Module: Psychological Foundations of Mental Health

Week 5
Psychological therapies: from behaviour modification to behaviour therapy

Topic 4
Evaluating the efficacy of cognitive therapy – Part 1 of 3

Professor Richard Brown
Department of Psychology, King's College London

Lecture transcript

Slide 4
Psychotherapy is a broad term that captures any psychologically-based approach that seeks to help a person change or overcome problems that are having a negative impact on their well-being, both mental and physical. We have looked at behavioural and cognitive psychotherapy as examples in this module. Other approaches include psychoanalysis and humanistic psychotherapies, also informed by their own, if very different, psychological models and methods. Here is a recent, formal definition of psychotherapy from the American Psychological Association. Take a moment to read it, and reflect on some of the essential elements of the definition.

Slide 5
A very important feature to note in this statement is that psychotherapy is based on established psychological principles. How these principles are established is at the core of a debate within psychotherapy that has run from its earliest beginnings. At its heart, from the rationalist traditions of Freud to the empiricist traditions of behavioural and cognitive therapy, the debate centres around the following two points.

First, what evidence do we need to support the underpinning or foundational principles of a particular psychotherapy approach? And second, what evidence do we need to be confident that the therapy is effective for the individuals that receive it? This latter need is driven by both ethical and economic considerations.

Ethically, we should only be using therapeutic methods that we know, rather than simply believe to be useful. Economically, the cost of psychotherapy can be high, whether paid by the individual, their insurers, or the state health care system. Those resources need to be spent well on the most effective or cost effective treatment approach, while acknowledging that a one-size-fits-all approach to therapy is not appropriate.

Today we refer to the need for evidence-based therapy, or evidence-based practise. We are not going to consider the debate between different therapy traditions in this module. Instead, we will work from the stance that such evidence is best derived from the application of robust and reliable empirical methods. In this module, we will focus on such methods and the evidence that we derive from them, specifically in considering cognitive therapy, or CBT.

Transcripts by 3Playmedia

Slide 6

Evidence-based practise in psychotherapy is part of a wider approach within health care. First and foremost, we need evidence that a treatment which is in use or planned for possible use is effective in reducing a particular physical or mental care problem. However, not all evidences equal value. The better the quality of the evidence, the more faith we have in the conclusions that we draw from it. In particular, we need to be confident that the evidence is not only robust, but that it is also, as much as possible, objective and unbiased.

What has emerged is a hierarchy of evidence types or levels of evidence. We see here an example of an evidence hierarchy, with the strongest evidence at the top and the weakest at the bottom. There are many such hierarchies in circulation and none that is universally accepted. We use this one as an example to discuss the basic principles.

At the base of the pyramid, and offering the weakest evidence, are anecdotal reports and practitioners, expert opinion, or testimonials from patients. This is not based on the systematically collected evidence using standard methods, but is subjective and highly prone to bias, from the experts, the therapists, or the clients receiving the treatment. The second level refers to reports based on individual patients, series of patients, or larger groups that are being observed and outcome measured.

Comparisons may be made between those who receive the treatment and those who didn't, or comparisons before or after treatment. These are typically naturalistic studies, collecting data from the available sources, such as clinical records. Such studies are uncontrolled. This means that there is no attempt to control for factors other than the treatments that may be causing a clinical improvement. They are also very subjective and subject to possible bias from multiple sources.

Better are controlled case studies, or case series. These often apply single case experimental designs in an attempt to control for at least some of the factors that may contribute to observable outcomes. Such methods can be useful in the early stages of therapy development as a prelude for a future clinical trial. Cohort studies are another observational approach in which typically large groups of patients are compared, one of which received the therapy and the other of which did not.

The value of such methods can be increased by the careful matching of the characteristics of the patients in the two groups, either overall or on a case by case basis-- so-called case-controlled studies. Objective and unbiased methods of evaluation can be built into such methods. However, cohort studies do not control for biases that might have led to one patient receiving treatment and another not.

This potential source of bias is addressed in the randomised controlled trial, or RCT. This allocates a patient to treatments or non-treatments or to treatment A versus treatment B in a random way. This and several other features of RCTs make them the so-called gold standard for evaluating therapy outcomes. However, no single RCT will provide conclusive evidence for a therapy's effectiveness.

RCTs vary in terms of their quality, just as any other form of evidence, and may differ in the specific methods used in treatment. For this reason, combining the evidence from multiple RCTs, and indeed from all available evidence types, represents the top of the hierarchy, the use of systematic reviews and meta-analyses. We will only look at RCTs and meta-analysis in this topic. We will consider the methods themselves and the application to evaluating cognitive therapies, before considering some of the evidence for the effectiveness of cognitive therapy for depression.

Slide 7

The randomised controlled trial is a general paradigm for treatment evaluation. However, the aims

of the trial can differ from study to study, even if they all apply the same basic methods that we will look at in a moment. Let's look at some of the aims. Clinical trials can be hugely expensive, costing hundreds of thousands to millions of pounds.

Before committing such money, it is essential to demonstrate that the trial is feasible, with the resources that are available. For example, it is essential to know whether it will be possible to recruit all the patients that will be needed within the time frame. There may not be sufficient suitable patients, or potential patients may not agree to take part. It is also necessary to show that the trained staff are available to deliver the treatment and that everything can be done with the funding and within the time available.

A feasibility RCT seeks to assess these uncertainties. A feasibility trial is a mini-RCT, run under the same conditions as the planned full RCT, with clear operational plans and objectives. Crucially, however, it does not seek to assess whether or not the treatment works, only that the planned trial can be delivered.

A pilot RCT is similar and may incorporate some assessment of feasibility. However, its aims are somewhat broader and designed to inform the design and plans for the full RCT. It may include work on developing and refining the treatment and provide information on the likely size of any treatment effect. This latter information, in turn, helps to plan the number of patients that will be needed in the full trial.

Most full RCTs are efficacy trials. Efficacy in this context refers to whether the active treatment shows a significant effect compared to a controlled treatment under the precise conditions of the clinical trial. Once a treatment has been shown to be effective, perhaps over the course of a number of separate RCTs, the question often remains whether it works equally effectively in the real world of clinical care, where there can be multiple factors that can complicate outcome. Such RCTs are called effectiveness RCTs. These are typically the most expensive of all and, for that reason, also one of the least commonly done.

Slide 8
Let's look at some of the key features that, together, distinguish the RCT from other types of treatment evaluation. There are many variations and refinements in the design of RCTs, but all have the following key features.

Slide 9
The typical RCT defines in advance the precise characteristics of patients that will take part. This includes the characteristics they must show, the inclusion criteria, but also any characteristics that rule them out from taking part, the exclusion criteria.

This serves a number of important purposes. First, it is likely to reduce the amount of variability in patient response to treatment. The less variability between patients, the easier it is to show a statistical effect of the treatment. So, for example, a trial of cognitive therapy for depression may define an inclusion criteria of a Beck Depression inventory score of greater than 21. This ensures that patients taking part will have a similar range of depression scores before starting.

Second, it can exclude patients with characteristics that may interfere with the effective or safe delivery of the treatment. For example, patients with poor literacy skills who could not make use of written materials provided as part of treatment or keep written records. So it can seek to minimise the impact of features that may complicate treatment and interpretation of the outcomes. Thus, the trial may exclude patients with significant symptoms of both depression and anxiety. This may be because the anxiety and depression symptoms interact and so influence the efficacy of therapy targeted at depression only.

Despite the advantage of carefully defined criteria, the more restrictive they are, the narrower

the range of patients able to take part and the harder it can be to find patients suitable for the study. Many a randomised trial has failed to recruit the numbers needed, because the eligibility criteria have been too narrow. The alternative is to take more time and spend more money. The second disadvantage is that, while the criteria reduce between patient variability and eliminate some confounding factors, they can also mean that the sample of patients selected are not representative of the broad range of patients that might ultimately benefit in the treatment if it is shown to be effective. In other words, the conclusions from the trial may have limited generalizability to patients at large.

Slide 10

The fundamental aim of the RCT is to determine whether a treatment is effective. The control aspect in RCT is that the active treatment under investigation is compared to an alternative control treatment. The term treatment here does not necessarily refer to an effective treatment. The control may be a dummy or placebo treatment-- in other words, one that is outwardly the same as the active treatment, but lacks the ingredient presumed to make the active treatment effective.

This is easy to achieve in drug trials, where a dummy pill can be made that looks and tastes identical to the active test drug, such that no one can tell which is which. The placebo is essential, because even dummy pills can produce real change, the placebo response. This is not just the patient imagining that they are benefiting. Placebos can produce biological changes in disease processes, as well as psychological change. Depression is particularly susceptible to the placebo response.

A review of trials of antidepressant drugs by Walsh and colleagues in 2002 showed that, on average, while 50% of patients taking active drugs showed clinical benefit, so, too, did 31% of those allocated to placebo. Of course, it may not be the pills themselves, active or dummy. Patients taking part in clinical trials receive lots of attention, with researchers showing considerable interest in their symptoms, maintaining contact over the course of the study. Such attention effects can be potent in the context of an illness such as depression. As we will see shortly, this is a major challenge for trials of psychological interventions.

An alternative approach, and one often used in psychotherapy trials, is to compare the active treatment with what a patient would have received if not in the trial-- so-called standard care or treatment as usual. Unfortunately, in the case of help available for psychological problems, treatment as usual often means no treatment. Another approach is to compare a group of patients who receive treatment immediately and those who are on a waiting list to be treated. This has both practical and ethical advantages over a standard care comparison.

Practically, patients may be more likely to take part in a trial if they think that they will be offered a treatment sooner or later, than if they think they have only a 50% chance of receiving it. Second, it may be considered unethical to deny a treatment to half of the patients, even if it does not yet demonstrated that the treatment is effective. The final control is one used where a new treatment has already been shown to be effective. The question then becomes whether it is more effective than the current best or gold standard treatment.

Alternatively, a new treatment may be as effective but easier, cheaper, or quicker to deliver, or one that patients prefer or one that has less need for highly trained staff to deliver. Given that both treatments are known to be effective, unless the difference between them is very large, such trials typically require a large number of patients and so can be very expensive. Demonstrating that a new treatment is better than no treatment or an existing one is called a superiority trial. Demonstrating that it is no worse than an existing treatment is called a non-inferiority trial.

Slide 11

Along with the existence of two or more carefully selected treatments, the most critical element of an RCT is that the decision about which patient should receive which treatment is determined

randomly. Neither the researcher, the treating therapist, or the patient can decide or influence which treatment they receive, whether active treatment or placebo, immediate or delayed. This randomisation is designed to ensure that the overall characteristics of patients in the two groups are likely to be as similar as possible.

Even within the narrow eligibility criteria, patients will differ in the nature and severity of their symptoms, as well as in terms of their gender and ages. All of these may influence the treatment cycle. Therefore, any such influence needs to be equalised between the two conditions of the trial. Randomisation is the best way to achieve this.

The second reason for randomisation is to reduce the opportunity for investigative bias. Without randomisation, an investigator might inadvertently or deliberately allocate a patient to the active treatment if they thought about patient would benefit, and to the controlled treatment if they thought that patient would do less well. This would obviously produce a major systematic bias in the results and invalidate any conclusions that might be drawn.

The third advantage is that randomisation enables blinding that we will look at it in a moment, the ability to hide which condition a patient has been allocated to. While randomisation is essential, it can also be an obstacle to recruitment. Most patients taking part in a trial understandably want to receive what they hope to be an effective treatment, and to receive it as quickly as possible, rather than a dummy treatment or possibly the less effective one. This is one of the reasons why it's so important to undertake an assessment of the feasibility of recruitment before commencing a full scale trial.

Slide 12

A clinical trial is an experiment, in which no one knows the outcome. However, although unscientific, all of those involved will typically hope that it will show that the new treatment is effective or better than an existing treatment. This hope will typically be shared by the clinicians who develop the treatment, the researchers testing it, and the patients receiving it.

Furthermore, they are likely to not only hope that it is effective, but also believe that it will be. They may have good reason for this-- for example, from previous research. However, until the trial is completed, they will not know for sure. If a clinician, researcher, patient, or statistician knows which treatment a patient is receiving, there are multiple opportunities for bias.

This need not be deliberate and conscious, seeking to distort the results, although this could happen. More likely is the opportunity for unconscious bias. Think back to previous weeks' work, where we considered the influence of expectations and beliefs on our perceptions, memories, attitudes, and behaviour. Such very human cognitive biases are no less likely to be present in the context of a clinical trial.

For example, a researcher assessing a patient on a symptom scored on a one to 10 scale may be unable to decide whether it is a seven or an eight. An expectancy bias, based on knowledge of which treatment condition the patient is in, might nudge the researcher to give a score more indicative of improvement in the patients receiving the active treatment. The researchers may also inadvertently and subtly cue the patient to provide a particular response to a question, so when the patient comes in for an assessment the researcher may say, you're looking really well today; the treatment's obviously doing you lots of good; let's ask you some questions about how you are doing.

Patients taking part in the trial are equally prone to bias. Where an assessment is based on a subjective response of the patient, there is good evidence to suggest that the patient gives answers that they think the doctor or researcher would want them to hear, whether or not they are cued to do so. Patients who know they are receiving an active treatment would also be more likely to look for and attribute any improvements to the treatment, rather than to chance or other

events unrelated to the trial. Equally, if they know they are not receiving treatment, they may tend to attribute negative events to a persistence or worsening of their condition.

Finally, there is the potential for bias at the point of analysing the data. However systematic and objective the statistician, there remains the potential for bias to intrude into decisions made in the choice of analytical approaches and interpretation of the results. For all of these reasons, every effort is made in trials to ensure that investigators, researchers, patients, and statisticians are blind to which treatment a patient is receiving, and that this blindness is maintained until after the trial is complete and the results revealed. In some instances, full blindness is impossible. Ensuring the blindness of the statistician is an absolute minimum.

Slide 13

Finally, we should consider how we measure the treatment effects. All clinical trials require a robust and reliable way to measure the symptoms or condition of interest, the target of the intervention. Most trials will have one or possibly two primary outcome measures, the index used to determine whether the treatment is effective. This measure should be capable of measuring the symptom of the condition with minimal error, its reliability. The more reliable the measure, the easier it is to detect a treatment effect from the trial.

Additionally, trials will often have one or more secondary outcomes to provide additional, useful information about the value of treatment. The primary outcome measure will typically be administered before randomisation at baseline, after randomisation, and, at minimum, at the end of the treatment period of the trial. Some trials will have a number of intermediate assessment points during the treatment period. Comparison of the outcome measure before and after treatment is the main indicator of its efficacy.

However, in some treatments, particularly the psychological therapies, it is hoped that clinical benefit will persist after active treatment is stopped. This can be demonstrated by following up the patients for a period of time. The longer the period of time, the better, but also the more costly, particularly if blindness needs to be maintained during the follow-up period.

Finally, clinical trials need to demonstrate that results are robust and statistically significant. However, statistical significance does not mean that the results are necessarily clinically meaningful. Rather than mean difference between groups before and after treatment, clinical trials often have additional clinical criteria.

One of these is to define what is taken to indicate a treatment response. This may be a response greater than what would be expected by chance alone or normal recovery. This may be expressed as a number of points on the self-report scale, such as the Beck Depression Inventory, and is sometimes called the minimum clinically important change, or MCIC.

The second criterion may be an arbitrary reduction in depression severity relative to baseline. Separate definitions in criteria are used to assess clinical outcomes related to clinical response, to short-term remission of symptoms, and to full recovery. We will look at these in detail later.

Slide 14

Before we come to look at the randomised control trial evidence for cognitive therapy and CBT, it's worth considering some of the challenges posed in designing and conducting robust trials, particularly the ones that minimise the risk of bias. The first point to consider is that psychological therapy is an example of what is called a complex intervention. The contrast here is with an intervention such as a drug, which comprises a simple pill administered following a specific and easy to control dosing regime.

Psychological therapies are inherently complex in terms of what is involved in therapy, combined with the fact that treatment is, at least to a degree, tailored to the individual patients presenting

problems in a situation. The more complex the intervention, the more important it is to standardise how it is delivered to minimise variability between patients receiving treatment, between therapists delivering it and across the many different centres that may be involved. A number of strategies are typically used to achieve this. The extent to which they do so is an important determinant of the quality of the trial.

First, the therapy is typically manualised. This is a detailed description of the treatment and how it is to be delivered in the context of the trial. This is based on best practise and trial evidence. While essential in achieving standardisation, a manual has the potential disadvantage of constraining the therapist in the techniques they might use in routine practise with individual patients. It follows that the therapists involved in the trial should all be trained to a minimum acceptable level to deliver the therapy defined by the manual and retain those skills over the course of the trial.

The degree to which the therapist is following the manual also needs to be checked regularly-- so-called treatment fidelity checks. Therapists can drift away, where a therapist starts following the manual closely but starts to lapse back into prior habits as the trial progresses. Such checks are even more important, where the same therapist is delivering both an active and a control treatment, because of the risk of cross-contamination, the inadvertent introduction of elements of active treatment into the control conditions.

Achieving such standardisation is complex, costly, and time-consuming, but essential for robust evidence. However much the trial seeks to standardise treatments, some things cannot be controlled. Therapists will vary in terms of personal and interpersonal characteristics, such as expressed empathy, warmth, and compassion. Such nonspecific factors can be measured and, while hard to control, may need to be incorporated into the data analysis as additional explanatory factors for outcome.

Perhaps the biggest challenge for psychotherapy research is linked to the need to find a good control treatment and achieve blindness. There is no psychotherapeutic equivalent to a dummy pill. Indeed, the majority of trials have not attempted a dummy treatment. Some use treatments as usual or a waiting list as a control, while others attempt what is called an attention placebo.

It's showing that the patients have a similar amount of contact with the therapist or other researcher without any of the active components of the therapy itself. This can mitigate some of the potential placebo components in the active treatment, but not all. The only way in which a true control condition is possible is if there are two treatments that are identical in all aspects and which seem equally plausible and convincing to the patient receiving them. While possible in theory, this is extremely difficult to achieve in practise.

As we will see in trials of CBT for depression, the most common control is not a placebo, but another active treatment, typically antidepressant medication. The effectiveness of such medication will already have been tested blind against a pill placebo in previous trials. Additionally, such trials ensure that all patients participating are offered a treatment for their depression.

This has ethical and practical advantage for recruitment. Without a true match placebo condition, patients will always be aware whether they have been allocated to the active treatment or to a control. This introduces all of the risk of reporting and expectancy bias that we discussed previously.

Equally, the therapist will be aware which treatment they are delivering. This is one reason why regular assessments of treatment fidelity are necessary, particularly if therapists are delivering an alternative and presume less effective control treatment in which they have no faith. Some degree of blindness can be maintained by ensuring that independent investigators are unaware of treatment allocation when carrying out the assessments. Although care is needed, the patients do not reveal which treatment they have been receiving.

Transcripts by 3Playmedia

Finally, and most straightforwardly, the trial statistician can be kept blind when analysing the data. They will simply have group A and group B and have no idea which was which.

Slide 15
Despite these challenges, the randomised control trial remains the basis of evaluating the clinical evidence for the effectiveness of cognitive therapy and CBT. We will use the term CBT here for convenience to refer to a range of treatment approaches based on the cognitive model that combine both cognitive approaches and, more or less, behavioural ones.

In this section, we are going to look briefly at just a small fraction of the RCT evidence available for the treatment of adult depression. A review published in 2013 identified 115 trials, with more published since. We will look at two trials as examples-- one of the earlier, higher quality trials and one more recent one that looks at the application of CBT in the setting of a low-income economy. We will then turn to the synthesis of the wider body of evidence from recent systematic reviews and meta-analyses.

Slide 16
The first study that we will consider is one reported by Irene Elkin and colleagues in 1989, some 26 years after Beck first published his cognitive model and cognitive therapy for depression. This study was notable for a number of reasons. First, it was the largest study at that time and involved a number of different clinical research centres. Previous trials tended to be carried out as a single centre that had a strong association with the therapy being evaluated, thereby increasing the opportunity for unintentional bias.

Second, it assessed not just one psychotherapy, but two-- cognitive therapy and interpersonal therapy, or IPT. So it included a group of patients who were receiving the standard antidepressant treatment at that time, tricyclic antidepressants, and a group taking a pill placebo. As in all RCTs, participants were randomised to one of the four treatment conditions-- cognitive therapy, IPT, antidepressants, or pill placebo.

Including groups taking medication or medication placebo had two major advantages. First, medication was a standard treatment at that time. Any decision whether to recommend the use of a psychological treatment requires evidence that it is at least as effective as medication. Second, it allowed the use of an effective placebo condition, reducing dependence on a psychological placebo.

250 participants were randomised to one of the four treatment conditions, although 11 withdrew before starting treatment. That treatment was designed to last for 16 weeks, at which point outcome was assessed. In practise, 32% participants dropped out of treatment before that time for a number of reasons.

The main outcome measure reported in this study was remission, in other words that the patients no longer showed clinically significant symptoms of depression at the end of the treatment period, as measured on the primary outcomes. Remission was defined as a score of less than 10 on the Beck Depression Inventory or less than seven on the Hamilton Depression Rating Scale, or HDRS, sometimes called the HAM-D. In contrast to the Beck Depression Inventory, the HDRS is rated by the research clinician, who can be blind to the treatment received.

Slide 17
The published paper reports a large number of different analyses of the data, and we will look at just a couple of the main ones here. We will also only look at the results for the Hamilton Depression Rating Scale. The figure on the left shows the proportion of patients in each group showing recovery at the end of the treatment period. Just under 1/3 receiving placebo had recovered compared to over half in each of the three active treatment conditions. Although looking

impressive, the differences were not statistically significant, as indicated by the P-value, which was greater than 0.05.

A separate analysis was also reported for all patients that started the study, whether or not they completed. This is called an intention-to-treat analysis and is typically the preferred method of analysis in clinical trials. It avoids possible biases that might occur by patients dropping out of the control group and other treatment groups at different rates and for different reasons.

These results showed a lower response rate across all the conditions, as those that drop out of treatment understandably do less well. Overall, however, there was now a significant overall difference between the groups. Further analyses show that this was mainly due to the IPT and drug treatment groups doing better than the placebo.

The recovery rate in the CBT group, however, did not differ significantly from the placebo group. However, there was no evidence that either psychotherapy condition was less effective than the antidepressant medication, the standard treatment at the time. This study, therefore, offered support for IPT as an effective treatment for depression and, to a lesser extent, CBT.

Although we have taken this study as a good example of one of the earlier studies, it still has important flaws, most notably the sample size. The limiting factor here is the sample size needed to test for differences between the active treatments. If the aim is to show that a treatment is no worse than an existing one, a non-inferiority trial, much larger sample sizes are needed than when testing whether a treatment's effects are superior to placebo.

Slide 18

The vast majority of trials of CBT for depression have been carried out in countries with highly developed health care systems, including provision for mental health. These countries offer specialist centres and highly trained mental health practitioners and therapists. While the results of such trials provide evidence for the potential efficacy and effectiveness of CBT in such countries, they do not necessarily generalise to countries where mental health services are rudimentary or non-existent, the majority of the population.

This study, published by Atif Rahman and colleagues in 2008, aimed to test whether CBT could be used to help depressed mothers with infant children living in rural areas of Pakistan. Because of the lack of specialist mental health professionals, the intervention is delivered pragmatically via existing resources-- in this case, primary health care workers who serve the broad needs of the rural population. They were trained to deliver one-to-one CBT for 16 weeks, weekly in the final month of pregnancy and the first month after birth and monthly for the following nine months. The control condition had the same number of visits from their primary care workers, but these had not received training for CBT. This equated to enhanced typical care.

Over 900 mothers were randomised to the two conditions and assessed at baseline, at six months, and at 12 months. The participants were assessed by a psychiatrist to see if they still met criteria for a diagnosis of major depressive disorder. Independently, the study also assessed the quality of the social interactions between the mother and her baby and various indices of infant health, such as immunisation record, growth, episodes of infection, and so on. The purpose of this was to see if improving the mother's mental health would also improve the health and early development of her newborn child.

Slide 19

Let's look at some of the results. At baseline, all of the mothers met criteria for major depressive disorder, an inclusion criterion for the trial. At six months, this had dropped to 23% in the intervention group, compared to 53% of those receiving the enhanced usual care. These positive effects were maintained at 12 months.

In terms of infant health, there was a trend for a reduction in the number of infants with stunted growth, but measuring change over 12 months was probably a challenge. However, other important health indicators included a reduced rate of infant diarrhoea in the CBT-treated group and a greater uptake of infant immunisation and maternal contraception. Very significant changes were also observed in the amount of positive interactions with the infant in the CBT group, with 69% of mothers observed playing with their child during assessments, compared to only 43% in the control condition. Similar figures were also observed in the fathers. These latter outcomes are important, because they show the wider and potentially long-term impact of untreated maternal depression on infant health and social adjustment, and show that a cost-effective CBT intervention targeting these can have effects that go far beyond the mother's acute mental health.