

## Editorial

## Negative results in phase III trials of complex interventions: cause for concern or just good science?†

Mike J. Crawford, Kirsten Barnicot, Sue Patterson and Christian Gold

**Summary**

Not all interventions that show promise in exploratory trials will be supported in phase III studies. But the high failure rate in recent trials of complex mental health interventions is a concern. Proper consideration of trial processes and greater use of adaptive trial designs could ensure better use of available resources.

**Declaration of interest**

None.

**Copyright and usage**

© The Royal College of Psychiatrists 2016.

Mike J. Crawford (pictured) is the Professor of Mental Health Research at the Centre for Mental Health at Imperial College, London. Kirsten Barnicot is a research fellow at the Centre for Mental Health at Imperial College, London. Sue Patterson is principal research fellow at Metro North Mental Health in Queensland. Christian Gold is a professor at The Grieg Academy Music Therapy Research Centre at Uni Research Health in Bergen.

Negative symptoms have a major impact on the quality of life of people with schizophrenia and antipsychotic drugs do little to relieve them. Promising results from exploratory trials of arts therapies raised hopes that they could help people with these symptoms. The negative results of a high-quality phase III trial of one such treatment – body psychotherapy – published in this issue are therefore disappointing.<sup>1</sup> In this study, 275 people were randomised to 20 sessions of body psychotherapy or 20 sessions of Pilates. Attendance was good, but at the end of treatment negative symptoms were similar in both groups. These results echo those of a phase III trial of group art therapy that also failed to find clear evidence of patient benefit.<sup>2</sup>

Arts therapies are not the only promising interventions that have failed to show benefit in phase III trials in mental health in recent years. Negative results have been reported in a succession of trials of complex interventions including training for parents of children with autism, web-based cognitive-behavioural therapy for depression, individual cognitive stimulation therapy for dementia and joint crisis cards and community treatment orders for people with psychosis. Although the principle of equipoise that provides the ethical basis for trials inevitably means that not all those tested will be found to be effective, negative results of so many trials of complex interventions in mental health requires critical consideration. Are negative results more common than in trials of pharmacological and other ‘simple’ interventions? What factors might explain null findings in phase III trials when exploratory studies showed benefit, and what can be done to ensure that as much value as possible is gained from these costly studies?

**Reasons for differences**

As summarised in Table 1, negative results in phase III trials of interventions that showed benefit in exploratory trials may reflect methodological weakness in either or both trials, or reflect

genuine differences in the impact of the interventions when delivered at scale. The validity of trials may be compromised in multiple ways. Because it can be difficult to obtain funding for exploratory studies they are often run with limited budgets by committed ‘champions’ who were involved in developing them. Resource constraints may make it hard to minimise bias through steps such as masking researchers, minimising loss to follow-up and using independent teams to analyse study data. Small sample sizes in exploratory trials increase the chances of differences in baseline characteristics of patients, which increase the probability of confounding. In phase III trials, the measures needed to promote real-world validity may dilute effect: intervention fidelity may be compromised by implementation in diverse contexts by non-experts, practising in accordance with usual care. Thus, it is not surprising that even when interventions do show benefit in phase III trials, effect sizes are generally lower than those found in exploratory studies.<sup>3</sup> Small effect sizes found in exploratory studies may disappear altogether in the ‘noise’ of larger studies.<sup>4</sup>

Participant selection may be influential. People recruited to exploratory and phase III trials may differ in ways that shape engagement and effect. In exploratory studies participants are often recruited opportunistically from sites where investigators work. They may be more inclined to help the investigator/clinician or to trust the intervention being offered. The way that clinicians recruit patients for exploratory and phase III trials may also differ. For example, in a recent trial of community treatment orders, some clinicians outside academic centres were unwilling to take the chance that their most high-risk patients would be randomised to the control group, thereby excluding patients who may have benefited most from the intervention.<sup>5</sup>

Measurement error, related to inconsistent application of research measures, may also dilute or mask the benefits of effective treatments. Phase III trials employ larger numbers of researchers across multiple sites, and it is therefore harder to assure that outcome measures are properly applied. Another potential cause of negative results is the precipitous conduct of phase III studies. Enthusiasm about an intervention together with pressures on academics to generate research income may lead to initiation of studies before a full understanding of active ingredients, mechanisms of action and the characteristics of patients who may benefit has been developed. The importance of such factors, particularly where complex interventions are concerned, is increasingly recognised. Rather than ask ‘what works?’, researchers are being urged to ask, what works for whom in which circumstances.<sup>6</sup>

†See pp. 54–61, this issue.

**Table 1** Reasons for different findings in exploratory and phase III trials

Discrepancy	Possible reason
Differences in study design	Including, but not restricted to, differences in: outcome measures control treatment primary end point
False 'positive' result in the exploratory trial	Positive result produced by chance Positive result reflects bias Positive result reflects imbalance between study arms (confounding) Positive findings resulted from powerful placebo effects resulting from enthusiasm and sometimes pioneering efforts of researchers and therapists
False 'negative' result in the phase III trial	Negative result generated by chance (phase III study was underpowered) Negative result arises from failure to deliver treatment as intended (in terms of treatment content or context) Phase III trial recruited a population less likely to benefit from the intervention Negative result relates to measurement error (imprecise data on patient outcomes)
Intervention has efficacy but is not clinically effective	Typical therapists cannot deliver the quality of treatment over time that was provided by specialist centres during the first phase of developing the intervention Typical patients are unwilling to adhere to treatments to the extent that people do during short-term trials in specialist centres Treatment context (administrative support, out-patient care, transportation etc.) needed to ensure success of the intervention is not available in the wider healthcare system

When exploratory studies show benefit, interventions should be tested for clinical effectiveness, but costly phase III trials should only be embarked upon once a theory of how the intervention generates benefits for patients has been developed.<sup>7</sup>

### Specific to complex interventions in mental health?

Although interventions in other areas of health can also be 'complex', psychosocial interventions in mental health are often relational in nature and rely even more on a clinician's ability to navigate interpersonal relationships and actively engage people in treatment. Do these factors have an impact on the likelihood of finding positive effects in phase III trials in mental health? To explore this we examined the final reports of phase III trials published by National Institute for Health Research (NIHR): Health Technology Assessment (the main funder of phase III clinical trials in the UK) between 2008 and 2014. During this period 71% (35 of 49) of trials of complex interventions returned negative results compared with 58% (21 of 36) for trials of drugs and other 'simple' interventions. Moreover, negative results were returned in a greater proportion of phase III trials of complex interventions in mental health (80%, 12 of 15) than other medical fields (68%, 23 of 34). Although not a statistically significant difference, the failure of so many complex interventions in mental health trials is striking, especially given the promising results of the exploratory trials that often preceded them.

### Maximising returns from investment in trials of complex interventions

Process evaluations can add value to trials whether results are positive or negative.<sup>8</sup> In exploratory trials they can be used to develop understanding of recruitment, uptake of interventions and outcomes that may otherwise not have been considered, supporting design of larger-scale studies.<sup>6</sup> In phase III trials, well-designed process evaluations can be used to explore generation of outcomes, support interpretation and generalisation, and planning for further research. Qualitative data collected from researchers, referrers and study participants during the course of a phase III trial can be especially valuable.<sup>9</sup> For instance,

Thornicroft and colleagues drew on accounts of clinicians and participants to explain the failure of joint crisis cards to reduce compulsory treatment.<sup>10</sup> With these data showing that contrary to the aim of the intervention, sessions to develop crisis cards were often led by clinicians, they concluded that if crisis cards were to work, they needed to be delivered by services that were willing to actively incorporate patient preferences in care planning.

Although historically tight word limits in paper-based journals have precluded detailed reporting of process evaluations with clinical trials, circumstances are changing. The advent of web-based reporting should make it easier for authors to disseminate and stakeholders to access findings of process evaluations alongside the main trial results, improving interpretation of the results.

Efforts to improve the cost-effectiveness of drug trials have led to the development of 'adaptive' trial designs.<sup>11</sup> Such trials allow for pre-specified interim analysis and subsequent changes in design to stop early for success or futility, adjust drug doses or modify inclusion criteria. Such trials are not without their problems, not least the increased potential to generate false positive findings.<sup>12</sup> However, at a time when so many trials of complex interventions are generating negative results we believe there is a case for greater use of adaptive designs in phase III trials of complex interventions. Adaptive trials could reduce unproductive investment in interventions that do not work and increase the likelihood that interventions that do work are tested among those who could benefit most from them.

Meanwhile, when considering negative results in phase III trials of interventions that appeared to demonstrate efficacy it is important to ask whether this is good science (a failure of an efficacious intervention to demonstrate clinical effectiveness), or whether methodological weaknesses in the exploratory trial or insufficient understanding of treatment process and context meant that it was premature to conduct a phase III trial.

Given the promise of exploratory studies and the pressing need to improve outcomes for people with schizophrenia, results of the trial of group body psychotherapy are disappointing. However, the team wisely conducted a parallel process evaluation. It will be interesting to see how these additional data can inform the development of new approaches to help people with psychosis and the design of future studies aimed at treating negative symptoms of schizophrenia.

**Mike J. Crawford**, FRCPsych, **Kirsten Barnicot**, PhD, Centre for Mental Health, Imperial College, London, UK; **Sue Patterson**, PhD, Metro North Mental Health, Queensland, Australia; **Christian Gold**, PhD, GAMUT, The Grieg Academy Music Therapy Research Centre, Unii Research Health, Bergen, Norway

**Correspondence:** Mike J. Crawford, Centre for Mental Health Imperial College London Du Cane Road, London W12 0NN, UK. Email: m.crawford@imperial.ac.uk

First received 8 Dec, accepted 7 Mar 2016

## References

- Priebe S, Savill M, Wykes T, Bentall RP, Reininghaus U, Lauber C, et al. Effectiveness of group body psychotherapy for negative symptoms of schizophrenia: multicentre randomised controlled trial. *Br J Psychiatry* 2016; **209**: 54–61.
- Crawford MJ, Killaspy H, Barnes TRE, Barrett B, Byford S, Clayton K, et al. Group art therapy as an adjunctive treatment for people with schizophrenia: multicentre pragmatic randomised trial. *BMJ* 2012 **344**: e846.
- Tyrer P, Cooper S, Salkovskis P, Tyrer H, Crawford M, Byford S, et al. Clinical and cost-effectiveness of cognitive behaviour therapy for health anxiety in medical patients: a multicentre randomised controlled trial. *Lancet* 2014; **383**: 219–25.
- Gold C. Signal and noise in music therapy outcome studies. *Nord J Music Ther* 2014; **23**: 97–8.
- Geller J. Community treatment orders for patients with psychosis. *Lancet* 2013; **382**: 502.
- Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ* 2008; **337**: a1655.
- Campbell M, Fitzpatrick R, Haines A, Kinmonth AL, Sandercock P, Spiegelhalter D, et al. Framework for design and evaluation of complex interventions to improve health. *BMJ* 2000; **321**: 694–6.
- Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ* 2015; **350**: h1258.
- Crawford MJ, Weaver T, Rutter D, Sensky T, Tyrer P. Evaluating new treatments in psychiatry: the potential value of combining qualitative and quantitative research methods. *Int Rev Psychiatry* 2002; **14**: 6–11.
- Thornicroft G, Farrelly S, Szmukler G, Birchwood M, Waheed W, Flach C, et al. Clinical outcomes of joint crisis plans to reduce compulsory treatment for people with psychosis: a randomised controlled trial. *Lancet* 2013; **381**: 1634–41.
- US Department of Health and Human Services. *Guidance for Industry: Adaptive Design Clinical Trials for Drugs and Biologics*. Department of Health and Human Services, 2010.
- Berry DA. Adaptive clinical trials: the promise and the caution. *J Clin Oncol* 2011; **29**: 606–9.

poems  
by  
doctors

## In the Giraffe House

Pascale Petit

Visiting you in the hospital  
is like going into the giraffe house,  
to peer down into that deep pit  
where they overwinter.  
Your head sways towards me,  
a map of *terra incognita*.  
Your legs wade as if through the sea –  
my clown-on-stilts, sleepwalker  
in desert pyjamas, your eyes too soft,  
your mouth so slack the upper jaw  
moves away from the lower  
like you've taken out dentures  
but have to chew over the same word.  
If only you could remember who this visitor is  
high up in the viewing gallery.  
I want to commemorate your youth  
in the savannah, my giraffe mother.  
I'm only passing through to shelter  
from the cold. It's freezing outside  
and I wanted warmth  
but you are all the colours of drought,  
the cracked riverbeds of your skin  
a jigsaw no one can get right.  
I rest my palm against the partition  
and my breath blurs your lips, the long  
blue tongue that keeps licking the glass.

From *The Hippocrates Prize 2015: The Winning and Commended Poems*, selected by T Dalrymple, R Gross, F Oyeboode and S Rae, eds MW Hulse & DRJ Singer. The Hippocrates Press, 2015. © Pascale Petit. Reprinted with permission.

The British Journal of Psychiatry (2016)  
209, 8. doi: 10.1192/bjp.bp.116.181628