The Meaningful Assessment of Therapy Outcomes: 
Incorporating a Qualitative Study Into a Randomized Controlled Trial 
Evaluating the Treatment of Adolescent Depression

Nick Midgley  
University College London and Anna Freud Centre, London,  
United Kingdom

Flavia Ansaldo  
Southwark Targeted Services—CAMHS, London,  
United Kingdom

Mary Target  
University College London

For many years, there have been heated debates about the best way to evaluate the efficacy and effectiveness of psychological therapies. On the one hand, there are those who argue that the randomized controlled trial (RCT) is the only reliable and scientifically credible way to assess psychological interventions. On the other hand, there are those who have argued that psychological therapies cannot be meaningfully assessed using a methodology developed to evaluate the impact of drug treatments, and that the findings of RCTs lack “external validity” and are difficult to translate into routine clinical practice. In this article, we advocate the use of mixed-method research designs for RCTs, combining the rigor of quantitative data about patterns of change with the phenomenological contextualized insights that can be derived from qualitative data. We argue that such an approach is especially important if we wish to understand more fully the impact of therapeutic interventions within complex clinical settings. To illustrate the value of a mixed-method approach, we describe a study currently underway in the United Kingdom, in which a qualitative study (IMPACT-My Experience [IMPACT-ME]) has been “nested” within an RCT (the Improving Mood With Psychoanalytic and Cognitive Behavioral Therapy [IMPACT] study) designed to evaluate the effectiveness of psychological therapies in the treatment of adolescent depression. We argue that such a mixed-methods approach can help us to evaluate the effectiveness of psychological therapies and support the real-world implementation of our findings within increasingly complex and multidisciplinary clinical contexts.

Keywords: adolescent depression, randomized controlled trials (RCTs), qualitative research, mixed-methods design, outcome research

For many years, there have been heated debates about the best way to evaluate the efficacy and effectiveness of psychological therapies. On the one hand, there are those who argue that the randomized controlled trial (RCT) (and meta-analyses of such trials) is the only reliable and scientifically credible way to assess psychological interventions. The RCT has long been considered the “gold standard” approach, placed at the top of the “hierarchy of evidence” and given almost exclusive credence by bodies such as the Cochrane Collaboration and guideline developers such as the National Institute for Health and Clinical Excellence (NICE). On the other hand, there are those who have argued that psychological therapies cannot be meaningfully assessed using a methodology developed to evaluate the impact of drug treatments, and that the findings of RCTs lack “external validity” and are difficult to translate into routine clinical practice. Those who have offered this critique of RCT approaches (McLeod, 2011) have often argued for the greater use of qualitative research methods, which can provide more “contextual knowledge” and allow a greater focus on meaning-making and the perspective of service users.

More recently, there has been a shift toward what some researchers call a “third research paradigm” (Johnson & Onwuegbuzie, 2004), in which qualitative and quantitative methods are combined in the form of mixed-methods research. Although such mixed-methods approaches have a long history, they are only recently beginning to be used systematically in the study of psychological therapies, and there are still a number of conceptual and pragmatic challenges to designing and carrying out such research. In this article, we argue for a mixed-method approach to evaluating psychological therapies, suggesting that such an approach addresses some of the limitations of either a purely quantitative or qualitative design. To illustrate the value of such an approach, we will describe a study currently underway in the United Kingdom,
in which a qualitative study (IMPACT–My Experience [IMPACT-ME]) has been “nested” within an RCT (the Improving Mood With Psychoanalytic and Cognitive Behavioral Therapy [IMPACT] study) designed to evaluate the effectiveness of psychological therapies in the treatment of adolescent depression. After setting out the aims and designs of the original RCT, we will discuss some of the advantages of incorporating a qualitative study within an RCT design, and address some of the pragmatic and conceptual issues that this raised. We will end by arguing more broadly for the advantages of an RCT trial design in which qualitative data based on the client’s own perspective is interwoven with the more traditional quantitative data usually collected in clinical trials.

The IMPACT Study

Identifying effective treatments for depression early in life represents a public health priority in light of the worryingly high level of incidence during teenage years. In the United Kingdom, 1 in 10 young people referred to Child and Adolescent Mental Health Services receives a diagnosis of depression (Harrington, Fudge, Rutter, Pickles, & Hill, 1990), and around 80% of first episodes occur during the teenage period (Angold & Costello, 2001; Ford, Goodman, & Meltzer, 2003). The strong links between adolescent depression and recurrent depressive conditions and suicidal behavior later in life, as well as the subsequent emergence of personality disorders and substance misuse (Rudolph & Klein, 2009), suggest the importance of identifying and improving psychological treatments that are delivered early and that have long-term benefits in reducing the risk of relapse later in life.

Yet despite a considerable investment in studies that have evaluated psychological treatments, there are still major gaps in our understanding of what kind of treatment is most effective for young people, especially in terms of long-term prevention of relapse, and what it is that contributes to a successful (or unsuccessful) outcome. When the NICE guidelines on child depression were published in the United Kingdom in 2005, it was recommended that a range of psychological therapies (including cognitive–behavioral therapy [CBT] and short-term psychodynamic psychotherapy) could be helpful elements within a treatment package, but it was noted that the evidence available was still provisional and that some of the research findings were contradictory or inconclusive. In the “key research recommendations” section at the end of the guideline, it stated:

“An appropriately blinded, randomized controlled trial should be conducted to assess the efficacy (including measures of family and social functioning as well as depression) and the cost effectiveness of individual CBT, systemic family therapy and child psychodynamic psychotherapy compared with each other and treatment as usual in a broadly based sample of children and young people diagnosed with moderate to severe depression (using minimal exclusion criteria). The trial should be powered to examine the effect of treatment in children and young people separately and involve a follow-up of 12 to 18 months (but no less than 6 months)” (NICE, 2005, p. 40).

In the light of these recommendations, the Health Technologies Assessment (a U.K. government-backed funding agency) put out a call for bids to conduct such a study, and in 2007 a joint application between the University of Cambridge, University of Manchester, and University College London successfully won this bid (Goodyer et al., 2011).

The IMPACT study, as it came to be called, is the largest clinical trial of the psychological treatment of adolescent depression to have ever taken place in Europe (Goodyer et al., 2011), with preliminary findings due to be published in 2014. The IMPACT study is a pragmatic, relapse prevention, superiority, RCT comparing the effectiveness of two specialist treatments—CBT and short-term psychoanalytic psychotherapy (STPP)—with clinical care without psychotherapy (Specialist Clinical Care) routinely delivered across a Child and Adolescent Mental Health Service (CAMHS) in the United Kingdom (The recommendation in the NICE guidelines for further investigation of systemic family therapy to also be included in the study was not followed). During the course of the study, almost 500 young people who have been referred to 18 different CAMHS teams across the United Kingdom, and who meet the criteria for moderate to severe depression, are being randomized to one of the three treatment arms. The 5-year study (now in its third year) is designed to address some of the key questions that were left unanswered by previous research: chiefly, to identify the most effective treatment to reduce depressive symptoms among adolescents with moderate to severe depression both in the short term (6 and 12 weeks) and in the medium/long term (36, 52, and 86 weeks), thereby accounting for the long-term effects of different types of treatments in reducing risk of relapse and recurrence.

Advantages of Using an RCT to Evaluate the Effectiveness of Psychological Therapies

There are a number of advantages to addressing the issue of treatment effectiveness by means of an RCT. The core elements of RCTs include randomisation to different treatments, control over treatment fidelity, and comparison of outcomes in the treatment group (Spillane et al., 2010). RCTs are especially important as a way of evaluating the efficacy of therapeutic treatments because they produce results that in most cases can be confidently explained in relation to controlled and carefully analyzed sets of variables. Different outcomes can be attributed with greater confidence to the impact of the different treatments that are being evaluated and compared, while minimizing the bias deriving from extraneous factors. A well-designed RCT study ensures high levels of scientific rigor and validity by collecting data longitudinally from a large representative sample receiving treatments that are manualized and rated with fidelity measures.

The IMPACT study design (see Goodyer et al., 2011, for full details) aims to uphold a high level of empirical strength and scientific power through the careful implementation of randomization and blinding procedures (Blackwood, O’Halloran, & Porter, 2010) and full adherence to the CONSORT guidelines on the design of RCTs (Schulz, Altman, & Moher, 2010). The young people entering the IMPACT trial have an equal and unbiased chance of being randomized to any of the three treatment arms. During the course of the trial, the research assistants remain blind to participants’ treatment allocation, thereby ensuring that the theoretical and professional inclinations of researchers and clinicians are more likely to maintain a more neutral attitude toward outcomes related to each particular case and type of intervention. Furthermore, the extent to which different treatments are more or
less likely to prevent relapse and maintain therapeutic gains has been incorporated in the design by monitoring the patients’ symptomatology up to 18 months after receiving the intervention. This extended follow-up design begins to address the lack of evidence around the long-term effects of different psychological treatments.

Other key strengths of the IMPACT trial are the implementation of treatment manuals and measures of treatment adherence for the different treatment modalities, ensuring greater adherence to a standardized protocol across sites; this further reinforces the internal validity of the study by guaranteeing that the treatments being investigated are actually the treatments being delivered. Yet this is balanced by an emphasis on “external validity,” in so far as the study is a pragmatic one, with treatments delivered by qualified clinicians in real child and adolescent mental health services. Each treatment has at least preliminary evidence for effectiveness. In terms of generalizability of the findings, although many RCT studies have been criticized for excluding participants with multiple diagnosis, the IMPACT study includes young people with comorbid disorders alongside major depression—especially important, as comorbidity is the rule rather than the exception for young people with depression. This ensures that the sample is representative of the complexity and variety of the cases accessing CAMHS, reflecting more closely and accurately the reality of clinical settings. The large size of the sample also makes it possible to address some of the questions about moderators and mediators of effectiveness, such as the level of family support provided to the young people, or the role of the therapeutic alliance in supporting adherence to treatment.

**Limitations in Using RCTs to Evaluate Psychological Therapies**

Despite all of these strengths, there are clear limitations in the design of “gold standard” statistical methods such as RCT studies in so far as they aim to provide understanding of complex interventions and identify mechanisms of change in therapy, including mediators and moderators of treatment outcome (Dattilio, Edwards, & Fishman, 2010; Kazdin, 2009). Researchers and practitioners have identified an “implementation gap” (Britten, 2010) when it comes to translating RCT’s findings effectively into changes in clinical practice (Holllon, 2006). As Jane Noyes, from the Cochrane Collaboration, has recently put it:

“Many aspects of treatment and care cannot be evaluated by randomized trials . . . a Cochrane review may provide clear evidence of the effectiveness of an intervention, but does not include evidence on how people experience the intervention or how it fits with their lifestyle or matches with their preferred choices or expectations. The latter evidence about views, experiences, lifestyles, concordance, attrition and undesired effects is more likely to be qualitative” (Noyes, 2010, p. 526).

Researchers are increasingly calling for an integrated approach that combines the “hard science” of quantifiable outcomes with qualitative data about the meaning of therapeutic interventions and process of change (Hill, Chui, & Baumann, 2013). The “implementation gap” cannot be filled by data illustrating outcomes unless it also sheds light on the different paths that lead to those outcomes. As a stand-alone approach, the RCT design struggles to explain complex interventions and unpick which aspect of the treatment is key to efficacy and isolate its effect from the broader context. As Blackwood et al. (2010) argue:

“The power of the RCT is dependent upon its capacity to use probability theory to approximate the closed system of the experiment where, all other things being equal, there is only one putative causal factor acting upon the intervention group, and this is absent from the control group . . . [However] interventions in clinical arenas not controlled by trial protocols are open systems, in which many factors additional to the intervention itself, including those relating to organization structure, cultural mores, economic capacity . . . will all affect the effectiveness of the intervention . . . So for example, if the trial shows no effect, one has to question whether or not the intervention itself was ineffective, inadequately applied, applied in an inappropriate setting or the comparison was unsuitable. In health care research, what stops us from seeing an ‘effect’ may have more to do with the system than the intervention per se” (p. 517/18 and 513/14).

Similarly, important “nonspecific factors,” which are thought to contribute to a large proportion of variation in treatment outcomes, often remain unexplained when RCT results suggest the equal effects across treatment types (the so-called “dodo bird effect”). These factors are likely to include important characteristics of the patient–therapist relationship, the personality and skill of the therapist, the particular phase of the therapeutic process, and the individual clients’ responsiveness to such processes (Fishman, 2002).

Like almost all RCTs, the IMPACT study relies on a battery of widely used and validated outcome measures, including structured interviews and questionnaires. Although this makes it possible to have confidence in the reliability and validity of the measures, and to compare findings across a range of different studies, it also means that the only outcomes being measured are ones that have been predetermined by mental health professionals and researchers. However, as the study by Morris (2005) demonstrated, many outcomes that were identified by clients themselves using open-ended interviews (qualitative data) might never have been identified using standardized questionnaires, especially the kind of changes that go beyond symptom-relief and touch on the wider (but ultimately just as important) question about quality of life and the meaning that people attribute to their experiences. Although RCTs need to limit the number of variables that are observed and measured, in reality we are unlikely to truly understand what factors promote or hinder recovery without an understanding of the broader social, cultural, and organizational context in which therapy is taking place.

Moreover, in light of the high rates of dropout in adolescent outpatient psychotherapy (Pelkonen, Marttunen, Laippala, & Lönnqvist, 2000), it can be predicted that a considerable number of young people in the IMPACT RCT will drop out of therapy despite the clinicians’ and researchers’ best efforts. Standardized outcome measures are not designed to capture the complex processes that lead to treatment dropout (and failure), suggesting that the IMPACT study on its own could struggle to provide answers to pressing questions around what helps to facilitate (or hinder) the adolescents’ engagement and retention in therapy. Although psychotherapy research is often driven by a professional need to show the effectiveness of treatment, the reality is that we often learn most by understanding treatment failure.
In summary, although it remains vitally important to identify effective interventions through quantitative methods, it is increasingly just as important to implement RCT studies in which the findings can be “translated” into clinical practice more meaningfully. Nevertheless, the CONSORT criteria for assessing the quality of clinical trials still place much greater emphasis on “science” issues than on issues related to clinical relevance or transferability of findings, and as such, the limitations of randomized clinical trials as a way of moving the field forward continue to be felt.

The Advantage of Incorporating Qualitative Data Into an RCT Study

While the Health Technologies Association (HTA) was going ahead with the funding for a high-profile RCT to evaluate the effectiveness of psychological therapies in the treatment of adolescent depression, a subsequent paragraph from the same guidelines was given less attention. In the 2005 report, an “additional research recommendations” section argued: “A qualitative study should be conducted that examines the experiences in the care pathway of children and young people and their families (and perhaps professionals) in order to inform decisions about what the most appropriate care pathway should be” (p. 41).

The NICE guidelines did not imply that such a qualitative study should be incorporated into an RCT design, as qualitative and quantitative studies have traditionally been carried out relatively independently from each other, drawing on different research paradigms and different research skills. Nevertheless, at the time that the IMPACT Study was designed, some thought had been given to including some qualitative data collection, but this idea had been shelved because it was considered too time- and labor-intensive. There was perhaps also a question about the value of such qualitative data, and whether it could add anything meaningful to the findings of an RCT study.

Certainly, when reviewing the existing psychotherapy research literature, there is a striking absence of mixed-method studies when evaluating treatment efficacy or effectiveness, reflecting the fact, perhaps, that qualitative data are still considered to have a low standing in the “hierarchy of evidence.” In a recent review, McLeod (2011) identified only 12 studies of the outcome of psychotherapy (all with adults) in which the client’s perspective was given less attention. In the 2005 report, “a qualitative study should be conducted that examines the experiences in the care pathway of children and young people and their families (and perhaps professionals) in order to inform decisions about what the most appropriate care pathway should be” (p. 41).

The NICE guidelines did not imply that such a qualitative study should be incorporated into an RCT design, as qualitative and quantitative studies have traditionally been carried out relatively independently from each other, drawing on different research paradigms and different research skills. Nevertheless, at the time that the IMPACT Study was designed, some thought had been given to including some qualitative data collection, but this idea had been shelved because it was considered too time- and labor-intensive. There was perhaps also a question about the value of such qualitative data, and whether it could add anything meaningful to the findings of an RCT study.

Certainly, when reviewing the existing psychotherapy research literature, there is a striking absence of mixed-method studies when evaluating treatment efficacy or effectiveness, reflecting the fact, perhaps, that qualitative data are still considered to have a low standing in the “hierarchy of evidence.” In a recent review, McLeod (2011) identified only 12 studies of the outcome of psychotherapy (all with adults) in which the client’s perspective was given less attention. In the 2005 report, “A qualitative study should be conducted that examines the experiences in the care pathway of children and young people and their families (and perhaps professionals) in order to inform decisions about what the most appropriate care pathway should be” (p. 41).

The NICE guidelines did not imply that such a qualitative study should be incorporated into an RCT design, as qualitative and quantitative studies have traditionally been carried out relatively independently from each other, drawing on different research paradigms and different research skills. Nevertheless, at the time that the IMPACT Study was designed, some thought had been given to including some qualitative data collection, but this idea had been shelved because it was considered too time- and labor-intensive. There was perhaps also a question about the value of such qualitative data, and whether it could add anything meaningful to the findings of an RCT study.

Certainly, when reviewing the existing psychotherapy research literature, there is a striking absence of mixed-method studies when evaluating treatment efficacy or effectiveness, reflecting the fact, perhaps, that qualitative data are still considered to have a low standing in the “hierarchy of evidence.” In a recent review, McLeod (2011) identified only 12 studies of the outcome of psychotherapy (all with adults) in which the client’s perspective was given less attention. In the 2005 report, “A qualitative study should be conducted that examines the experiences in the care pathway of children and young people and their families (and perhaps professionals) in order to inform decisions about what the most appropriate care pathway should be” (p. 41).

The NICE guidelines did not imply that such a qualitative study should be incorporated into an RCT design, as qualitative and quantitative studies have traditionally been carried out relatively independently from each other, drawing on different research paradigms and different research skills. Nevertheless, at the time that the IMPACT Study was designed, some thought had been given to including some qualitative data collection, but this idea had been shelved because it was considered too time- and labor-intensive. There was perhaps also a question about the value of such qualitative data, and whether it could add anything meaningful to the findings of an RCT study.

Certainly, when reviewing the existing psychotherapy research literature, there is a striking absence of mixed-method studies when evaluating treatment efficacy or effectiveness, reflecting the fact, perhaps, that qualitative data are still considered to have a low standing in the “hierarchy of evidence.” In a recent review, McLeod (2011) identified only 12 studies of the outcome of psychotherapy (all with adults) in which the client’s perspective was given less attention. In the 2005 report, “A qualitative study should be conducted that examines the experiences in the care pathway of children and young people and their families (and perhaps professionals) in order to inform decisions about what the most appropriate care pathway should be” (p. 41).

The NICE guidelines did not imply that such a qualitative study should be incorporated into an RCT design, as qualitative and quantitative studies have traditionally been carried out relatively independently from each other, drawing on different research paradigms and different research skills. Nevertheless, at the time that the IMPACT Study was designed, some thought had been given to including some qualitative data collection, but this idea had been shelved because it was considered too time- and labor-intensive. There was perhaps also a question about the value of such qualitative data, and whether it could add anything meaningful to the findings of an RCT study.

Certainly, when reviewing the existing psychotherapy research literature, there is a striking absence of mixed-method studies when evaluating treatment efficacy or effectiveness, reflecting the fact, perhaps, that qualitative data are still considered to have a low standing in the “hierarchy of evidence.” In a recent review, McLeod (2011) identified only 12 studies of the outcome of psychotherapy (all with adults) in which the client’s perspective was given less attention. In the 2005 report, “A qualitative study should be conducted that examines the experiences in the care pathway of children and young people and their families (and perhaps professionals) in order to inform decisions about what the most appropriate care pathway should be” (p. 41).

The NICE guidelines did not imply that such a qualitative study should be incorporated into an RCT design, as qualitative and quantitative studies have traditionally been carried out relatively independently from each other, drawing on different research paradigms and different research skills. Nevertheless, at the time that the IMPACT Study was designed, some thought had been given to including some qualitative data collection, but this idea had been shelved because it was considered too time- and labor-intensive. There was perhaps also a question about the value of such qualitative data, and whether it could add anything meaningful to the findings of an RCT study.
depression and that the integration of both approaches offers something that neither a clinical trial nor qualitative data can offer when looked at separately.

IMPACT-ME—A Qualitative Study “Nest ed” Within an RCT

In the light of the issues set out above, we would like to describe how we have established a qualitative longitudinal study—IMPACT-ME—as a “nested” study within the IMPACT RCT, aiming to interweave qualitative data based on the client’s own perspective with the more traditional quantitative data collected in randomized clinical trials in a “double helix” design (Miller & Crabtree, 2008). We will first set out the overall design of the IMPACT-ME study in relation to the main RCT, before discussing some of the methodological issues that have arisen by attempting this kind of mixed-methods evaluation of psychological therapy as a means of reducing relapse among young people suffering from depression.

The primary aim of the IMPACT-ME study is to explore the experience of overcoming depression, in adolescents (and their parents) who have undergone a course of psychological therapy within a CAMHS setting as part of the IMPACT RCT. The IMPACT-ME study is spread across three time points of data collection, involving in-depth interviews with young people and their parents (when applicable) before the start treatment (Time 1), at the end of treatment (Time 2), and 1 year after the end of treatment (Time 3).

In the pretreatment phase of data collection, all young people and parents entering the IMPACT RCT will take part in a qualitative interview (the “Expectation of Therapy Interview,” Midgley et al., 2011a) investigating the way young people and their parents understand the difficulties that brought them to CAMHS and their hopes and expectations about therapy. The interview schedule is an adaptation of Elliott’s Change Interview (Elliott, Slatick, & Urban, 2001) and Werbart’s Private Theories Interview (Werbart & Levander, 2005). It can be used by the interviewer in a flexible way (in keeping with the principles of qualitative interviewing, e.g., Kvale, 1996) but covers: (a) what brought the young person to treatment and how these difficulties have been affecting the lives of the young person and those around them; (b) the interviewee’s understanding of those difficulties (how things came to be like this); (c) hopes for change and ideas about what could lead to meaningful change; (d) and ideas and expectations about therapy itself.

In the posttreatment phase of data collection (Time 2 and 3), all families in one of the regions (London) taking part in the main IMPACT study will be interviewed using the “Experience of Therapy Interview” (Midgley et al., 2011b), which revisits the topics explored in the earlier interview, but adds to them an exploration of the young people and their families’ experiences of therapy and change over time, with a focus in particular on the processes that led to positive or negative treatment outcomes as well as the broader cultural and contextual factors affecting those outcomes, and an exploration of the participant’s experience of being involved in the research study (e.g., the process of randomisation, research assessment meetings, audio-recording etc.). A sample of therapists will also be interviewed at time point 2, subject to the consent of the young people.

Although the primary aim of the study is to explore young people’s experiences of therapy and their own understanding of the process of change, a number of subsidiary questions will also be addressed, both at specific time-points (e.g., young people’s expectations of therapy, based on time point 1) and longitudinally. Specific attention will be paid to certain subgroups, such as those young people who withdrew from, or dropped out of therapy, and those who appeared to benefit from therapy but who subsequently “relapsed” by the 1-year follow-up point. In an RCT study looking at the role of psychological therapy in preventing relapse, and given the high levels of remission among those with depression who are untreated, it will be especially important to have this kind of interview data at the 1-year follow-up stage.

P ragmatic and Scientific Issues in the Design of the IMPACT-ME Study

Despite the obvious advantages of incorporating qualitative data into an RCT study in this way, there were a number of significant obstacles—both pragmatic and scientific—to making such a study possible. Not least of these was the issue of funding—with major funding bodies notoriously reluctant to fund qualitative studies in the field of psychotherapy research, despite the increasing emphasis in medical services more generally on service-user involvement and a recognition of the value of taking into account the user’s perspective and user’s experience. (Indeed, NICE published their first guidelines specifically on service-user experience in adult mental health settings in 2011). We were extremely fortunate, however, that the Monument Trust—part of the Sainsbury Family Charitable Trusts—agreed to provide funds for an additional study, to be conducted alongside (and incorporated within) the main IMPACT Study. One of the co-Principal Investigators of the RCT study (M.T.) is also a PI on the IMPACT Study, while the other (N.M.) is a Senior Research Fellow on the main RCT.

The first step toward setting up the IMPACT-ME study within the broader frame of the IMPACT RCT was to create a “third research community” among the research team, within which quantitative and qualitative researchers could embrace and value an integrated and flexible approach combining different aims, methods, and procedures. This challenge was heightened by the fact that the IMPACT-ME study officially started 2 years after the RCT study was up and running, with just >100 out of the 540 young people already recruited to the main study, and with the plans for data collection and analysis already clearly set out (Goodyer et al., 2011). This meant that the IMPACT-ME qualitative researchers were given the challenging yet stimulating task to integrate themselves as part of a quantitative research “community,” which needed molding and adaptation to embrace new objectives and methodologies that felt both complementary and contrasting.

This first phase of theoretical and philosophical integration, which entails the adoption of a common set of lenses through which shared research aims could be understood and contextualized, required a delicate process of negotiation. The aim of this process was not only to reach a democratic acceptance of both paradigms under one roof, but rather to achieve a shared conviction that a mixed-method approach could best address the multi-layer complexity of our investigation. In practice this also involved deciding on data-collection procedures as well as data analysis and
integration procedures (Hanson, Creswell, Clark, Petska, & Creswell, 2005). After a meeting between the principal investigators of both studies, a presentation about the qualitative study was made to a national meeting of the IMPACT study, after which two of the authors (N.M. and F.A.) visited each of the study sites and offered a half-day training in qualitative research interviewing, which was followed up three months later with a review meeting, once qualitative interviewing had begun. The focus at this stage, perhaps inevitably, was more on the collection of data rather than on how best to integrate at the stage of data analysis. It was agreed that each research assistant would contribute at least to some extent to both quantitative and qualitative data collection. This meant that each research assistant in the main IMPACT team had to learn a new set of skills, which needed to be integrated as part of their existing package of skills and experiences. Regular supervision sessions helped the researchers to adopt a “pragmatic” approach aimed at collecting both quantitative and qualitative data, which fulfilled high standards of scientific rigor while addressing the complexity of individual experiences within ever changing contexts. In practice, research assistants often spoke about needing to shift their focus in moving from the qualitative to quantitative data, as more open-ended and fluid interviewing gave way to a style that was more structured and focused on gathering specific information.

When deciding on data-collection procedure, the key decisions revolve around the order in which the qualitative and quantitative data are collected (concurrently or sequentially) and the priority or emphasis that is given to each type of data (equal or unequal). For example, when setting up a mixed-method psychotherapy evaluation research, the researchers can decide for quantitative data (such as standardized questionnaires), and qualitative data (such as in-depth interviews), to be collected simultaneously or at different points in time, before, during, and after the intervention. As far as the priority of the data is concerned, it has to be decided whether equal status is given to the two sets of data or whether one type of data is used to inform or support the findings of the other.

In line with Creswell and colleagues’ classification system (Creswell et al., 2003; as cited by Hanson et al., 2005), the IMPACT RCT design can be described both as a concurrent triangulation design and a concurrent nested design study. As a concurrent triangulation design study, qualitative and quantitative data are collected simultaneously and priority is given equally to both sets of data, which are analyzed separately and will be integrated at the interpretation stage. The data sets will be “triangulated” to establish the degree of convergence or divergence of the findings with the aim to address the main topics of investigation: namely, what treatments of adolescent depression have the most extended therapeutic benefits over time and how do these treatments achieve what they appear to achieve (or not achieve). However, it could also be argued that the IMPACT-ME study is “nested” within the main RCT in that the embedded qualitative data are collected to help a subset of questions, such as exploring the experiences of stakeholders and carrying out case study interviews with young people, parents, and clinicians involved in the RCT. In this view, IMPACT-ME qualitative data do not hold equal status to the quantitative data, as it is aimed at addressing a subset of questions, which branch off the main research questions.

One immediate concern, when considering how best to nest the qualitative element within the clinical trial, was the issue of “assessment burden” on the young people and their families. It was agreed that adding further assessment time points would be burdensome, so the qualitative data should be collected alongside the quantitative data at the same time points. Three time points were considered of particular importance: pretherapy (to explore the expectations of therapy and young people’s experience of depression), posttherapy (to explore the nature of change and the experience of treatment), and at 1-year follow-up (to explore what factors contribute to relapse or to the maintenance of treatment gains). However the principal investigators of the IMPACT study were concerned that adding a further qualitative interview to the heavy assessment load would be unmanageable for these vulnerable families, and could lead to families withdrawing their participation from the study. It was therefore agreed to pilot the baseline qualitative interview with a small number of families in one of the study sites (London), in order to assess its impact.

When this pilot period came to an end, we were interested to discover that young people, their parents, and the research assistants were all extremely positive about the qualitative interview, which we decided to use right at the start of the baseline assessment. Young people and their families reported that it made them feel as if they were being listened to and that their “story” mattered; research assistants said that it helped them to engage families and get to know them better, before launching into a more structured psychiatric assessment and/or a set of questionnaires. Although the overall length of the baseline assessment slightly increased, the feedback was so positive in terms of the establishment of rapport between researcher and participants that it was decided to build the qualitative interview in to all baseline assessments (rather than a subgroup, as had originally been proposed). This greatly increased the sample size for the qualitative study at the baseline, with implications for how data were to be analyzed (see below).

Some concern was raised at this stage about the way in which the qualitative interviews could impact on the outcome of the interventions—would they have a “therapeutic” effect in their own right? It can also be assumed (and the families themselves confirmed this) that in many cases the interviews were experienced by the families as a place where their thoughts and ideas about the treatment were especially valued and desired by the research staff, as opposed to any feeling that they were just the subject of observation—and that this is likely to have had an impact on their involvement in the study. Similar findings have been found in studies that have looked at the impact of “therapeutic assessments” on treatment alliance (Hilsenroth, Cromer, & Ackerman, 2012). Although this might raise concerns about the way in which the research impacts on findings, it is almost certain that any involvement with data collection as part of an RCT influences a family’s experience of treatment. At least in the IMPACT study we hope to explore this impact, by means of examining the experience of participating in a research study as part of the IMPACT-ME interview, and whether families themselves felt that this contributed to change.

Slightly different concerns were raised about the qualitative interviews at time points 2 and 3 (posttherapy). The research team discussed whether IMPACT-ME data collection at the end of treatment and at the 1-year follow-up could also be embedded in the existing IMPACT follow-up meetings, but concluded that the “Experience of Therapy Interview” required a separate (and op-
In the case of the IMPACT study and IMPACT-ME, we are aiming to analyze the data at a number of different levels, some separately and some in a combined way. For example, the baseline qualitative interviews will be analyzed alone as a way of investigating young people’s experiences of depression and their expectations about therapy, but the qualitative data on expectations can also be coded quantitatively to look at correlations between expectations and outcome. Alongside this, a study examining a subsample of young people (such as those who dropped out of therapy) can draw on both the quantitative data (looking at moderators or mediators of dropout) as well as sampling from the qualitative interviews to carry out an in-depth analysis of a small number of young people who dropped out of treatment. Over the course of time, we would anticipate a series of studies, some of which will report independently on findings from the qualitative or quantitative data, and others which will incorporate both, whether in the form of systematic case studies or studies focused on specific subgroups (e.g., those who responded well to CBT, or those whom therapists had initially thought were not “suitable” to the type of treatment they were randomized to, but subsequently made good use of therapy).

At a practical level, the integration of qualitative and quantitative data—especially in studies that involve collecting large amounts of data, as with the IMPACT/IMPACT-ME study—has long been complicated by the fact that different types of data analytic software are used for quantitative and qualitative analysis. Whereas software packages such as SPSS are widely used for quantitative data analysis, qualitative researchers have made use of software such as Atlas.ti (http://www.atlasti.com/index.html) and NVivo (http://www.qsrinternational.com/default.aspx), which until recently were purely textual. However, recent developments in qualitative data programming means that it is now much easier to integrate data from a package such as NVivo with statistical packages such as SPSS. A wider range of methods has also been developed to allow different “levels” of qualitative analysis. For example, within IMPACT-ME we plan to analyze large data sets using Framework Analysis (Ritchie & Spencer, 1994), which was developed to work both as a “data management tool” to systematically organize large data sets in sizable “chunks” and as an analytical process whereby the researchers maintain a creative and interpretative stance throughout. Framework Analysis (which is compatible with the NVivo software) can either pave the way toward developing a (quantitative) coding system, or can be the basis for a more in-depth exploratory form of qualitative analysis, such as Interpretative Phenomenological Analysis (IPA, Smith, Flowers, & Larkin, 2009), which explores the lived experience of participants in an idiographic way. IPA is especially helpful for studies such as this, focusing as it does on trying to “explore in detail the participant’s view of the topic under investigation . . . an individual’s perception or account of an object or event as opposed to an attempt to produce an objective statement of the object or event itself” (1999, p. 218). The method allows the researcher to...
build up the analysis from the reading of individual cases to the theorizing of themes at a group level, while retaining the focus on personal perceptions. (For further details, see Smith et al., 2009).

The challenges (and opportunities) of a mixed-methods approach in this particular case are also increased by the fact that data are collected longitudinally, across a period of almost 2 years, and that for IMPACT-ME, we will be integrating data from different perspectives (i.e., the young person, the parent, and the therapist). Despite a growing interest among social science researchers in using qualitative longitudinal research designs (Holland, Thomson, & Henderson, 2006), data analysis of longitudinal qualitative data still navigates in uncharted waters. Researchers are faced with the challenge of developing innovative and experimental strategies to integrate cross-sectional analysis, capturing the nature of the sample at a particular data-collection point, with longitudinal analysis following individual trajectories over time. The ultimate aim is to bring together time, change, and process within complex multidimensional data sets (Holland et al., 2006).

Our research team decided to embrace a flexible approach shifting between cross-sectional analyses, focusing on identifying the key theoretical framework and themes around particular research questions (such as the mechanism of change in good outcome adolescent patients) and individual longitudinal analysis holding the individual as a core unit of analysis within their unique psychosocial context. Pragmatic case studies (Dattilio et al., 2010; Fishman, 2002) have been shown to be useful in examining unexpected outcomes as well as for throwing light on the mechanisms of change. In conducting multiperspective case studies research, bringing together detailed narratives from young people, parents, and therapist, there are ethical issues around the privacy and confidentiality of the individual cases, but there are now helpful guidelines on “best practice” in relation to systematic case study designs (e.g., McLeod, 2011), and several journals, including this one, now have special sections or regular features on “Evidence-Based Case Studies.” Studies such as the one by Lunn et al. (2012) suggest that there is an increasing interest in undertaking case studies as part of RCTs.

Despite the considerable challenges involved in the analysis of such a complex multimethod data set, it is within our team ethos to welcome a research endeavor that contains the potential for methodological and analytical development and innovation throughout the entire research process. How successfully we are able to achieve this within the IMPACT-ME study remains to be seen.

Conclusion

In this article, we have outlined the conceptual advantages of a mixed-methods approach, specifically, incorporating a qualitative study within an RCT, to the evaluation of psychotherapy. We have also described the scientific and pragmatic challenges of design and data analysis we have faced to date in the context of an ongoing study of the effectiveness of psychological treatment of adolescent depression currently underway in the United Kingdom (IMPACT ME and IMPACT).

As yet, there have been few attempts to integrate qualitative data and data from RCTs in evaluating a psychological therapy (Lunn et al., 2012). As recently as 2005, Hanson et al. were lamenting that “virtually nothing has been written about mixed methods research designs in applied psychology generally” (p. 224). Although there are numerous and broadly recognized advantages in implementing mixed methodologies as part of mental health research (Hanson et al., 2005), and social science research at large (Creswell, 2003), in this article we have aimed to specifically illustrate the benefits of combining quantitative and qualitative research as part an RCT of psychological and psychiatric interventions.

Such a mixed-methods approach, we have argued, has several advantages. The “triangulation” of the findings combines qualitative outcome data about the effectiveness of treatment with a deeper understanding of the therapeutic process and mechanisms of change that lead to such outcomes. The collection of in-depth interview data alongside the battery of standardized outcome measures will shed light on important questions around the factors facilitating or hindering the young people’s engagement and retention in therapy, including dropout and treatment failure. This way we can go beyond a set of predefined outcomes to include unexpected broader social, cultural, and contextual factors to build a more complex reality-based model of adolescent depression and process of change inside and outside the therapy. Furthermore, qualitative data add a “zoom” on individual differences in the young people, families, and therapists’ beliefs and preferences affecting treatment alliance, retention, and outcome. In this way, combing the scientific rigor of quantitative data about patterns of change with the phenomenological contextualized strength of qualitative data that can help us to understand the meaning of therapeutic interventions will increase the transferability of the findings into improvements in clinical practice, thereby addressing some elements of the implementation gap in psychotherapy research.

Nevertheless, we want to acknowledge that there are real challenges—both conceptual and pragmatic—to nesting qualitative research within an RCT study. In setting up the IMPACT-ME study as part of the IMPACT RCT, we first encountered a conceptual challenge in integrating the theoretical and philosophical framework of a newly formed research team including both quantitative and qualitative researchers. The creation of a collaborative and effective “mixed method” research team required the careful negotiation of a shared system of values to embrace a pragmatic, flexible stance in approaching our epistemological query. The classical positivist paradigm of most quantitative researchers, relying on the “hard science” of statistical and numerical assessments, had to be integrated with the rather different epistemology of qualitative research. Although this has been successfully done at the level of data collection, through a series of meetings and trainings, it remains to be seen how successfully this can be achieved at the level of data analysis. Given that the main RCT and the qualitative study were set up and funded separately, it is likely that the initial stages of data analysis will be independent of each other, and a true integration at the level of data analysis will be at the stage of secondary data analyses and sub-studies. Whether this would have been different if the qualitative component had been established from the start is an important question.

Certainly there is now much greater common ground between qualitative and quantitative researchers, and there are an increasing number of researchers who have been trained in, and are comfortable working with, both types of data. It is now widely accepted that the most fitting philosophical basis to support mixed-method
research is found in pragmatism (Tashakkori & Teddlie, 2003) or critical realism (Blackwood et al., 2010), although some argue that mixed-methods approaches sit most comfortably with a “postparadigm” generation of researchers for whom the “paradigm wars” of previous generations have been replaced by a view that there are multiple paths toward knowledge, and that no one approach can address all the questions we wish to address (Wheel Eton, 2010). The “third research community” (Teddlie & Tashakkori, 2009) argue that the understanding of reality is provisional and ever changing and equal value should be given to both objective and subjective knowledge; different methods, techniques, and procedures, which ought to be flexibly tailored to the purposes of each epistemological query, can lead to a more balanced and complete view of social phenomena, by drawing on the strengths of both approaches and increasing the internal and external validity of findings (Dures, Rumsey, Morris, & Gleseson, 2010).

Alongside these conceptual challenges, the research team has had to sustain a continuous self-reflective and monitoring process to ensure a smooth and flexible transition between more structured form-filling data-collection procedures (quantitative) and an open-ended in-depth style of interviewing in which the aim is less to categorize and more to explore the meaning of human experience. Other aspects that required careful consideration include the implementation of longitudinal qualitative data-collection procedures as part of an RCT whose plans for data collection and analysis had already been set out (Goodyer et al., 2011). This process has entailed pragmatic decisions about sampling and the timing of the different waves of data collection to minimize the “assessment burden” on the participants and creative attempts to manage the integration of large quantitative and qualitative data sets and the analysis of a large set of qualitative longitudinal data. The management and analysis of complex multidimensional data will require a flexible approach bringing together mixed-method statistical analysis, including the process of “quantifying” qualitative data and vice versa, as well as cross-sectional and individual longitudinal analysis of large qualitative data sets.

To conclude, there are ongoing challenges involved with incorporating qualitative data within RCTs focusing on the effectiveness of complex health care interventions. Yet, it is our view that only multimethod research can truly help us to evaluate the effectiveness of psychological therapies, in such a way that such studies can also support the effective implementation of our findings within increasingly complex and multidisciplinary clinical contexts. In psychotherapy research, we suggest, the “gold standard” can no longer be identified as the use of one methodology in isolation, whether that methodology is the RCT or even meta-analysis of RCTs. For modern psychotherapy researchers, the term “gold standard” (if we wish to retain it at all) can only refer to the reflective and critical integration of a range of methods (Dattilio et al., 2010). The conception of RCTs as belonging to the highest tier of the hierarchy of evidence, whilst single-case studies, expert opinion and qualitative investigations are placed in the lowest ranks, is not only out of date but also potentially harmful and misleading; if we wish to address real-world issues such as how best to help depressed young people whose difficulties can potentially have long-term consequences, we need to see more mixed-methods studies in which qualitative data are nested within RCT designs.

References


Received June 16, 2013
Accepted June 18, 2013