in conducting
randomised controlled trials, other methods need to be developed which encourage practitioners to take a more active role in the measurement and management of their patients’ outcomes. An essential part of this effort could be to turn an appreciable proportion of healthcare workers into scientist-practitioners, a model used since the 1940s by clinical psychologists both in the United Kingdom and in the United States.2 This model integrates research with practice in ways which allow clinicians to monitor the process and outcomes of their interventions. Its benefits are that it encourages clinicians to be aware of existing research, develop their own—and, it is hoped, to apply it more regularly because they are a part of it. Also, it takes account of the individuality of the therapist; of the context; and, most importantly, of the patient. The research methods used rarely use randomised controlled trials but instead involve methods which fit alongside the ordinary practice of the clinician so that, among other things, funding for such research can be reduced considerably. In fact, it is at the somewhat blurred boundary between research and audit that such methods can best be developed.

Audit and research
Although research in the form of well designed randomised controlled trials remains the best means of showing whether a treatment is effective or not, clinical audit is a method for showing, among other things, whether an effective treatment has been applied in practice. In other words randomised controlled trials and clinical audit can to some extent be seen as two ends of a single process involving dissemination, education, and change management.11 Nevertheless, there are many reasons to see clinical research and clinical audit as two very different activities, despite the argument that both are science.12 Their fundamental purpose varies, as does what is derived from their findings. For example, to show reliability, it is necessary to be able to replicate research findings, whereas audit is about showing beneficial change each time the cycle is completed—that is, it is hoped that the audit results are not the same as they were the first time round.

Perhaps more important than these differences is that the use of randomised controlled trials in most clinical research, and the inferential statistics that support them, go against many of the ideas behind audit. In randomised controlled trial the statistics used in testing hypotheses about differences between the groups come from agriculture in which the need was to show, for example, that a particular type of grain or fertiliser used in one field did or did not fare better than that used in the neighbouring field, and that any such difference was unlikely to be due to chance. Such inferential statistics take no account of the small area of wheat that might have died in one corner of the field, just as the results of a randomised controlled trial may not emphasise any small group of patients who do not respond to treatment or who deteriorate, as its purpose is to look at the benefits to patients within the experimental group overall. This is the very antithesis of audit, where the thrust of concern will be with those patients who fail to meet the successful outcomes expected in the criteria and standards set.

Nevertheless, there are methods in which audit and research have been successfully combined in ways that assess clinical outcomes for example, that described by Richards,13 which audits variation, tracks it to root causes, applies a protocol according to the audit findings, educates the outliers, and then reaudits outcomes, analysing for significant differences between time 1 and time 2 since the protocol was introduced. Likewise, within psychotherapy and clinical psychology, effectiveness audit has been actively pursued alongside research for several years14 and has also been assessed by different research methods designed to keep the patient foremost. No doubt the urgency of this was the impetus to a decade of this being undertaken to show the usefulness of treatments; in the United States, for example, the development of psychotherapy research came from pressure from insurance companies just as much as it did from the challenge presented by Eysenck15 in 1952 when, on the basis of reviewing 24 studies, he concluded that there was no evidence to support the efficacy of psychotherapy, in particular, psychoanalysis. The huge body of research which resulted (see, for example, Bergin and Garfield’s Handbook of Psychotherapy and Behavior Change16) began with the randomised controlled trial model to investigate general effectiveness, and this continues but has gradually moved towards creating new approaches both to recognise the importance of individual differences (of both patient and clinician) and to investigate process questions such as which particular elements of treatment matter most and for which particular patients.17

Because of the recognition within psychology and psychotherapy that traditional research methods only rarely capture the complexity of human experience (something that is also true increasingly within medicine) and because of the expense of clinical trials, much of this research has been carried out with the scientist-practitioner model already described, and, as a result, several methods have evolved.

Asking the patient
There are perhaps fewer objective tests within psychotherapy, psychology, or psychiatry than in physical health care, and so research has largely relied on the opinions of patients or clinicians, or both, about the state of the patient’s wellbeing. However, this is by no means a limitation as the patient’s view is said to be central to the new health system and should perhaps be taken into account in all considerations of effectiveness of treatment.18

Moreover, adding patients’ perceptions of their symptoms can radically alter research findings. For example, in a recent meta-analysis of antidepressant drug treatment, findings indicated that when clinicians rated patients’ outcomes, there was a small but significant
Looking at effectiveness: ideas from the couch

difference between patients taking placebos and those taking antidepressants; however, when patients were asked to rate their own outcomes, there was no longer any difference,\textsuperscript{19} even when assessing only the subsample with endogenous depression.\textsuperscript{20}

There is no doubt that involving patients complicates things – some patients are quite capable of sabotaging the research whereas others can be so keen to help that they bias their opinions in the direction of improvement.\textsuperscript{21} But using methodologies which leave out this important factor just because it is somewhat unreliable is like looking under the street lamp for the key you dropped, just because it is light there – despite the fact that you dropped it elsewhere in the dark!

Within health care as a whole, the individual patient’s views and estimates of change are now beginning to be captured through questionnaires such as the SF-36\textsuperscript{22} and the Nottingham health profile,\textsuperscript{23} but their use is comparatively rare. Within clinical psychology, as well as problem specific measures, the personal questionnaire has been used for the past three decades in varying degrees of complexity both to assess change (for better or for worse) and to help the therapist constantly assess the progress of therapy. It is formed of the patient’s actual statements (for example, “my back pain...” or “my sense of exhaustion...”), agreed with the therapist, about what the problems are that they most want to change, or the goals that they most want to achieve. Each item on the questionnaires is usually followed by an analogue scale for patients to mark their progress or otherwise. Completed regularly throughout treatment, such a questionnaire can provide invaluable information for linking any changes with various interventions within treatment or external events,\textsuperscript{24} thus acting as an antidote to any idea that simple before and after measurement can show the cause of change. This detailed study of the single case has again been a mainstay of clinical psychological research,\textsuperscript{25} 30 after a similar route to the patient trial methods and to “N of 1” randomised controlled trials,\textsuperscript{27} but with patient expectations and assessment as key. The continual assessment involved makes it an ideal way to assess a series of interventions as well as single ones, and the involvement of both clinician and patient in assessing progress, or otherwise, can allow research and clinical practice to be carried out alongside each other.

Another methodology, much more elaborate and individualised than any of the others, is the use of repertory grids.\textsuperscript{28} This requires a detailed interview with the patient, which is generally seen as part of the treatment as well as part of the methodology to assess change. Nevertheless, it has been used, for example, to assess quality of life in groups of patients at different stages of a rehabilitation programme.\textsuperscript{29} Unlike patients who had still not entered the programme, those who had completed it evaluated their past selves in less positive terms than their present and future selves. In other words, the tool seems useful in showing important aspects of change in terms of the aims of rehabilitation, that would be difficult to capture in other ways. Therapists and others could use such a methodology to evaluate change in self image before and after physical treatments such as surgery, neurology, and so on.

**Using outcomes to create change**

Within psychiatry, psychotherapy, and clinical psychology good reliable questionnaire measures exist with which patients can assess symptoms before treatment (often as part of a general assessment to help with treatment plans) and after it to measure change. A traditional group research design offers no information about individual patients, as all the analyses are conducted at the group level. However, a method developed by Jacobson and colleagues,\textsuperscript{30} 31 now quite widely used, visually presents the before and after scores of individual patients as a scatterplot, enabling clinical teams not only to identify patients, diagnoses, types of treatments (or however you wish to interpret the dots) but also to assess the clinical and statistical evidence for change.

In this method changes are evaluated without using a control group, recognising that whatever change exists (whether an improvement or a decline) is not inevitably the result of the intervention, but perhaps of life events outside treatment. The method also recognises that the variation in any group of patients is so considerable that it is the particular context which needs to be considered rather than an unrealistic comparison with another group in a potentially different context.

With this method patients are assessed before and after treatment, with whatever instrument (including personal questionnaires) is most appropriate and defining “after” in whatever way is most useful to show that change has been maintained. These scores are logged in the form of a scattergram (fig 1). If no change occurs, patients would be on the diagonal. If the patients improve, they would be to the right of the diagonal (presuming that higher scores indicate more symptoms), and if they deteriorate they would be on the left. At a glance, it is obvious who has improved or
become worse. Because this method aims to show clinical change within this particular group, we take the mean of time 1 assessment data (measured over a sufficient period to show that this represents the sample well enough) and draw it across horizontally (fig 2), and also draw below the mean one and two standard deviations from that mean. The widened line of the diagonal represents the standard error of the instrument used, so people falling within the wider diagonal are seen as unchanged.

This research method in fact blends seamlessly into audit. In audit terms the criterion of effective change might be that someone has moved at least one standard deviation, or, more stringently, at least two standard deviations. For our standards, we need also to decide what proportion of patients would be expected to move one or two standard deviations, and this might be decided, for example, from research literature. As each patient has a name, a condition, a consultant, an age, particular life events, and so on, we can colour the dots in any way we like to reflect these categories and to see which patient is not doing well enough and why, so that change can be introduced to improve care. This method is described in more detail elsewhere.32

Such a method is right on the border between research and audit. Although a research methodology in its own right, as a form of audit it uses findings from randomised controlled trials and other research, whenever possible, to set its standards. It can be used as part of routine clinical work with minimum effort and so it encourages the application of research findings both directly and indirectly but keeps the individual patient central to the analysis. In this way it helps to maintain the motivation for change among healthcare workers: seeing people who have not moved into the hatched areas makes the need for bringing about improvements much more immediate than do the mean scores and significance tests in research papers, especially if these patients are our own. Educational initiatives can be designed based on clinically effective research findings to assess conditions or groups in which beneficial change is not occurring. This is not an alterna-

tive to traditional methods; it does not establish the effectiveness of a treatment. Like the other methods described, the results cannot, by design, be extrapolated beyond the particular therapist or context. Rather, the method assesses the parts randomised controlled trials cannot reach. Best of all both methods should be used together.

Conclusions
Establishing effectiveness and managing outcomes are both high on healthcare agendas, as is the involvement of patients in the process. However, methods to do this on a large scale and in a way that is integrated with clinical care have not been used very often in physical medicine. Although, as Sheldon15 points out, “no short cuts exist for obtaining reliable information on effectiveness,” by using or adapting the methods I have described, developed by mental health research to overcome some of its own difficulties, we may be able to tackle the agenda in different ways which would hopefully make practitioners more involved in the research questions they need to consider — which may narrow the gap between research and practice. Moreover, we would open up the debate into assessing the use of a wider variety of research methods that will contribute both to the individual clinician’s own interest and development as well as considering wider issues of effectiveness, in particular those which involve the patient more fully.

I acknowledge the helpful comments of two anonymous referees.

Looking at effectiveness: ideas from the couch