

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS


<b>TITLE (PROVISIONAL)</b>	Long-term outcome of therapist-guided internet-based cognitive behavioural therapy for body dysmorphic disorder (BDD-NET): A naturalistic 2-year follow-up after a randomised controlled trial
<b>AUTHORS</b>	Enander, Jesper; Ljótsson, Brjánn; Anderhell, Lina; Runeborg, Martin; Flygare, Oskar; Cottman, Oskar; Andersson, Erik; Dahlén, Sofia; Lichtenstein, Linn; Ivanov, Volen; Mataix-Cols, David; Rück, Christian

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Charlotte Ulrikka Rask Center for Child and Adolescent Psychiatry, Risskov, Aarhus University Hospital
<b>REVIEW RETURNED</b>	19-Jun-2018

<b>GENERAL COMMENTS</b>	<p>Review: bmjopen-2018-024307</p> <p>Reviewer comments This paper presents results from a large long-term study on the effect of internet-based CBT in patients suffering from BDD. Overall, the paper is well-written and the results are clearly presented and nicely discussed – also in relation to future implications and perspectives. There are some features that need refinement and which could improve the paper even more, particularly more emphasis on the potential benefits of implementation of internet-based therapy (also possible cost-effectiveness advantages), further information on handling of missings in some of the analyses, more discussion of the potential issue with the lack of blinding of assessors in relation to the long term assessments and finally, more discussion of the lack of positive findings with regard to self-reported quality of life. Below follow more detailed comments which I would like the authors to address.</p> <p>Introduction This section is clearly written and really makes the case that research in this area is lacking with only three quite small studies on the long-term outcome of psychological treatment of BDD. It would be nice for the more naïve reader to have a more detailed presentation of the disorder at the beginning, such as information on prevalence, typical age of onset, core clinical characteristics etc.</p> <p>The authors suggest that internet-based therapy could increase access to evidence-based specialized care provided by experienced clinicians. This case could be made even stronger if the potential cost-effectiveness of the treatment format shown in</p>
-------------------------	--

	<p>relation to other disorders is also mentioned (including that few therapists can treat many patients simultaneously).</p> <p><b>Method and results</b>  A cut-off of at least 20 on the BDD-YBOCS was used in the assessment of eligible patients. Please provide more information on the clinical significance, i.e. does it correspond to cases with mild, moderate or severe BDD? Also, was this assessment supervised and was the inter-reliability among the assessors determined?</p> <p>Please provide a bit more information on the treatment setup, such as how many therapists were involved in the project and how many patients did each of them treat?</p> <p>It is stated that all patients had unlimited access to their individual therapist. Is it possible to provide some more information on the mean number of patient contacts and how much time the therapist spent on average on each patient? If yes, could these numbers be related to the patients' long term outcome and responder status?</p> <p>The patients were assessed by a structured diagnostic interview several times. Who did these assessments? The same persons who also performed the therapy (which could bias the results)? The potential limitation of the lack of blinding these assessments needs to be discussed in more detail in the later discussion of the overall findings. For example it is surprising that the gains found on the clinician-rated GAF score are not reflected in the results for the patients' self-reported quality of life. This could also be explained by lack of sensitivity of the used quality of life measure. Therefore, please provide more information on its psychometric properties.</p> <p>It is not clear how missing data were handled in the trajectory analyses of responder status. Please provide more information here.</p> <p>Only 49 patients participated in the brief interview about reasons for applying for participation. How were these patients recruited/selected and did they differ from the non-participants in this part of the study?</p> <p><b>Discussion</b>  Again, this is a well-written section. However, the mentioned possible limitations with regard to lack of blinding of assessors and problems with missing data in some of the analyses need to be further discussed. Also, the unexpected (?) finding of no positive results with regard to the patients' self-rated quality of life should be further elaborated.</p> <p><b>Final comment</b>  The authors mention that these self-referred patients may be especially motivated for treatment and that this treatment format may also support self-efficacy to a higher degree which may again support a sustainable long term treatment effect. Just out of curiosity: based on the performed post hoc interviews, was any participant feedback received on the treatment components that were actually utilized or found especially successful by the patients? If yes, this could be nice to add.</p>
--	--

<b>REVIEWER</b>	David Veale South London and Maudsley NHS Foundation Trust  Centre for Anxiety Disorders and Trauma 99 Denmark Hill London, SE5 8AZ
<b>REVIEW RETURNED</b>	25-Jun-2018

<b>GENERAL COMMENTS</b>	<p>This is a naturalistic follow up but is still a helpful addition to the literature in BDD. It is well written and the story is clear. The following are therefore minor queries.</p> <p>We are told that participants who sought further treatment in the follow up period had on average higher scores on the BDD-YBOCs, MADRs and have a past history of depression. Are there significant differences in variables between Responders and Non-responders at 2 year follow up?</p> <p>Was the observer rated Clinical Global Impression/ Improvement not used in the follow up as I thought It was used in the original study and help determine responder status</p> <p>On Table 2 it would be helpful to have the n at each follow up as it is not clear how many participants had multiple imputation for their data for each variable. I think this is a slight limitation as 12-24 month had reduced to 54-56 out of 88 and whether those who had dropped out are likely to bias the results.</p> <p>Comment - why does the effect size of the Euroquol and MADRs decrease from 1 to 2 years when BDD YBOCS and GAF continue to improve? I appreciate Quality of Life is often affected by mood, but it suggests the mood in a few patients had deteriorate</p>
-------------------------	---

### VERSION 1 – AUTHOR RESPONSE

Reviewer(s)' Comments to Author:

Reviewer: 1

Reviewer Name: Charlotte Ulrikka Rask

Reviewer comments

This paper presents results from a large long-term study on the effect of internet-based CBT in patients suffering from BDD. Overall, the paper is well-written and the results are clearly presented and nicely discussed – also in relation to future implications and perspectives. There are some features that need refinement and which could improve the paper even more, particularly more emphasis on the potential benefits of implementation of internet-based therapy (also possible cost-effectiveness advantages), further information on handling of missings in some of the analyses, more discussion of the potential issue with the lack of blinding of assessors in relation to the long term assessments and finally, more discussion of the lack of positive findings with regard to self-reported quality of life. Below follow more detailed comments which I would like the authors to address.

Introduction

COMMENT: This section is clearly written and really makes the case that research in this area is lacking with only three quite small studies on the long-term outcome of psychological treatment of BDD. It would be nice for the more naïve reader to have a more detailed presentation of the disorder at the beginning, such as information on prevalence, typical age of onset, core clinical characteristics etc.

RESPONSE: Thank you for this great suggestion. A more detailed description of BDD has been added to the introduction on page 4.

COMMENT: The authors suggest that internet-based therapy could increase access to evidence-based specialized care provided by experienced clinicians. This case could be made even stronger if the potential cost-effectiveness of the treatment format shown in relation to other disorders is also mentioned (including that few therapists can treat many patients simultaneously).

RESPONSE: We agree with reviewer 1 that the potential cost-effectiveness of internet-based CBT is important to mention. In general, cost-effectiveness research of internet-based CBT shows that this treatment modality has a high probability of being cost-effective when compared to no treatment or conventional CBT. We have added a statement about the cost-effectiveness of ICBT on page 5.

Reference: Hedman, Erik, Brjánn Ljótsson, and Nils Lindefors. "Cognitive behavior therapy via the Internet: a systematic review of applications, clinical efficacy and cost-effectiveness." *Expert review of pharmacoeconomics & outcomes research* 12.6 (2012): 745-764.

#### Method and results

COMMENT: A cut-off of at least 20 on the BDD-YBOCS was used in the assessment of eligible patients. Please provide more information on the clinical significance, i.e. does it correspond to cases with mild, moderate or severe BDD? Also, was this assessment supervised and was the inter-reliability among the assessors determined?

RESPONSE: We thank the reviewer for raising an important question. BDD-YBOCS has a total score of 0-48, with a higher score indicating more severe disorder. A score of 20 corresponds to moderate symptom severity. To ensure quality of assessments, clinicians in this trial practiced together on case examples with excellent reliability between raters (intraclass correlation 0.95, 95% confidence interval 0.89 to 0.98). We have added information about the clinical significance of BDD-YBOCS on page 6, and inter-rater reliability on page 7 of the manuscript.

COMMENT: Please provide a bit more information on the treatment setup, such as how many therapists were involved in the project and how many patients did each of them treat?

RESPONSE: In total, four therapists guided the participants through the treatment with patients equally split among them. This information had been added on page 6.

COMMENT: It is stated that all patients had unlimited access to their individual therapist. Is it possible to provide some more information on the mean number of patient contacts and how much time the therapist spent on average on each patient? If yes, could these numbers be related to the patients' long term outcome and responder status?

RESPONSE: The treatment content consists of self-help texts and worksheets that are delivered in eight interactive modules in the BDD-NET platform. The role of the therapists is mainly to guide the participants through the treatment, review homework and provide support. The duration of therapist contact and sent emails was automatically recorded by the BDD-NET platform. Median therapist time

spent weekly per participant reading and answering messages was 13.2 minutes. In total, therapists sent an average of 25.0 (SD=24.2) messages to participants during the treatment. This data has been added on page 6.

We are currently preparing a manuscript on predictors and moderators of responder status and improvement. However, we do not feel that it is within the focus of the current manuscript, and we would like to cover this important topic with the depth it deserves.

COMMENT: The patients were assessed by a structured diagnostic interview several times. Who did these assessments? The same persons who also performed the therapy (which could bias the results)? The potential limitation of the lack of blinding these assessments needs to be discussed in more detail in the later discussion of the overall findings. For example it is surprising that the gains found on the clinician-rated GAF score are not reflected in the results for the patients' self-reported quality of life. This could also be explained by lack of sensitivity of the used quality of life measure. Therefore, please provide more information on its psychometric properties.

RESPONSE: Thank you for the opportunity to clarify some aspects of the trial. Independent assessors performed the assessments. All assessors were blind to treatment allocation, except at the 1-year and 2-year follow-up (as all participants had received BDD-NET at those time points). I.e., blinding of these assessments were not possible considering the nature of the trial. This has been added as a limitation on page 14. EQ5D is primarily designed to assess quality of life in patients with somatic illnesses. It includes questions about anxiety/depression, but also mobility, pain, and self-care; none of which are usually impaired in BDD. Considering this, EQ5D may not be the ideal instrument for assessing quality of life in BDD. For example, Lamers et al. (2006) found EQ5D to be sensitive to change over time for patients with mild depression/anxiety, whereas de Willige et al. (2005) did not find a significant change on EQ5D after symptom improvements in more severely ill psychiatric patients. However, it is widely used in health economic evaluations to calculate quality-adjusted life years, allowing for future cost-effectiveness studies of BDD-NET. A thorough discussion of the non-significant findings on the EQ5D has been added on page 12.

#### References:

Van de Willige, Gerard, et al. "Changes in quality of life in chronic psychiatric patients: a comparison between EuroQol (EQ-5D) and WHOQoL." *Quality of Life Research* 14.2 (2005): 441-451.

Lamers, L. M., et al. "Comparison of EQ-5D and SF-6D utilities in mental health patients." *Health economics* 15.11 (2006): 1229-1236.

COMMENT: It is not clear how missing data were handled in the trajectory analyses of responder status. Please provide more information here.

RESPONSE: Data were imputed using multiple imputations with chained equations (including auxiliary variables to increase precision of the estimates). Responder status was calculated using data derived from the multiple imputations. This has been clarified in the manuscript on page 8.

COMMENT: Only 49 patients participated in the brief interview about reasons for applying for participation. How were these patients recruited/selected and did they differ from the non-participants in this part of the study?

RESPONSE: We agree with reviewer 1 that this section was a bit unclear. At the 12-month follow-up, 49 of 54 participants that took part in the follow-up agreed to also be briefly interviewed about reasons for applying to the original trial (the rest were lost to follow-up). This has been clarified in the

manuscript on page 11. Participants with missing data at follow-up had on average higher levels of depressive symptoms at baseline on the MADRS-S (M=22.6, SD=8.6), compared to participants who provided follow-up data (M=16.8, SD=8.7;  $t(86)=-30$ ,  $p=0.003$ ), but did not differ significantly in BDD symptom severity, level of insight or in any other demographic characteristics.

## Discussion

COMMENT: Again, this is a well-written section. However, the mentioned possible limitations with regard to lack of blinding of assessors and problems with missing data in some of the analyses need to be further discussed. Also, the unexpected (?) finding of no positive results with regard to the patients' self-rated quality of life should be further elaborated.

RESPONSE: Thank you for these great suggestions. We have further elaborated on the non-significant findings in EQ5D on page 12. We agree that the amount of participants lost to follow-up is a limitation and may introduce bias. Therefore, we carefully imputed the missing data in order to reduce bias and improve parameter estimates. In data simulation models, the use of multiple imputations has shown to perform well even with 50% of missing observations, and may reduce bias (Soullier et al., 2010). Compared to the complete-case analysis, the analysis using multiple imputations resulted in more conservative treatment effects. However, the estimates should be interpreted with caution. We have added a discussion of the missing data under the limitations section on page 14, as well as follow-up not being blinded.

Reference: Soullier, Noémie, Elise de La Rochebrochard, and Jean Bouyer. "Multiple imputation for estimation of an occurrence rate in cohorts with attrition and discrete follow-up time points: a simulation study." *BMC medical research methodology* 10.1 (2010): 79.

## Final comment

The authors mention that these self-referred patients may be especially motivated for treatment and that this treatment format may also support self-efficacy to a higher degree which may again support a sustainable long term treatment effect. Just out of curiosity: based on the performed post hoc interviews, was any participant feedback received on the treatment components that were actually utilized or found especially successful by the patients? If yes, this could be nice to add.

RESPONSE: Unfortunately we did not systematically record this during the post hoc interviews.

## Reviewer: 2

Reviewer Name: David Veale

Institution and Country: South London and Maudsley NHS Foundation Trust, Centre for Anxiety Disorders and Trauma, 99 Denmark Hill, London, SE5 8AZ

Please state any competing interests or state 'None declared': None

Please leave your comments for the authors below

This is a naturalistic follow up but is still a helpful addition to the literature in BDD.

It is well written and the story is clear.

The following are therefore minor queries.

COMMENT: We are told that participants who sought further treatment in the follow up period had on average higher scores on the BDD-YBOCs, MADRs and have a past history of depression. Are there significant differences in variables between Responders and Non-responders at 2 year follow up?

RESPONSE: Thank you for raising this question. At 2-year follow-up, non-responders had on average higher scores on MADRS, were more likely to be diagnosed with depression, and more likely to have had previous contact with psychiatric services (all Ps <0.05). No other significant differences on demographic or clinical variables emerged. This data has been added on page 10. We are currently preparing a manuscript on predictors and moderators of responder status and improvement, as we would like to cover this important topic with the depth it deserves.

COMMENT: Was the observer rated Clinical Global Impression/ Improvement not used in the follow up as I thought it was used in the original study and help determine responder status

RESPONSE: Thank you for giving us the opportunity to clarify this. In total, 56% (95% CI 45 to 69%) of participants were classified as much improved or very much improved on the Clinical Global Impression scale. This data has been added to the results on page 10. We choose the 30% cut-off based on the current literature where a >30% reduction on the BDD-YBOCS has a high sensitivity and specificity for correctly classifying treatment response.

Reference: Phillips, Katharine A., Ashley S. Hart, and William Menard. "Psychometric evaluation of the yale–brown obsessive-compulsive scale modified for body dysmorphic disorder (BDD-YBOCS)." *Journal of Obsessive-Compulsive and Related Disorders*, 3.3 (2014): 205-208.

COMMENT: On Table 2 it would be helpful to have the n at each follow up as it is not clear how many participants had multiple imputation for their data for each variable. I think this is a slight limitation as 12-24 month had reduced to 54-56 out of 88 and whether those who had dropped out are likely to bias the results.

RESPONSE: Number of participants at each follow-up has been clarified in Table 2. We thank reviewer 2 for giving us the opportunity to discuss this important issue as we agree that the amount of participants lost to follow-up is a limitation and may introduce bias. Therefore, we carefully imputed the missing data in order to reduce bias and improve parameter estimates. Compared to the complete-case analysis, the analysis using multiple imputations resulted in more conservative treatment effects; however, the estimates should be interpreted with caution. We have elaborated on the potential bias that may have been introduced due participants being lost to follow-up as a limitation on page 14.

COMMENT: Why does the effect size of the Euroquol and MADRs decrease from 1 to 2 years when BDD YBOCS and GAF continue to improve? I appreciate Quality of Life is often affected by mood, but it suggests the mood in a few patients had deteriorate.

RESPONSE: Thank you for raising this question. Most participants had mild to moderate depressive symptoms at baseline, and there was a slight increase in MADRS from 1 to 2 years, however this increase was non-significant. One could assume that there will be some fluctuation in depressive symptoms considering the time passed between the follow-ups. Moreover, the amount of missing data was somewhat higher for the self-reported outcomes compared to the clinician-assessed outcomes and this may have an impact on the parameter estimates. EQ5D was significantly improved from baseline to the 12-month follow-up (although the effect size was small), but EQ5D slightly decreased between 1 to 2 years and the improvement that had earlier been observed was no longer significant (p=.06). EQ5D is primarily designed to assess quality of life in patients with somatic illnesses, and includes questions about mobility, pain, and self-care; none of which are usually impaired in BDD. Considering this, EQ5D may not be the ideal instrument for assessing quality of life in BDD; however, it is widely used in health economic evaluations to calculate quality-adjusted life

years, allowing for future cost-effectiveness studies of BDD-NET. A discussion about the non-significant EQ5D findings has been added on page 12 of the manuscript.

#### VERSION 2 – REVIEW

<b>REVIEWER</b>	Charlotte Ulrikka Rask Department of Child and Adolescent Psychiatry, Research Unit, Psychiatry, Aarhus University Hospital, Denmark
<b>REVIEW RETURNED</b>	02-Aug-2018
<b>GENERAL COMMENTS</b>	Thank you for giving me the opportunity to review this revised manuscript..  The authors have addressed the concerns and points of clarification from the first review in a good and comprehensive manner. I have no further comments.
<b>REVIEWER</b>	David Veale South London and Maudsley NHS Foundation Trust, UK
<b>REVIEW RETURNED</b>	20-Aug-2018
<b>GENERAL COMMENTS</b>	I thought the authors have responded well and in the revision and this publication deserves to be published