

Sample size in clinical trials on anorexia nervosa: a rejoinder to Jenkins

Timo Brockmeyer¹, Hans-Christoph Friederich², Beate Wild² and Ulrike Schmidt³

Invited Letter Rejoinder

Cite this article: Brockmeyer T, Friederich H-C, Wild B, Schmidt U (2019). Sample size in clinical trials on anorexia nervosa: a rejoinder to Jenkins. *Psychological Medicine* **49**, 1581–1582. <https://doi.org/10.1017/S0033291719000679>

First published online: 12 April 2019

Author for correspondence:

Timo Brockmeyer,
E-mail: timo.brockmeyer@uni-goettingen.de

¹Department of Clinical Psychology and Psychotherapy, Institute of Psychology, University of Goettingen, Goettingen, Germany; ²Department of General Internal Medicine and Psychosomatics, Center for Psychosocial Medicine, Heidelberg University Hospital, Heidelberg, Germany and ³Section of Eating Disorders, Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK

We thank Dr Jenkins for his comment on our review on advances in the treatment of anorexia nervosa (AN) (Brockmeyer *et al.*, 2018). We agree with Dr Jenkins' main statement, i.e. adequate sample size calculations are a necessary prerequisite in designing treatment studies, and that having limited statistical power can lead to erroneous conclusions, including so-called false negatives (i.e. failures to statistically identify real differences between the efficacy of treatments). Given the only moderate treatment response, particularly in adult patients with AN, and its severity, high mortality rate, long-term impairments in social functioning and employment, low quality of life, high burden on caregivers, and huge societal costs (Giel *et al.*, 2016; Schmidt *et al.*, 2016; Zipfel *et al.*, 2016), we entirely agree that adequate clinical trials with large enough samples of patients with AN are urgently needed. Hence, we do not disagree with Dr Jenkins regarding the main message of his comment – which is also the basic principle of sample size calculation. However, we would like to comment on a few of his statements, which, in our view, might lead to erroneous conclusions themselves.

The dearth of large-scale randomised controlled trials in AN arises not only from underestimations of statistical power but can be explained by limited funding for AN research (Schmidt *et al.*, 2016), low prevalence rates of AN, and high treatment ambivalence in this population (Abbate-Daga *et al.*, 2013; Williams and Reid, 2010; Gregertsen *et al.*, 2017). For instance, it took 4 years and 10 participating centres to recruit $n = 242$ eligible patients with AN for the ANTOP study (Zipfel *et al.*, 2014). These factors should be taken into account when judging small sample sizes in clinical trials on AN.

Furthermore, Dr Jenkins states that null findings in superiority trials on AN are often interpreted in a way to suggest that the examined treatments are equivalent. Indeed, such an interpretation of a null finding in a superiority trial would be improper. However, Dr Jenkins comment lacks any reference for such interpretations in the AN literature. In our review we do *not* interpret findings in this way. In contrast, we clearly state (as cited by Dr Jenkins in his comment) that 'there is no single psychotherapy that is substantially superior to another'. This is also the common tone in other reviews on treatments for AN (Hay, 2013; Kass *et al.*, 2013; Le Grange, 2016). Likewise, in the original studies (Zipfel *et al.*, 2014; Schmidt *et al.*, 2015) it was clearly communicated that there was no significant difference between the treatment conditions. No conclusions have been drawn that any treatment is not inferior or equivalent to another. Thus, Dr Jenkins presses charges where no crime has been committed.

Dr Jenkins further argues that differences between treatments for AN rarely exceed effect sizes around $d = 0.30$, without providing any proper reference for this specific number [actually, the sample size calculation for the ANTOP study was, for instance, based on an effect size of $d = 0.59$ which was deduced from a previous trial on AN (Dare *et al.*, 2001)]. He then states that, given conventions of power analysis ($\alpha = 0.05$; power = 80%), one would need a sample size of $n = 139$ in each treatment arm to detect an effect of this size. Unfortunately, we cannot reconstruct how this specific sample size results from the given parameters. The needed sample size very much depends on the statistical test that is applied (e.g. for an independent samples *t* test, a sample size of $n = 176$ per condition would be necessary, for a mixed ANOVA it could be $n = 45$ per condition, given the parameters suggested by Dr Jenkins). Dr Jenkins' line of argumentation suggests that previous clinical trials on AN have not utilised appropriate a priori sample size calculations, but this is definitely not the case. For instance, in the ANTOP study (Zipfel *et al.*, 2014) it was expected that one of the two specific treatments (focal psychodynamic therapy and/or enhanced cognitive behaviour therapy) would result in an improvement in body mass index (BMI) of 1.0 kg/m^2 compared with optimised treatment as usual, which was considered a clinically meaningful difference that translates into a between-groups effect size of $d = 0.59$. Given this expected effect size, an alpha of 0.025 (corrected for multiple comparison), and 80% power, one would need $n = 55$ per condition. Expecting an attrition rate of 30%, this sample size was increased to $n = 80$ per condition. In our view, this is a reasonable rationale for a clinical trial on AN. In addition, Dr Jenkins argumentation that previous treatment studies on AN have been insufficiently powered to

detect effect sizes of $d = 0.30$ seems to neglect the issue of clinical significance (Jacobson and Truax, 1991; Bauer et al., 2004). Taking into account the standard deviation in BMI at end of treatment in the ANTOP study, for instance, such an effect size of $d = 0.30$ would translate into a mean difference of 0.513 BMI points, equalling 1.43 kg (given the mean height of the sample in this study). Thus, studies that are sufficiently powered to detect an effect size of $d = 0.30$ as suggested by Dr Jenkins, would render two treatments significantly different if they result in a mean difference in body weight of 1.43 kg. It can be questioned whether such a difference should be considered clinically meaningful.

In sum, we would like to emphasise once again that we agree with Dr Jenkins' point about the need for large enough sample sizes in AN research. However, this valuable discussion should neither discount the obstacles AN researchers have to face when planning a clinical trial (including low funding, low prevalence, high treatment ambivalence in patients) nor the efforts researchers in the field have undertaken to design and conduct methodologically rigorous randomised controlled trials in the past. Finally, discussions around sample size in psychotherapy research should generally not only take statistical but also clinical significance into account.

Author ORCIDs.  Timo Brockmeyer, 0000-0003-2544-7610.

References

- Abbate-Daga G, Amianto F, Delsedime N, De-Bacco C and Fassino S** (2013) Resistance to treatment and change in anorexia nervosa corrected. A clinical overview. *BMC Psychiatry* **13**, 294.
- Bauer S, Lambert MJ and Nielsen SL** (2004) Clinical significance methods. a comparison of statistical techniques. *Journal of Personality Assessment* **82**, 60–70.
- Brockmeyer T, Friederich H-C and Schmidt U** (2018) Advances in the treatment of anorexia nervosa. A review of established and emerging interventions. *Psychological Medicine* **48**, 1228–1256.
- Dare C, Eisler I, Russell G, Treasure J and Dodge L** (2001) Psychological therapies for adults with anorexia nervosa – Randomised controlled trial of out-patient treatments. *British Journal of Psychiatry* **178**, 216–221.
- Giel K, Schmidt U, Fernandez-Aranda F and Zipfel S** (2016) The neglect of eating disorders. *Lancet* **388**, 461–462.
- Gregertsen EC, Mandy W and Serpell L** (2017) The egosyntonic nature of anorexia. An impediment to recovery in anorexia nervosa treatment. *Frontiers in Psychology* **8**, 2273.
- Hay P** (2013) A systematic review of evidence for psychological treatments in eating disorders: 2005–2012. *International Journal of Eating Disorders* **46**, 462–469.
- Jacobson NS and Truax P** (1991) Clinical significance. a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology* **59**, 12–19.
- Kass AE, Kolko RP and Wilfley DE** (2013) Psychological treatments for eating disorders. *Current Opinion in Psychiatry* **26**, 549–555.
- Le Grange D** (2016) Anorexia nervosa in adults: the urgent need for novel outpatient treatments that work. *Psychotherapy (Chicago, Ill.)* **53**, 251–254.
- Schmidt U, Magill N, Renwick B, Keyes A, Kenyon M, Dejong H, Lose A, Broadbent H, Loomes R, Yasin H, Watson C, Ghelani S, Bonin EM, Serpell L, Richards L, Johnson-Sabine E, Boughton N, Whitehead L, Beecham J, Treasure J and Landau S** (2015) The maudslay outpatient study of treatments for anorexia nervosa and related conditions (MOSAIC): comparison of the Maudsley model of anorexia nervosa treatment for adults (MANTRA) with specialist supportive clinical management (SSCM) in outpatients with broadly defined anorexia nervosa: a randomized controlled trial. *Journal of Consulting and Clinical Psychology* **83**, 796–807.
- Schmidt U, Adan R, Böhm I, Campbell IC, Dingemans A, Ehrlich S, Elzackers I, Favaro A, Giel K, Harrison A, Himmerich H, Hoek HW, Herpertz-Dahlmann B, Kas MJ, Seitz J, Smeets P, Sternheim L, Tenconi E, van Elburg A, van Furth E and Zipfel S** (2016) Eating disorders: the big issue. *The Lancet Psychiatry* **3**, 313–315.
- Williams S and Reid M** (2010) Understanding the experience of ambivalence in anorexia nervosa. The maintainer's perspective. *Psychology & Health* **25**, 551–567.
- Zipfel S, Wild B, Gross G, Friederich HC, Teufel M, Schellberg D, Giel KE, de Zwaan M, Dinkel A, Herpertz S, Burgmer M, Lowe B, Tagay S, von Wietersheim J, Zeck A, Schade-Brittinger C, Schauenburg H and Herzog W and group, A. s** (2014) Focal psychodynamic therapy, cognitive behaviour therapy, and optimised treatment as usual in outpatients with anorexia nervosa (ANTOP study). randomised controlled trial. *Lancet* **383**, 127–137.
- Zipfel S, Giel KE, Bulik CM, Hay P and Schmidt U** (2016) Anorexia nervosa. Aetiology, assessment, and treatment. *The Lancet Psychiatry* **2**, 1099–1111.