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

TITLE (PROVISIONAL)	An investigation of general predictors for cognitive behavioral therapy outcome for anxiety disorders in a routine clinical setting
AUTHORS	Nielsen, Sara; Vangkilde, Signe; Wolitzky-Taylor, Kate; Daniel, Sarah; Hageman, Ida

VERSION 1 - REVIEW

REVIEWER	Jessica Freshour Michael E. DeBakey Veteran Affairs Medical Center, Baylor College of Medicine, Houston, TX USA
REVIEW RETURNED	23-Dec-2015

GENERAL COMMENTS	Nicely presented study protocol examining emotional regulation and attentional control as predictors of CBT for anxiety outcomes.

REVIEWER	Toomas Timpka Division of Social Medicine and Public Health Department of Medical and Health Sciences Faculty of Health Sciences Linköping University SE-581 83 Linköping Sweden
REVIEW RETURNED	12-Jan-2016

GENERAL COMMENTS	This protocol describe a study that is to examine attentional control and facets of emotion regulation as general predictors of outcome in CBT for anxiety disorders in a community-based mental health setting. Following predictors are included: - awareness and understanding of emotions, - acceptance of emotions, and - the ability to engage in goal-directed behavior, and refrain from impulsive behavior, when experiencing negative emotions. A secondary aim is to examine atheoretical variables (i.e., socio-demographic data and clinical characteristics, such as comorbidity) as predictors of CBT outcome General comments The protocol describes an interesting study. I have two suggestions to the authors regarding areas in which the manuscript could be improved. 1. The term "naturalistic study" is vague. Today, the concepts of practice-based research (Holmqvist et al 2015) and observational

	studies (Benchimol et al 2015) are well established and provide guidance for study design and reporting. The authors are encouraged to consider changing the term "naturalistic" to one of the above and/or comment on these in the manuscript. 2. The explanation of the sample size calculation could be expanded with a couple of sentences. As it stands, it is not clear whether and how the relations between outcome measures, predictors, and effect sizes are realistic and that the study is likely to provide useful information. References Holmqvist R, Philips B, Barkham M. Developing practice-based evidence: benefits, challenges, and tensions. Psychother Res. 2015;25(1):20-31. Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM; RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015 Oct 6;12(10):e1001885. doi: 10.1371/journal.pmed.1001885.

REVIEWER	Michael E. Miller Professor, Department of Biostatistical Sciences Wake Forest School of Medicine USA
REVIEW RETURNED	24-Jan-2016

GENERAL COMMENTS	There are several aspects of the statistical plan that could be clarified. First, the sample size sentence indicates that the chosen sample size provides “sufficient” power. The specific level of power (i.e. 80%, 90%) should be provided. The initial two sentences of the Data Analysis Section are a bit confusing. The design is a pre-post design where all patients will receive some planned amount of treatment, but most certainly not all patients will attend all sessions. The initial sentence indicates that repeated measures ANOVA will be used to investigate the effect of treatment on the two outcome measures. Some definition of “treatment” should be provided. Is this the change from baseline to post-treatment based on a contrast or is “treatment” based on how many sessions one attends? In addition, are both outcomes being used in the repeated measures ANOVA at each time point (i.e. 3 time points by 2 outcomes for 6 repeated measures) or is this one repeated measures ANOVA for each outcome? The 2nd sentence indicates that a mixed effects, ANCOVA will be used to explore the effect of attention control and emotion regulation on the results at the post-treatment. It would be good to mention what assumptions will be made about the covariance matrix within this model for the repeated outcomes. Is this the same model as is described in the 1st sentence that also includes the therapist-specific random intercept to account for clustering? For the analysis of treatment completers versus dropouts, a multivariate ANOVA is described. It would be helpful if more clarity was added here as to what the multiple outcomes are that require a

	multivariate ANOVA or is the term “multivariate” being used to indicate that multiple predictor variables will be used to predict treatment drop-out (a binary outcome)? This reviewer is confused as to what the model is that will be fit and why something like logistic regression for binary outcomes is not being used. Finally, additional details on the multiple imputation approach would be helpful. With the mixed model using maximum likelihood estimation being appropriate for data with “missingness” that is dependent on previously measured outcomes and covariates in the model (i.e. MAR), there may be little to gain from a multiple imputation approach that relies on those same covariates and previously collected outcomes.

REVIEWER	Victoria Harris Institute of Psychiatry, Psychology and Neuroscience at King's College London
REVIEW RETURNED	25-Jan-2016

GENERAL COMMENTS	I have been asked to provide a statistical review and therefore my comments are restricted to the statistical methods. With regard to your sample size calculation you have not specified your power. I have assumed 90% as is typical but this should probably be explicitly stated. Have you accounted for loss to follow up in your sample size? With regard to your missing data approach you may wish to look at linear mixed models if missingness is only expected to occur on the outcome. Linear mixed models have the advantage that the model can be estimated in the presence of missing data on the outcome at certain time points, as long as it is measured at least one time point. A multiple imputation approach would therefore only be required to estimate the model if there is missingness on the covariates. The abstract mentions that need for identifying reliable predictors of treatment effects. The data analysis section only mentions testing for main effects on post-treatment outcome, rather than testing for modification of the treatment effect, for which an interaction test would be required. An interaction test would be more informative as to whether patients with higher levels of emotion dysregulation benefit less from the treatment. The main limitation in the formal testing of interaction terms is likely to be due to a larger sample size being required to achieve adequate power.

VERSION 1 – AUTHOR RESPONSE

Reviewer 2:
Reviewer Name: Toomas Timpka

Reviewer 2 suggests that we change term "naturalistic study" to either “observational” or “practice-based” and that we elaborate on the explanation of the sample size calculation. The suggestions for literature were very relevant and helpful. We agree, and we have changed the term “naturalistic” to “practice-based study” and included Holmquist et al., 2015 as a reference. Furthermore, we have elaborated the section of sample size calculation.

Reviewer: 3
Reviewer Name: Michael E. Miller

Reviewer 4:
Reviewer Name:Victoria Harris

Both Reviewer 3 and Reviewer 4 comment on the plan for statistical analyses, which they find needs clarification. Furthermore, reviewer 4 suggests that a linear mixed modeling approach is advantageous. We agree and have edited the descriptions of statistical analyses to more clearly describe how we will implement the analyses in a mixed models framework. We thank the reviewers for drawing our attention to this and find that the manuscript has improved accordingly.

VERSION 2 – REVIEW

REVIEWER	Toomas Timpka Division of Social Medicine and Public Health Department of Medical and Health Sciences Linköping University SE-581 83 Linköping Sweden
REVIEW RETURNED	22-Feb-2016

GENERAL COMMENTS	The authors have done a proficient job with the revision. I have no further comments. The protocol is a valuable addition to this field of research.

REVIEWER	Michael E. Miller Wake Forest University School of Medicine USA
REVIEW RETURNED	22-Feb-2016

GENERAL COMMENTS	The revision to the statistical methods section is appropriate. The authors might also consider if their plan for baseline adjustment for the values of the outcomes (e.g. BAI score) will provide unbiased estimates of the relationships of interest in this non-randomized setting. There have been a number of papers written on this topic that the authors might find of interest (e.g. Fitzmaurice, A Conundrum in the Analysis of Change, Nutrition 17:360 –361, 2001; Glymour et al., When Is Baseline Adjustment Useful in Analyses of Change? An Example with Education and Cognitive Change. Am J Epidemiol 2005;162:267–278)

REVIEWER	Victoria Harris Institute of Psychiatry, Psychology and Neuroscience at King's College London
REVIEW RETURNED	03-Mar-2016

GENERAL COMMENTS	This is a statistical review. I am satisfied that the authors have now addressed all points raised with regard to the statistical analysis.