PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Which patients benefit specifically from short-term psychodynamic psychotherapy (STPP) for depression? Study protocol of a systematic review and meta-analysis of individual participant data.
AUTHORS	Driessen, Ellen; Abbass, Allan; Barber, Jacques P.; Connolly Gibbons, Mary Beth; Dekker, Jack; Fokkema, Marjolein; Fonagy, Peter; Hollon, Steven; Jansma, Elise; de Maat, Saskia; Town, Joel; Twisk, Jos; Van, Henricus; Weitz, Erica; Cuijpers, Pim

VERSION 1 – REVIEW

REVIEWER	Prof.ssa Cinzia Bressi Department of Pathophysiology and Transplantation DePT - School of Medicine, University of Milan Department of Neuroscience and Mental Health Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico Milan, Ialy Psychiatry, Major and minor clinical psychiatric disorders, Neuroscience, Psychotherapy, Psychoanalysis Reflective Functioning, Alexityhimia
REVIEW RETURNED	27-Aug-2017

GENERAL COMMENTS	Authors propose a significant and valuable study protocol about the systematic review and meta-analysis of individual participant data of Short Term Psychodynamic Psychotherapy (STPP) for depression. Previous works presented in literature adopted many different methodologies and techniques, with huge differences in sample sizes, pre-and-post variables and length of trials. A new protocol adopting standardized measures, with the aim of studying potential predictors and moderators in STPP for depressed patients, is welcome in the scientific psychotherapeutic field. However, I have doubts about the criteria that will define the STPP (which is described in p.10,11; line 203, 204 ,205). “First, STPP is based on psychoanalytic theories and practices; second, is time – limited from the onset (i.e. not a therapy that is a brief only in retrospect); third, it applies verbal techniques (e.g. STPP does not consider therapies applying art as an expression form)”. In my opinion, this work is very significant because of the value of the study protocol. The STPP criteria should be delineated in a more defined and specific frame.

	Indeed protocol parameters, such as number of sessions, length of session, face to face model of intervention, psychotherapeutic techniques and psychodynamic process variables, should be defined at the very beginning. About the measures, I would like to add alexythimia as a psychological patient characteristics because of the importance of such factor as a predictor of outcome in psychodynamic psychotherapy (Leweke F, Bausch S, Leichsenring F, Walter B, Stingl M, 2009). Indeed, alexithymia is a predictor of outcome of psychodynamically oriented inpatient treatment (Psychotherapy Research 19, 323-331). For ethics approval I thinks that all institutions must approve the study protocol. In other words must not be an optional choice. I really appreciate the limits that were considered but I hope they are not too optimistic Finally, I believe that the protocol presented in this study has both clinical and scientific relevance.

REVIEWER	Gilles Ambresin CHUV centre hospitalier universitaire vaudois Institut Universitaire de Psychothérapie du Département de psychiatrie-CHUV Bâtiment Les Cèdres Site de Cery - 1008 Prilly +41 (0)21 314 21 98 TEL +41 (0)21 314 05 86 SEC +41 (0)21 314 27 84 FAX gilles.ambresin@chuv.ch I practice and provide training in STPP for depressed inpatients. No other declared
REVIEW RETURNED	03-Sep-2017

GENERAL COMMENTS	Thank you for giving me the opportunity to review the protocol by Driessen et al., which I read with much interest. I found no major limitations regarding originality, importance, and scientific reliability of the protocol. The design appears ethically and procedurally sound. You will find comments, which might warrant minor clarifications in the method section, in the text below. Review The manuscript by Driessen et al. describes a study protocol aiming at identifying and examining predictors and moderators of short-term psychodynamic psychotherapy (STPP) efficacy for depression. To investigate this, authors will adopt an adequate systematic review methodology with a subsequent sound meta-analytic procedure. Meta-analyses will be conducted at the individual participant-level data. Authors conclude that increased knowledge of such predictors and moderators may have important clinical implications as it offers new evidence that is relevant to clinicians and patients. * Originality Understanding the effects of predictors and moderators is important, as suggestions of the efficacy of STPP in the treatment of depression exist.

	Authors are aware of previous literature related to their research question indicating the importance of predictors and moderators for the treatment of depression (e.g. Driessen et al., 2015; Driessen et al., 2016; Barber et al., 2012). However, these results are only preliminary and a systematic assessment is lacking. This protocol is also original in the methods used. Traditional meta-analyses have considerable limitations in testing for moderators and predictors of treatment outcomes. Using individual patient-level data will address this issue as it provides greater power to investigate interactions between predictors or moderators and treatment effects. (Fischer et al., 2017) * Importance of work Short-term psychodynamic psychotherapy is commonly used in the treatment of depressed patients. Clear indications have been found to show that STPP is effective in the depression for adults. (Driessen et al., 2015; Fonagy, 2015) As its reported effects are usually moderate rather than large, it is of interest to better identify which patient may benefit more specifically of STPP. The proposed study does matter to clinicians and patients as it may give indications for treatment planning and valuable information to both of them. The current protocol is also important for researchers in the field of psychoanalytic psychotherapy. It will provide them with detailed information that is not given in the summary of the research posted in PROSPERO. It may also help them develop their actual or future research. Last, researchers who shared their data will be offered co-authorship. This will foster collaboration between researchers in the field. Scientific reliability * Research Question Aim of the study is clearly defined. Scientific reliability of the answer is very likely to be sound. Authors will use up-to-date methods for the systematic review and for the meta-analysis of the collected data. * Overall design of study Description of the overall design of study is thorough and accurate. * Participants studied and studies included The method section provides a description of who participants will be. Authors write that: 'Participants are considered depressed if they meet specified criteria for major depression' (Line 199). Could they be more specific? Have they decided to include studies that assessed participants with semi-structured interview, or based on clinicians' assessment or both? What do they mean by an 'elevated score on a standardized measure of depression' (Line 200)? Does it mean that any score above the 'no depression' cut-off will be considered? Alternatively, it may mean that a certain degree of severity is required for inclusion in the review. Authors used predefined criteria for the inclusion of interventions in their review. They present a definition of STPP. * Methods Method section follows current recommendations. Methods are adequately described and they comply with relevant reporting standards (ie PRISMA-P 2015). I would like to add a special note to the efforts authors will make to identify available studies outside the PubMed Search.

	Such a wide search is of invaluable help for the development of a proper and comprehensive literature review. Table 1 reports the results of the PubMed Query with the final number of items found at the top. It may be more amenable to the reader to have the results reported according to the logical sequence of the query. Authors describe a very thorough check of data integrity. Measures are clearly stated and primary, secondary and tertiary outcomes are accurately delineated. They will deal with missing data by performing multiple imputation which is relevant. In this study, authors will use an increasingly popular meta-analytic approach which is meta-analysis of individual participant data. The raw individual level data for each study will be obtained and used for synthesis. This can facilitate the derivation of the information desired, it may also increase the number of participants and the length of the follow-up compared to those reported in the original publication. Authors will adopt a one-step approach and they specify their assumption of the meta-analysis. This approach can inform how treatment effect is modified by study level characteristics and patient level characteristics. This procedure is increasingly popular in medicine and has been used in psychiatry and depression (eg Fournier et al., 2010; Geddes et al., 2009). One protocol (Weitz et al., 2017 BMJ Open) with three authors of this protocol will use an IPD meta-analysis to examine the effect of individual patient characteristics as moderators on the efficacy of combined treatment and comparator treatment for depression. To the best of my knowledge it has not been used for the analysis of the treatment effect of STPP for depressed patients. I am not a specialist in IPD meta-analysis, a review by a statistician may be recommended here. IPD meta-analysis is resource intensive and may require advanced statistical experience, which may not be a problem for this large research group. Authors have set a procedure to collect data as extensively as possible, which should prevent major bias due to poor provision of individual data. They have also set a check of the quality of the data that should address the potential bias due to the eventual poor quality of the original studies. I will not develop further my comments on this section as authors really followed current recommendations and present their methodology clearly and thoroughly. * Conclusion Overall, the conclusion section is warranted by and sufficiently derived from the methods. * References References seem up to date and relevant. *Abstract Abstract reflects accurately the content of the paper. Review of protocol based on the notes from the Editors for study protocols The protocol paper reports on an ongoing study running since 1st of December 2016 and planned to be completed by the 30th of November, 2018. The dates of the study are included in the manuscript. The PRISMA-P 2015 Checklist is a checklist that has been adapted for use with protocol submissions. The protocol by Driessen et al. complies with all items of the PRISMA_P 2015 Checklist. Item 2 ‘Registration’ can now be filled in.

	Since the manuscript has been submitted to BMJ Open a registration number has been assigned to the protocol (PROSPERO 2017:CRD42017056029). This should be updated (Line 70) No results or conclusions are present in the study protocol. I found no major flaw in the study that would prevent a sound interpretation of the data. Ethics Authors of the current protocol will invite authors of the included studies to share the participant-level data of their studies. Included studies should have received local IRB approval. This is usually the case but I would encourage authors of the current protocol to double check in the published studies. A line could be added in the protocol under Ethics and Dissemination. I would also suggest that this should be reported in their main paper.

REVIEWER	Olavi Lindfors National Institute for Health and Welfare, Helsinki, Finland
REVIEW RETURNED	07-Sep-2017

GENERAL COMMENTS	The study protocol is sound, comprehensive and clearly written, and focused on the important issue of which patients specifically benefit from short-term psychodynamic psychotherapy. The methods and analysis and specific procedures regarding joint analysis of patient-level data have been described in detail. This meta-analysis is a needed contribution for evaluating the applicability of short-term psychodynamic psychotherapy, based on patient characteristics.

REVIEWER	Johannes C. Ehrenthal Department of Psychology Alpen-Adria University Klagenfurt Austria The author practises and writes about various forms of psychodynamic psychotherapy.
REVIEW RETURNED	11-Sep-2017

GENERAL COMMENTS	The authors present a concise and generally well-written manuscript about a highly important endeavor: To examine possible patient-characteristics that predict a better outcome specifically in short-term psychodynamic psychotherapy (STPP), and to examine similar variables that predict a better outcome in STPP as compared to another treatment or non-treatment-conditions by means of individual participant data (IPD) meta-analysis. The manuscript may benefit from some clarifications with regard to content, or specifications with regard to methods. Please find some comments below. 0. Registration Just to be formally correct: As the authors are of course aware of, the trial is now registered: PROSPERO 2017:CRD42017056029. 1. Introduction I would suggest to reframe parts of the introduction to increase coherence of the deduction of the research-questions (especially p. 5). In particular, it may be helpful to put less emphasis on a lack of RCTs on STPP, but rather describe today's more solid empirical basis that calls for predictor- and moderator-analyses.

	Otherwise, to some readers it could be a bit confusing to call for research on predictors and mediators before establishing general efficacy. (On a side note, the German national consensus guidelines for unipolar depression do include STPP as a valid treatment option similar to for example CBT.) It may be helpful to use a clearer wording when describing the definition of predictors and moderators (pp. 5-6). Some general words on patient variables may be good to contextualize the research as well as cited research findings (p. 6). For example, the effect of pre-treatment symptom severity on treatment-effect is well known and probably not specific to STPP. A little bit more theory, models, and assumptions about general vs. specific predictors/moderators would strengthen the manuscript. 2. Methods I wonder why the search strategy excludes letters. Some journals, for example Psychotherapy & Psychosomatics, do publish possibly relevant original studies in the format of a letter (p. 9). Although study selection criteria are quite clear (pp. 10-11), I am wondering if the authors should expand their definition more toward a bona fide treatment, and include, or at least control for therapist variables such as allegiance, competence, training, and the like. With regard to data collection (pp. 11-12) I do have some concerns, especially for more recent trials. 1. As a researcher providing data, I would probably not release any measures that I intend to publish myself. It would be great to have an opt-out option for certain measures, or an option to hold back analyses until own manuscripts have been published, and something like a steering committee to take care of these issues. 2. This is especially relevant if the authors of the current study intend to pass the dataset to third-party-researchers as well. In the 'Measures' section, I wonder if the authors would also consider looking into initiatives like the PROMIS system, to make measures comparable by more sophisticated means than just z-standardization (p. 14). At least it may be important to recognize that not all depression-measures measure the same, not even all versions of the same instrument (see for example the different versions of the Hamilton Rating Scale). The same is true for other measures, for example concerning attachment. In general, taking into account differences in healthcare-systems – either conceptually or empirically - could also prove to be of value.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1

Reviewer Name: Prof.ssa Cinzia Bressi

Institution and Country: Department of Pathophysiology and Transplantation, DePT - School of Medicine, University of Milan, Department of Neuroscience and Mental Health, Fondazione IRCCS Cà Granda Ospedale Maggiore Policlinico, Milan, Italy Please state any competing interests or state

‘None declared’: Psychiatry, Major and minor clinical psychiatric disorders, Neuroscience, Psychotherapy, Psychoanalysis Reflective Functioning, Alexityhimia

Please leave your comments for the authors below: Authors propose a significant and valuable study protocol about the systematic review and meta-analysis of individual participant data of Short Term Psychodynamic Psychotherapy (STPP) for depression.

Previous works presented in literature adopted many different methodologies and techniques, with huge differences in sample sizes, pre-and-post variables and length of trials.
A new protocol adopting standardized measures, with the aim of studying potential predictors and moderators in STPP for depressed patients, is welcome in the scientific psychotherapeutic field.

Comment 3:
However, I have doubts about the criteria that will define the STPP (which is described in p.10,11; line 203, 204 ,205). “First, STPP is based on psychoanalytic theories and practices; second, is time – limited from the onset (i.e. not a therapy that is a brief only in retrospect); third, it applies verbal techniques (e.g. STPP does not consider therapies applying art as an expression form)”.

In my opinion, this work is very significant because of the value of the study protocol. The STPP criteria should be delineated in a more defined and specific frame. Indeed protocol parameters, such as number of sessions, length of session, face to face model of intervention, psychotherapeutic techniques and psychodynamic process variables, should be defined at the very beginning.

Reply:
We agree with this reviewer that the inclusion criteria for STPP in this review are quite broad. We also agree that STPPs can vary with regard to the abovementioned treatment characteristics and that it is important to take such differences into account. However, we prefer not to use defined and specific STPP criteria, as specifying such narrow inclusion criteria might result in a considerable number of relevant studies being excluded from consideration, which we find undesirable.

Rather than specifying narrow STPP inclusion criteria, we choose to apply quite broad inclusion criteria, as doing so would result in including as many relevant studies as possible with varying STPP treatment characteristics. This will allow us to study the consistency of predictor/moderator relationships across different STPP treatment settings. It will also allow us to examine the relationships between certain therapy characteristics (e.g., number of sessions) and STPP efficacy.

For these purposes, we actually do define STPP protocol parameters for each study. However, we can see how this might not have been reported clearly enough in the previous version of the manuscript, where we stated:

We will also extract multiple STPP characteristics and study design characteristics (for an overview see[12])

We have now changed this as follows in hopes of clarifying this matter (lines 270-274):
For each study, we will list all predictor/moderator variables that were assessed, as well as all outcome variables, intermediate, and follow-up assessments. We will also extract multiple STPP characteristics (e.g., number of sessions, treatment format, STPP mode) and study design characteristics (e.g., therapist training, treatment integrity check, use of a treatment manual; for a complete overview see[4]).

We will also conduct sensitivity analyses to examine the robustness of our prediction/moderation findings with regard to STPP treatment characteristics, which we now note explicitly in the methods section (389-392):

Finally, we will examine the impact of STPP characteristics (e.g., STPP type, delivery mode) and study design characteristics (e.g., therapist training, use of a treatment manual) by adding these variables to the mixed effects models too.

Comment 4:
About the measures, I would like to add alexythimia as a psychological patient characteristics because of the importance of such factor as a predictor of outcome in psychodynamic psychotherapy (Leweke F, Bausch S, Leichsenring F, Walter B, Stingl M, 2009). Indeed, alexithymia is a predictor of outcome of psychodynamically oriented inpatient treatment (Psychotherapy Research 19, 323-331).

Reply:
We agree that alexithymia is an important psychological patient characteristic and we aim to examine this variable as a potential predictor/moderator of STPP efficacy provided that sufficient data are available. We now note this explicitly in the methods section (lines 307-312):

Potential predictors and moderators include socio-demographic variables (e.g., gender, age, education level, marital status, employment status, ethnicity), clinical variables (e.g., number of previous depressive episodes, previous exposure to treatment, comorbid Axis I and II psychopathology, global assessment of functioning), and psychological patient characteristics (e.g., personality organization, attachment, interpersonal styles, childhood maltreatment, alexithymia).

Comment 5:
For ethics approval I thinks that all institutions must approve the study protocol. In other words must not be an optional choice.
I really appreciate the limits that were considered but I hope they are not too optimistic Finally, I believe that the protocol presented in this study has both clinical and scientific relevance.

Reply:
We agree that institutional ethics approval for investigators sharing patient-level data is necessary. We, therefore, have the following clause in our data sharing agreement:

Institution represents and warrants that the Data has been collected and is transferred to VU in accordance with all applicable local and international laws and regulations.

Thus, by signing the data sharing agreement, the investigators confirm that they share their data according to all applicable local and international laws and regulations, including but not limited to institutional ethics approval.

We now note this in the manuscript as follows (lines 421-427):
IRB approval was not required for this project. IRB approval may be required for the investigators to share their primary data depending on their institution's policies. It is the responsibility of the investigators to obtain IRB approval if their institution's policies require them to do so. By signing the data sharing agreement, the authors who share their data declare that those data were collected and transferred to our research group according to all applicable local and international laws and regulations, including but not limited to local IRB approval.

Reviewer: 2
Reviewer Name: Gilles Ambresin
Institution and Country: CHUV, centre hospitalier universitaire vaudois, Institut Universitaire de Psychothérapie du, Département de psychiatrie-CHUV, Bâtiment Les Cèdres, Site de Cery - 1008 Prilly Please state any competing interests or state 'None declared': I practice and provide training in STPP for depressed inpatients. No other declared

Dear Miss Gray,
Thank you for giving me the opportunity to review the protocol by Driessen et al., which I read with much interest.

I found no major limitations regarding originality, importance, and scientific reliability of the protocol. The design appears ethically and procedurally sound.

You will find comments, which might warrant minor clarifications in the method section, in the text below.

Review
The manuscript by Driessen et al. describes a study protocol aiming at identifying and examining predictors and moderators of short-term psychodynamic psychotherapy (STPP) efficacy for depression. To investigate this, authors will adopt an adequate systematic review methodology with a subsequent sound meta-analytic procedure. Meta-analyses will be conducted at the individual participant-level data. Authors conclude that increased knowledge of such predictors and moderators may have important clinical implications as it offers new evidence that is relevant to clinicians and patients.

* Originality
Understanding the effects of predictors and moderators is important, as suggestions of the efficacy of STPP in the treatment of depression exist. Authors are aware of previous literature related to their research question indicating the importance of predictors and moderators for the treatment of depression (e.g. Driessen et al., 2015; Driessen et al., 2016; Barber et al., 2012). However, these results are only preliminary and a systematic assessment is lacking.

This protocol is also original in the methods used. Traditional meta-analyses have considerable limitations in testing for moderators and predictors of treatment outcomes. Using individual patient-level data will address this issue as it provides greater power to investigate interactions between predictors or moderators and treatment effects. (Fischer et al., 2017)

* Importance of work
Short-term psychodynamic psychotherapy is commonly used in the treatment of depressed patients. Clear indications have been found to show that STPP is effective in the depression for adults. (Driessen et al., 2015; Fonagy, 2015) As its reported effects are usually moderate rather than large, it is of interest to better identify which patient may benefit more specifically of STPP.

The proposed study does matter to clinicians and patients as it may give indications for treatment planning and valuable information to both of them. The current protocol is also important for researchers in the field of psychoanalytic psychotherapy. It will provide them with detailed information that is not given in the summary of the research posted in PROSPERO. It may also help them develop their actual or future research. Last, researchers who shared their

data will be offered co-authorship. This will foster collaboration between researchers in the field.

Scientific reliability

* Research Question
Aim of the study is clearly defined. Scientific reliability of the answer is very likely to be sound. Authors will use up-to-date methods for the systematic review and for the meta-analysis of the collected data.

* Overall design of study
Description of the overall design of study is thorough and accurate.

* Participants studied and studies included

Comment 6:

The method section provides a description of who participants will be. Authors write that: 'Participants are considered depressed if they meet specified criteria for major depression' (Line 199). Could they be more specific? Have they decided to include studies that assessed participants with semi-structured interview, or based on clinicians' assessment or both? What do they mean by an 'elevated score on a standardized measure of depression' (Line 200)? Does it mean that any score above the 'no depression' cut-off will be considered? Alternatively, it may mean that a certain degree of severity is required for inclusion in the review. Authors used predefined criteria for the inclusion of interventions in their review. They present a definition of STPP.

Reply:
We agree that we could have been more specific in our description of the depression inclusion criterion. We have rewritten the relevant sentence as follows in hopes of clarifying this matter (lines 202-206):

Participants are considered depressed if they meet specified criteria for major depressive disorder or another mood disorder as assessed by means of a semi-structured interview or clinicians' assessment, or if they present an elevated score above the 'no depression' cut-off on a standardized measure of depression.

* Methods

Comment 7:
Method section follows current recommendations. Methods are adequately described and they comply with relevant reporting standards (ie PRISMA-P 2015). I would like to add a special note to the efforts authors will make to identify available studies outside the PubMed Search. Such a wide search is of invaluable help for the development of a proper and comprehensive literature review. Table 1 reports the results of the PubMed Query with the final number of items found at the top. It may be more amenable to the reader to have the results reported according to the logical sequence of the query.

Reply:
Thank you for recognizing our extensive literature search. We agree that Table 1 could benefit from presenting the query in a more logical sequence and have adjusted this Table as suggested (line 185).

Authors describe a very thorough check of data integrity.
Measures are clearly stated and primary, secondary and tertiary outcomes are accurately delineated. They will deal with missing data by performing multiple imputation which is relevant.

In this study, authors will use an increasingly popular meta-analytic approach which is meta-analysis of individual participant data. The raw individual level data for each study will be obtained and used for synthesis. This can facilitate the derivation of the information desired, it may also increase the number of participants and the length of the follow-up compared to those reported in the original publication. Authors will adopt a one-step approach and they specify their assumption of the meta-analysis. This approach can inform how treatment effect is modified by study level characteristics and patient level characteristics.
This procedure is increasingly popular in medicine and has been used in psychiatry and depression (eg Fournier et al., 2010; Geddes et al., 2009). One protocol (Weitz et al., 2017 BMJ Open) with three authors of this protocol will use an IPD meta-analysis to examine the effect of individual patient characteristics as moderators on the efficacy of combined treatment and comparator treatment for depression. To the best of my knowledge it has not been used for the analysis of the treatment effect of STPP for depressed patients. I am not a specialist in IPD meta-analysis, a review by a statistician may be recommended here.

IPD meta-analysis is resource intensive and may require advanced statistical experience, which may not be a problem for this large research group. Authors have set a procedure to collect data as extensively as possible, which should prevent major bias due to poor provision of individual data. They have also set a check of the quality of the data that should address the potential bias due to the eventual poor quality of the original studies. I will not develop further my comments on this section as authors really followed current recommendations and present their methodology clearly and thoroughly.

* Conclusion
Overall, the conclusion section is warranted by and sufficiently derived from the methods.

* References
References seem up to date and relevant.

*Abstract
Abstract reflects accurately the content of the paper.

Review of protocol based on the notes from the Editors for study protocols
The protocol paper reports on an ongoing study running since 1st of December 2016 and planned to be completed by the 30th of November, 2018. The dates of the study are included in the manuscript.

Comment 8:
The PRISMA-P 2015 Checklist is a checklist that has been adapted for use with protocol submissions. The protocol by Driessen et al. complies with all items of the PRISMA_P 2015 Checklist. Item 2 ‘Registration’ can now be filled in. Since the manuscript has been submitted to BMJ Open a registration number has been assigned to the protocol (PROSPERO 2017:CRD42017056029). This should be updated (Line 70)
No results or conclusions are present in the study protocol.
I found no major flaw in the study that would prevent a sound interpretation of the data.

Reply:
We have now updated the manuscript to include the PROSPERO registration number in the abstract and methods section. Please see Comment 2 by the Editor above for more details.

Comment 9:

Ethics
Authors of the current protocol will invite authors of the included studies to share the participant-level data of their studies. Included studies should have received local IRB approval. This is usually the case but I would encourage authors of the current protocol to double check in the published studies. A line could be added in the protocol under Ethics and Dissemination. I would also suggest that this should be reported in their main paper

Reply:
We agree that the primary studies of which we collect patient-level data should have received local IRB approval. We, therefore, have the following clause in our data sharing agreement:

Institution represents and warrants that the Data has been collected and is transferred to VU in accordance with all applicable local and international laws and regulations.

Thus, by signing the data sharing agreement, the investigators who share their data confirm that those data were collected according to all applicable local and international laws and regulations, including but not limited to local IRB approval.

We now note this in the manuscript as follows (lines 421-427):

IRB approval was not required for this project. IRB approval may be required for the investigators to share their primary data depending on their institution's policies. It is the responsibility of the investigators to obtain IRB approval if their institution's policies require them to do so. By signing the data sharing agreement, the authors who share their data declare that those data were collected and transferred to our research group according to all applicable local and international laws and regulations, including but not limited to local IRB approval.

Reviewer: 3
Reviewer Name: Olavi Lindfors
Institution and Country: National Institute for Health and Welfare, Helsinki, Finland Please state any competing interests or state 'None declared': None declared

Please leave your comments for the authors below
The study protocol is sound, comprehensive and clearly written, and focused on the important issue of which patients specifically benefit from short-term psychodynamic psychotherapy. The methods and analysis and specific procedures regarding joint analysis of patient-level data have been described in detail. This meta-analysis is a needed contribution for evaluating the applicability of short-term psychodynamic psychotherapy, based on patient characteristics.

Reviewer: 4
Reviewer Name: Johannes C. Ehrenthal

Institution and Country: Department of Psychology, Alpen-Adria University Klagenfurt, Austria Please state any competing interests or state ‘None declared’: The author practises and writes about various forms of psychodynamic psychotherapy.

Please leave your comments for the authors below The authors present a concise and generally well-written manuscript about a highly important endeavor: To examine possible patient-characteristics that predict a better outcome specifically in short-term psychodynamic psychotherapy (STPP), and to examine similar variables that predict a better outcome in STPP as compared to another treatment or non-treatment-conditions by means of individual participant data (IPD) meta-analysis.

The manuscript may benefit from some clarifications with regard to content, or specifications with regard to methods. Please find some comments below.

0. Registration

Comment 10:
Just to be formally correct: As the authors are of course aware of, the trial is now registered: PROSPERO 2017:CRD42017056029.

Reply:
During the review process, PROSPERO did indeed assign a registration number to our project. We have now updated the manuscript to include this registration number in the abstract and methods section. Please see Comment 2 by the Editor above for more details.

Comment 11:

1. Introduction
I would suggest to reframe parts of the introduction to increase coherence of the deduction of the research-questions (especially p. 5). In particular, it may be helpful to put less emphasis on a lack of RCTs on STPP, but rather describe today’s more solid empirical basis that calls for predictor- and moderator-analyses. Otherwise, to some readers it could be a bit confusing to call for research on predictors and mediators before establishing general efficacy. (On a side note, the German national consensus guidelines for unipolar depression do include STPP as a valid treatment option similar to for example CBT.)

Reply:
Thank you for this suggestion. We can see how our mentioning the discussion regarding the general efficacy of STPP for depression might be confusing for the reader. We have now removed this section and rewritten the first paragraph of the introduction as follows (lines 95-105):

Depression is a highly prevalent and disabling disorder associated with major personal and societal costs.[1] Affecting more than 300 million people worldwide, depression is ranked as the single largest contributor to global disability by the World Health Organization.[2] Given the tremendous burden of disease, there is a great need for effective and efficient treatments for depression. Antidepressant medications and different psychological therapies constitute the predominant treatments for depressive disorders.[3] Concerning psychological treatments, there is a clinical tradition of short-term psychodynamic psychotherapies (STPPs) being used to treat depression. STPP is an empirically supported treatment for depression.[4] However, it is unlikely that any treatment will work for equally well for all depressed patients[5] and it remains largely unclear if certain subgroups of patients can benefit specifically from STPP.

Comment 12:
It may be helpful to use a clearer wording when describing the definition of predictors and moderators (pp. 5-6).

Reply:
We can see how our definitions of predictors and moderators could benefit from further clarification. We have added examples in hopes of achieving this (lines 106-118):

Two types of information are relevant to this question: predictors and moderators. Predictors (or prognostic factors) predict outcome to a given treatment and can be used to determine which patients are more likely to respond to STPP relative to other patients. For instance, if age were found to be a positive predictor of STPP efficacy, this might indicate that older patients would be more likely to benefit from STPP than younger patients. Predictors can inform expectations of STPP efficacy, but are of little use in deciding which treatment to select. On the other hand, moderators (or prescriptive factors) can detect different patterns of outcomes between different treatments for different types of patients and provide a basis for choosing the best treatment for a given patient.[6] For instance, if age were found to be a moderator of STPP efficacy versus antidepressant medication, this might indicate that older patients might benefit more from STPP than from medication, while younger patients might benefit more from medication than from STPP.

Comment 13:

Some general words on patient variables may be good to contextualize the research as well as cited research findings (p. 6). For example, the effect of pre-treatment symptom severity on treatment-effect is well known and probably not specific to STPP. A little bit more theory, models, and assumptions about general vs. specific predictors/moderators would strengthen the manuscript.

Reply:
We agree that predictors and moderators of STPP efficacy can either apply specifically to STPP or can be more general factors associated with depression treatment efficacy, and that it is important to make this distinction. We have therefore added the following (lines 119-124):

Some preliminary empirical findings concerning predictors and moderators of STPP efficacy for depression do exist. With regard to predictors, meta-regression analyses alongside a ‘conventional’ meta-analysis (based on results extracted from publications[4]) showed that mean pre-treatment depression scores were positively associated with pre- to post-treatment depression effect size, although this effect might not be specific to STPP[7].

We can see the value of providing a more elaborate discussion of theory, models, and assumptions about general vs. specific predictors/moderators. However, since we already exceeded the suggested word limit for this manuscript, we decided to refrain from such a discussion at this point. In the articles describing the outcomes of this research project, we aim, however, to provide an explicit discussion of general versus specific in the context of the predictors/moderators identified. We now note this as follows in the discussion (lines 504-508):

Fifth, the predictors and moderators identified in this study can either apply specifically to STPP or can be more general factors associated with depression treatment efficacy. We intend to address this distinction in this study's outcome reports in the context of the predictors/moderators identified.

2. Methods

Comment 14:
I wonder why the search strategy excludes letters. Some journals, for example Psychotherapy & Psychosomatics, do publish possibly relevant original studies in the format of a letter (p. 9).

Reply:
We excluded letters from the PubMed search to reduce the number of irrelevant search results. However, this reviewer is correct in noting that some journals, such as Psychotherapy & Psychosomatics, publish original studies in the format of a letter. In fact, we just started the literature search update and came across such a study (Kolaitis et al., 2014). We think these studies will be picked up in our literature search nevertheless, because we search in multiple electronic databases, not all of which exclude letters. Moreover, we perform the following checks on our database searches (lines 191-198):

Fourth, we will search an Internet database of controlled and comparative outcome studies on psychological treatments of depression (http://www.psychotherapyrcts.org[14]) for studies examining STPP. Fifth, reviews and meta-analyses concerning the efficacy of psychodynamic treatments for depression or for psychiatric disorders in general retrieved from the first search method will be screened for relevant references not located by means of the other search methods. Sixth, we will contact an email list of researchers in the field of psychodynamic therapy to ask for ongoing or unpublished studies.

Thus, although letters are excluded from the PubMed search, we consider it unlikely that relevant studies published in the format of a letter will be missed, for the abovementioned reasons.

Comment 15:
Although study selection criteria are quite clear (pp. 10-11), I am wondering if the authors should expand their definition more toward a bona fide treatment, and include, or at least control for therapist variables such as allegiance, competence, training, and the like.

Reply:
This relates to Comment 3 by Reviewer 1 above. As mentioned in the reply to that comment, we agree that STPPs can vary with regard to the abovementioned characteristics and that it is important to take such differences into account. We aim to do so, but we can see how this might not have been reported clearly enough in the previous version of the manuscript, where we stated:

We will also extract multiple STPP characteristics and study design characteristics (for an overview see[12])

We have now changed this as follows in hopes of clarifying this matter (lines 270-274):

For each study, we will list all predictor/moderator variables that were assessed, as well as all outcome variables, intermediate, and follow-up assessments. We will also extract multiple STPP

characteristics (e.g., number of sessions, treatment format, STPP mode) and study design characteristics (e.g., therapist training, treatment integrity check, use of a treatment manual; for a complete overview see[4]).

We will conduct sensitivity analyses to examine the robustness of our prediction/moderation findings with regard to varying treatment characteristics, which we now note explicitly in the methods section (389-392):

Finally, we will examine the impact of STPP characteristics (e.g., STPP type, delivery mode) and study design characteristics (e.g., therapist training, use of a treatment manual) by adding these variables to the mixed effects models too.

Comment 16:
With regard to data collection (pp. 11-12) I do have some concerns, especially for more recent trials. 1. As a researcher providing data, I would probably not release any measures that I intend to publish myself. It would be great to have an opt-out option for certain measures, or an option to hold back analyses until own manuscripts have been published, and something like a steering committee to take care of these issues. 2. This is especially relevant if the authors of the current study intend to pass the dataset to third-party-researchers as well.

Reply:
These are valid concerns that researchers of more recent trials might indeed have when being asked to share their data. To increase chances that relevant data can be obtained for this project, we aim to be flexible with regard to our agreements with the investigators in such cases. For instance, it is certainly possible for researchers to share parts of their dataset. It is also possible to coordinate the timing of publication, such that the trial results will be published before publication of the meta-analyses' results. As preparing IPD meta-analyses' reports usually takes quite a while (given the large number of co-authors, who all need to approve the manuscript), this can be quite feasible.

All such issues can be negotiated and included in the data sharing agreement. We think such an agreement is preferable over a steering committee, which might be considered biased towards the needs of this research project, while a data sharing agreement is legally bound and more balanced towards the needs of both parties.

With regard to use of the collective database by third parties, the data sharing agreement states the following:

The Data is used by VU for the sole purpose of the Analyses and for no other purpose. The Data shall not be transferred to third Parties by VU without obtaining prior written consent from Institution.

Thus, our research group cannot pass the dataset to third-party-researchers without the investigators' consent and investigators can always decline sharing their data for purposes other than the current project.

Comment 17:
In the 'Measures' section, I wonder if the authors would also consider looking into initiatives like the PROMIS system, to make measures comparable by more sophisticated means than just z-

standardization (p. 14). At least it may be important to recognize that not all depression-measures measure the same, not even all versions of the same instrument (see for example the different versions of the Hamilton Rating Scale). The same is true for other measures, for example concerning attachment. In general, taking into account differences in healthcare-systems – either conceptually or empirically - could also prove to be of value.

Reply:
From what we know of the PROMIS system, we can see its value for future IPD meta-analyses of trials with PROMIS as their outcome measure. However, we find it more difficult to see how the outcome data previously collected by studies already completed could be transformed in a way that fits this system. However, if such a possibility exists, we would be very interested in applying it to the current project.

We recognize that measures of depression and other variables can vary. Indeed, standardizing such variables is a core challenges in IPD meta-analyses. One of the ways in which we aim to handle this, is by conducting the sensitivity analyses that are described at lines 299-301:

Sensitivity analyses will be conducted using unstandardized scores for each depression measure that is assessed in the majority of studies included in the meta-analysis (e.g., Hamilton Depression Rating Scale[19], Beck Depression Inventory[20]).

Thus, if depression is assessed by means of different measures across studies, we will examine if the moderation/prediction results hold when only one (version of the) depression outcome measure is considered.

To account for differences in health-care systems, we will conduct sensitivity analyses to examine the robustness of our prediction/moderation findings with regard to varying STPP treatment characteristics described in lines 389-392.

Reference cited:

Kolaitis G, Giannakopoulos G, Tomaras V, et al. Self-esteem and social adjustment in depressed youths: a randomized trial comparing psychodynamic psychotherapy and family therapy. Psychother Psychosom 2014;83:249-51.

VERSION 2 – REVIEW

REVIEWER	Gilles Ambresin CHUV, centre hospitalier universitaire vaudois, Institut Universitaire de Psychothérapie du, Département de psychiatrie-CHUV, Bâtiment Les Cèdres, Site de Cery - 1008 Prilly
REVIEW RETURNED	26-Oct-2017

GENERAL COMMENTS	The Authors have satisfactorily addressed my comments and relevantly altered the manuscript.