STUDY PLANS FOR THE INDIVIDUAL PARTICIPANT DATA META-ANALYSIS OF ANTIDEPRESSANT TRIALS FOR MAJOR DEPRESSION IN JAPAN

Steering Committee

Prof Shigeto Yamawaki (Hiroshima University, Former President of CINP)
Prof Emeritus Jun Ishigooka (Tokyo Women's College of Medicine, Former President of JSNP)
Dr Kazutaka Ikeda (Tokyo Metropolitan Institute of Medical Science, President of JSNP)
Prof Toshiaki Furukawa (Kyoto University)
Prof Shiro Tanaka (Kyoto University)
A/Prof Norio Watanabe (Kyoto University)
A/Prof Hisashi Noma (The Institute of Statistical Mathematics)
Dr Kazushi Maruo (National Center of Neurology and Psychiatry)
A/Prof Andrea Cipriani (University of Oxford)

(Data Manager)
Dr Hissei Imai (Kyoto University)

Principles of data sharing and data analyses

1. Data will be securely managed and only members of the Steering Committee (SC) may have access to individual data. A medium (either a hard disk or a USB memory) containing the individual participant data will be stored in a locked room and data will be analyzed on a computer that is not connected to the internet, except for the data from GSK which will be analyzed online.
2. All analyses will follow the study plans as specified and agreed-upon below. A new contract will be required if any other analyses are to be undertaken.
3. The members of SC will conduct the analyses, write up and publish the papers.
4. When any doubt arises, SC will request clarification from the relevant manufacturer. All data used in the PMDA submission and CSR will be used as such.
5. SC retains the right to publish all the results of the analyses below, both primary and secondary, that have been conducted according to the following study plans.
6. None of the following analyses, primary or secondary, will contain any direct, overall comparison of efficacy or acceptability among active drugs, as the network is a so-called star-shaped network and will therefore not allow any informative analyses allowing such comparisons. Specific drug names will be mentioned in the section of the included studies but not in the results section for global efficacy or acceptability comparison. However, some subgroups may be identified who are particularly responsive or unresponsive to particular drugs.
7. A number of trials have active drug comparators but the data for these active comparators may not be shared due to the license agreement with the manufacturer of these drugs. Because the number of participants allocated to such active drug comparators is usually small, the present study can proceed without these data.
8. The names of the members of the Working Group and of the manufactures will be acknowledged in the acknowledgement section of the publications but they will not be co-authors or investigators.
9. All manufacturers agreeing to the above principles promise to provide all data, if present, that will be requested in the data template accompanying the study plans.
10. Primary analyses #1-#2 will be registered in PROSPERO. Primary analysis #3 and secondary analyses will be more exploratory and will not be pre-registered. We intend to publish all of these analyses in the end.

Primary analyses

The main purposes of the present study are threefold:

A) Identify subgroups of patients that show greater differentiation between antidepressants and placebo, in order to contribute to the planning of more efficient trials in the future
B) Identify subgroups of patients that show smaller differentiation between antidepressants and placebo, in order to identify unmet medical needs
   ① It would be meaningful to examine biomarkers for this subset of patients too.
C) Specify placebo responders (narrowly defined), in order to distinguish between true drug responders and placebo responders among those who apparently respond to drugs and thus help identify bio-markers of drug responses

and we will conduct the following primary analyses to address these questions. None of the following analyses will contain any direct, overall comparison of efficacy or acceptability among active drugs.

1 The first study will focus on baseline depression severity as effect modifier for efficacy, because this is an unresolved, hot issue in clinical psychopharmacology and will have a great clinical implication.

1.1 The primary outcome will be the Hamilton Rating Scale for Depression (HAMD) or Montgomery-Asberg Rating Scale (MADRS) as specified as primary in the original trial. When different versions of HAMD or MADRS are used, the scores will be converted using the conversion table based on the item response theory [1].

1.2 All arms within the following licensed dose range (FDA, EMA, PMDA) will be treated as effective dose.

	licensed dose	very low	low	moderate to high	very high
bupropion	300-400	150	300		
desvenlafaxine	50-100	25	50		
duloxetine	40-120		40, 60		
escitalopram	10-20		10	20	
mirtazapine	15-45		15	30, 45	
paroxetine	20-50		20	40	
paroxetine CR	25-62.5		25	50	
venlafaxine	75-375		75	75-225 (mean=202)	

1.3 We will then conduct IPD-MA to examine the relationship between baseline symptom severity and the differences in change scores between the drugs and placebo using a three-level mixed-effects model with maximum likelihood or restricted maximum likelihood estimation [2,3]. The levels will account for the data structure such that: level 1 represented time, level 2 the participant, and level 3 the trial. The following competing models with increasing complexity will be applied: (Model 1) time and treatment and the two-way time by treatment interaction; (Model 2) Model 1 plus baseline symptom score and all two-way interactions among time, treatment and baseline score; (Model 3) Model 2 plus the three-way interaction of linear time by treatment by baseline score; and (Model 4) Model 3 plus the two-way and three-way interactions among quadratic time, treatment and baseline score. The time variable will be treated as the categorical or continuous variable. These models will be tested unadjusted and adjusted for confounders. [4]

2 The second study will look at all other effect predictors and effect modifiers for efficacy and dropouts.

2.1 IPD-MA of efficacy

2.1.1 The primary outcome will be the Hamilton Rating Scale for Depression (HAMD) or Montgomery-Asberg Rating Scale (MADRS) as specified as primary in the original trial. When different versions of HAMD or MADRS are used, the scores will be converted using the conversion table based on the item response theory [1].

2.1.2 The literature suggests many candidates for effect predictors (variables associated with response regardless of the treatment) and for effect modifiers (variables associated with differential response depending on the treatment) in the treatment of depression [5]. We have listed the possible candidate variables for effect predictors and effect modifiers based on the literature in the following. The variables to be actually examined will first be limited by their availability in the included original studies but when several variables that measure similar concepts are available, the research team will discuss which ones we believe are the most important predictors and which should be included in the model.

Demographics

1) Age [6]

Life and social history

2) Childhood maltreatment [7]
3) Education [8]
4) Employment [9,10]
5) Marital status [9-11]
6) Recent life events and difficulties [9,10]
7) Social adjustment/function [12]

History of present illness

8) Age at onset [13]
9) Chronicity [6]
10) No of previous episodes [8,14]
11) Prior treatments with antidepressants [10]

Present illness: symptomatology

12) Baseline severity [15-17]
13) Baseline psychomotor symptoms [12,18]
14) Baseline anxiety symptoms [18,19]
15) Comorbid personality disorder [10]
16) Comorbid substance use/abuse [18]

Therapeutic process

17) Early response [20]
18) Co-prescriptions other than antidepressants

2.1.3 We will then conduct IPD-MA to examine the relationship between each independent variable and the differences in change scores between the drugs and placebo using a linear mixed-effects model with maximum likelihood or restricted maximum likelihood estimation as delineated above. The between-studies and between-drugs heterogeneities will be modelled by random effects. The effect modifications are evaluated by involving treatment-by-covariates interaction terms in the mixed-effects model analysis. The strategy of modeling variables will be explanatory determined because there have been only limited prior information what variables have predictive abilities and how these variables should be modelled in the statistical models. However, we will use the candidate variables noted above and these analyses would be conducted within the framework of linear mixed model analyses. We expect to assemble 4000 participants' individual data which will provide enough power to examine the available variables.

2.1.4 In addition, we would attempt other data-driven subgroup identification methods which might detect predictive factors and responder subgroups more efficiently, if necessary. We will adopt the SIDES (subgroup identification based differential effect search) procedure and its companion methods [21-23] that effectively identify responder subgroups. The SIDES and its companion procedures are flexible searching algorithms that split the patient population for exploring the subpopulations that the treatment effects would be particularly different. These algorithms can also control the random errors effectively in the comprehensive explanatory analyses. The candidate variables and the analysis statistical models will be limited as the same with Section 2.1.3.

2.1.5 The obtained model will be used to build subgroups of those who are likely to respond on drug but not on placebo, those who are likely to respond both on drug and on placebo, and those who are likely not to respond to drug or placebo.

2.1.6 The following sensitivity analyses will be conducted using meta-regression in order to examine the robustness of the obtained results.

1) Use of placebo run-in phase
2) Probability to be allocated to placebo[24]

2.2 IPD-MA of response, dropouts for inefficacy, dropouts for side effects, and dropouts for remission on antidepressants and placebo

2.2.1 The average dropout rate in antidepressant trials hovers around 35-40% [25] and casts serious doubt on the validity of their analyses. Identifying predictors and moderators of dropouts will contribute to recruitment of subjects less likely to dropout and thus increase the internal validity of the future trials of antidepressants.

2.2.2 We expect predictors and modifiers for dropouts due to inefficacy or remission to be similar across drugs. Those for dropouts due to side effects will likely be different across drugs.

2.2.2.1 Cytochrome P450 subtypes were measured for 2C19 in escitalopram and 2D6 for venlafaxine only.

2.2.3 The IPD-MA should be conducted using logistic mixed effects model that accommodates between-studies and between-drugs heterogeneities. The outcome variables are response, dropouts for inefficacy, dropouts for side effects, and dropouts for remission at the follow-up periods. The effect modifications are evaluated by involving treatment-by-covariates interaction terms and the strategy of selection of modeling variables will be explanatory determined as the same with the analyses for efficacy in Section 2.1.3.

- Studies #1 and #2 will be registered in PROSPERO.

3 Predictors of response on placebo

3.1 The primary outcome will be response defined as 50% or greater reduction on any observer-rated depression scale specified as primary in the original study.

3.2 Logistic regression with random effects for explaining between-studies heterogeneity will be used to predict response on placebo. Candidates of covariates in the model (we will use the same list as above) will be screened through a backward variable selection with the critical value of $p = 0.15$.

3.3 In addition, we would attempt other data-driven subgroup identification methods which

might detect predictive factors and responder subgroups more efficiently, if necessary. [21-23]

3.4 External validity of the final multi-variate logistic regression will be evaluated with the use of leave-one-trial-out cross-validation. Hosmer-Lemeshow test will be used as a measure for calibration and C-statistics will be used as a measure for discrimination.

3.5 If the model is proven to be stable and efficient, it will then be used to distinguish true drug responders and placebo responders among those who apparently respond on drugs.

Secondary analyses

Because the data that we envision to obtain in this study is so rich in contents, we would like to conduct the following additional analyses to advance the clinical trial methodology and the understanding of depression symptomatology. The following analyses are mainly exploratory in nature. None of the following analyses will contain any direct, overall comparison of efficacy or acceptability among active drugs.

4 Based on pre-post data

4.1 Will raising the baseline eligibility threshold decrease the standard deviation and thereby increase the effect size of the trial? We have already undertaken a similar analysis in schizophrenia antipsychotic trials and found that it cannot. We will use the same methodology and examine if the results are different or similar for antidepressants.
--Furukawa TA & Leucht S (2013) Can we inflate effect size and thus increase chances of producing "positive" results if we raise the baseline threshold in schizophrenia trials? Schizophrenia Research, 144, 105-108.

4.2 Do effect size estimates differ depending on the definitions of response and remission? We have already conducted a similar study in schizophrenia antipsychotic trials and shown that relative indices of effect remain constant, regardless of the definitions of response and remission. We will use the same methodology and examine if the results are different or similar for antidepressants.
--Furukawa TA, Akechi T, Wagenpfeil S & Leucht S (2011) Relative indices of treatment effect may be constant across different definitions of response in schizophrenia trials. Schizophrenia Research, 126, 212-219.

4.3 Change scores vs endpoint scores: do they lead to different estimates in individual RCTs and can they be mixed in meta-analyses? There are conflicting reports on this methodological question important in evidence synthesis. Using the IPD we can examine the possible differences in the pooled results of change scores and endpoint scores.
--Fu R & Holmer HK (2016) Change score or follow-up score? Choice of mean difference estimates could impact meta-analysis conclusions. Journal of Clinical Epidemiology.
--da Costa BR, Nuesch E, Rutjes AW, Johnston BC, Reichenbach S, Trelle S, Guyatt GH & Juni P (2013) Combining follow-up and change data is valid in meta-analyses of continuous outcomes: a meta-epidemiological study. Journal of Clinical Epidemiology, 66, 847-855.

4.4 Which items in HAMD or MADRS are most sensitive to change in depression trials? A recent re-analysis based on IPD has revealed that the first item of HAMD may be the most sensitive. We will use the same methodology and examine which items of HAMD or MADRS (or QIDS-SR and IDS-SR) are most sensitive to change.
--Hieronymus F, Emilsson JF, Nilsson S & Eriksson E (2016) Consistent superiority of selective serotonin reuptake inhibitors over placebo in reducing depressed mood in patients with major depression. Molecular Psychiatry, 21, 523-530.

4.5 The old "myths" of antidepressant responsiveness can now be more accurately addressed from the view point of (i) absolute response rates (i.e. are the response rates in the following subgroups higher than those in the others) and (ii) relative response (i.e. are the RR of response in the following subgroups higher than in the others). The "myths" to be addressed are

4.5.1 Antidepressants are ineffective in depression due to exogenous stressors (stressors were measured in escitalopram trials)

4.5.2 Antidepressant are particularly effective in endogenous depression (DSM-IV melancholic features were measured in almost all trials, diurnal variation is measured in SIGH-D)

4.6 Variability in response rates, dropout rates, and efficacy across sites/region

4.6.1 It would be impossible to identify the same sites across studies.

5 Based on longitudinal, repeated measurements

5.1 Trajectory of symptoms of depression: Each symptom and symptom factors of HAMD and MADRS will be examined as to its time course. Similar items of self-rated symptoms from QIDS and IDS may also be compared.

5.2 What early signs within the first 1 or 2 weeks predict later response and remission on antidepressants and on placebo? e.g. Early non-response has been shown to predict later non-response. Those who complain of typical side effects, such as nausea on SSRI/SNRI and sleepiness on mirtazapine, may show higher response than those without any such side effects.
This is also a hot topic in recent literature and the following studies have each examined related but different predictors.
--Mitchell AJ (2006) Two-week delay in onset of action of antidepressants: new evidence. Br J Psychiatry, 188, 105-106.
--Tylee A & Walters P (2007) Onset of action of antidepressants. BMJ, 334, 911-912.
--Henkel V, Seemuller F, Obermeier M, Adli M, Bauer M, Mundt C, Brieger P, Laux G, Bender W, Heuser I, Zeiler J, Gaebel W, Mayr A, Moller HJ & Riedel M (2009) Does early improvement triggered by antidepressants predict response/remission? Analysis of data from a naturalistic study on a large sample of inpatients with major depression. J Affect Disord, 115, 439-449.
--Lam RW (2012) Onset, time course and trajectories of improvement with antidepressants. European Neuropsychopharmacology, 22 Suppl 3, S492-498.

5.3 Sustainability of response on drug and placebo: What happens to early responders to placebo (with or without placebo run-in) and drug? We will identify early responders on placebo, defined as >25% decrease within the first one week and some similar but variable definitions, and describe their later course in comparison with the other patients.
--Quitkin FM, McGrath PJ, Rabkin JG, Stewart JW, Harrison W, Ross DC, Tricamo E, Fleiss J, Markowitz J & Klein DF (1991) Different types of placebo response in patients receiving antidepressants. The American Journal of Psychiatry, 148, 197-203.
--Quitkin FM, Stewart JW, McGrath PJ, Nunes E, Ocepek-Welikson K, Tricamo E, Rabkin JG & Klein DF (1993) Further evidence that a placebo response to antidepressants can be identified. The American Journal of Psychiatry, 150, 566-570.

5.4 Methodological comparison of LOCF, MI and MMRM
Several studies have already examined the results of the above-mentioned analyses in clinical trial data. However, no similar comparison has been conducted with antidepressant

trials.
-- Mallinckrodt CH, Clark WS & David SR (2001) Type I error rates from mixed-effects model repeated measures versus fixed effects analysis of variance with missing values imputed via last observation carried forward. Drug Information Journal, 35, 1215-1225.
--Cook RJ, Zeng L & Yi GY (2004) Marginal analysis of incomplete longitudinal binary data: a cautionary note on LOCF imputation. Biometrics, 60, 820-828.
--Leucht S, Engel RR, Bauml J & Davis JM (2007) Is the superior efficacy of new generation antipsychotics an artifact of LOCF? Schizophrenia Bulletin, 33, 183-191.
--Siddiqui O, Hung HM & O'Neil R (2009) MMRM vs LOCF: a comprehensive comparison based on simulation study and 25 NDA datasets. Journal of Biopharmaceutical Statistics, 19, 227-246.
-- Grobler AC, Matthews G & Molenberghs G (2014) The impact of missing data on clinical trials: a re-analysis of a placebo controlled trial of Hypericum perforatum (St Johns wort) and sertraline in major depressive disorder. Psychopharmacology, 231, 1987-1999.

5.5 Low dose vs high dose: is high dose clinically meaningful? The following new study is the first study ever to show dose-response in antidepressants. Because the studies to be included in the current study has often examined two different dosages of the same drug in the fixed regimen, it will be possible to examine the difference between the lower dose vs the upper dose within the licensed dosage. We will conduct a three-node IPD network meta-analysis of placebo vs low dose vs moderate to high dose in order to estimate the relative efficacy of placebo vs low dose vs moderate to high dose arms.

--Jakubovski E, Varigonda AL, Freemantle N, Taylor MJ & Bloch MH (2016) Systematic Review and Meta-Analysis: Dose-Response Relationship of Selective Serotonin Reuptake Inhibitors in Major Depressive Disorder. American Journal of Psychiatry, 173, 174-183.

5.6 Methodological research on new methods for discovering subgroup differences within IPD-MA: In the literature, several effective data-driven subgroup identification methods have been proposed, which will be used in Section 2. However, these methods generally focus on applications within a randomized clinical trial, which has limited statistical powers for detecting treatment-by-covariates interactions. IPD-MA are expected to gain statistical powers via synthesizing much larger statistical information for multiple randomized trials. However, there is ample room of improving such methods and we would like to develop new methods to identify subgroup differences within IPD-MA, and their effectiveness will be evaluated via applications to this IPD-MA. These analyses will be conducted within the framework of Section 2.

-- Lipkovich I, Dmitrienko A, B. R. & D' Agostino S (2017) Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials. Statistics in Medicine, 36, 136-196.

list of available studies

Overview of Placebo-controlled MDD studies conducted in Japan (based on the published data)

	Fluoxetine [8]	Venlafaxine ER [10]	Vortioxetine [9]	Bupropion SR [7]	Paroxetine CR [2]	Escitalopram [4]	Desvenlafaxine [5]	Duloxetine [1]	Miltazapine [6]	Escitalopram [10]
Phase	Ph3	Ph3	Ph3	Ph3	Ph3	Ph3	Ph3	Ph3	Ph3	Ph2
Region	Japan	Japan	Japan	Japan (80%) Korea (20%)	Japan (89%) Korea (11%)	Japan	Japan (52%) US (48%)	Japan	Japan	Japan
Study Period	Mar. 2013- Jul. 2014	Nov. 2011- Mar. 2014	May 2011- Dec. 2012	Jun. 2010- Sep. 2012	Apr. 2009- Feb. 2010	Apr. 2008- Dec. 2010	Dec. 2008- Apr. 2010	Jun. 2006- Sep. 2007	Nov. 2004- Dec. 2005	Nov. 2004- Dec. 2005
Primary Measure	HAMD21 / GRID-HAMD	HAMD17	MADRS	MADRS	HAMD17 / SIGH-D	MADRS	HAMD17	HAMD17	HAMD17 / SIGH-D	HAMD17 / SIGH-D
% of Patients Received Placebo	51% (259/510)	34% (184/535)	34% (123/364)	33% (186/564)	42% (171/412)	26% (124/484)	33% (231/699)	34% (148/440)	26% (70/270)	34% (100/297)
Active Control	N	N	N	N	Y (Paroxetine)	Y (Paroxetine)	N	Y (Paroxetine)	N	N
Primary Result	**Negative**	**Positive**	**Negative**	**Negative**	**Positive**	**Positive**	**Positive**	**Positive**	**Positive**	**Negative**
Age (SD) /placebo	38.5 (11.7)	38.6 (11.1)	37.6 (10.7)	37.9 (11.1)	36.8 (10.07)	36.4 (10.8)	40 (12)	38.7 (10.5)	39.9 (12.8)	34.7 (8.6)
first or reccurent eposode /placebo	N.A.	N.A.	F: 61.3% R: 38.7%	F: 53% R: 46%	F: 53% R: 47%	F: 57.3% R: 42.8%	N.A.	F: 61.4% R: 38.7%	F: 64.3% R: 27.2% unknown 8.6%	F: 59.0% R: 41.0%
Current MDD Episode /placebo	N.A.	N.A.	N.A.	Mean:26.8 (19.4)week Median: 21.0 week	Mean:33.5 (23.65) week	Mean: 13.2 (22.8) months	Mean: 18 (28) months	N.A.	Mean: 12.4 (17.7) months	Mean: 5.8 (11.0) months
Mean total score (SD) at Baseline /placebo	N.A.	N.A.	MADRS 32.5(4.5) HAMD17 21.5 (4.48)	N.A.	HAMD17 22.6 (2.75)	MADRS 29.0 (5.6)	HAMD17 23 (3)	HAMD17 22.7 (3.3)	HAMD17 22.5 (3.6)	HAMD17 22.5 (3.6)
Key Inclusion Criteria	Total HAMD21 ≥ 20 CGI-S ≥ 4	Total MADRS ≥ 26 CGI-S ≥ 4 Current single episode ≥ 90 days Recurrent episode ≥ 28 days QIDS16-J-SR ≥ 16	Total MADRS ≥ 26 CGI-S ≥ 4 Current episode ≥ 3 months	Total HAMD17 ≥ 20 CGI-S ≥ 4 Current episode ≥ 8 weeks and < 24 months IDS-SR ≥ 25 Subscale (Item 5) IDS-SR ≥ 7	Total HAMD17 ≥ 20 Depressed mood ≥ 2	Total MADRS ≥ 22 CGI-S ≥ 4 Current episode ≥ 4 weeks	Total HAMD17 ≥ 20 CGI-S ≥ 4 Depressed mood ≥ 2 Current episode ≥ 30 days	Total HAMD17 ≥ 19 Depressed mood ≥ 2	Total HAMD17≥ 18	Total HAMD17 ≥ 18 Depressed mood ≥ 2

*We expect to obtain IPD from all the studies except for fluoxetine and vortioxetine.

Time schedule

December 21, 2016	1st meeting for the workgroup
January 18, 2017	2nd meeting for the workgroup, with A/Prof Andrea Cipriani The study plan will be finalized in English. Japanese abstract will be prepared.
By end of February, 2017	The non-disclosure agreement will be signed off by respective manufacture and the Steering Committee. For drugs by Pfizer, we will need to make requests at Pfizer Inspire https://iirsubmission.pfizer.com/_layouts/InspiirePortal/ （TAF has checked that NCT01441440 (venlafaxine trial) and NCT00798707 (desvenlafaxine trial) are available as IPD at this site. The web page provides the email address to send the study protocol to.） For drugs by GSK, we will make requests through clinialstudydatarequest.com as “CINP/JSNP PPP Joint Task Force study plans”.
By end of April, 2017	We would like to obtain the following data. ● Individual participant data as used in the PMDA submission and CSR (in csv format) ● List of definitions of variables
By end of June, 2017	Finish formatting the obtained data in the common format, ask for any necessary additional clarifications and fix the dataset

July 2017 -	Analyses and publications

References

1. Carmody TJ, Rush AJ, Bernstein I, et al. The Montgomery Asberg and the Hamilton ratings of depression: a comparison of measures. *Eur Neuropsychopharmacol* 2006; **16**(8): 601-11.
2. Hedeker D, Gibbons RD. Longitudinal Data Analysis. Hoboken, New Jersey: John Wiley & Sons, Inc.; 2006.
3. Gibbons RD, Hur K, Brown CH, Davis JM, Mann JJ. Benefits from antidepressants: Synthesis of 6-week patient-level outcomes from double-blind placebo-controlled randomized trials of fluoxetine and venlafaxine. *Arch Gen Psychiatry* 2012.
4. Schwarz G, Gideon E. Estimating the dimension of a model. *Annals of Statistics* 1978; **6**(2): 461-4.
5. Kessler RC, Bossarte R, Brenner L, et al. Using patient self-reports to study heterogeneity of treatment effects in major depressive disorder. *Epidemiology and Psychiatric Sciences* in press.
6. Cuijpers P, Reynolds CF, 3rd, Donker T, Li J, Andersson G, Beekman A. Personalized treatment of adult depression: medication, psychotherapy, or both? A systematic review. *Depress Anxiety* 2012; **29**(10): 855-64.
7. Nemeroff CB, Heim CM, Thase ME, et al. Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. *Proc Natl Acad Sci U S A* 2003; **100**(24): 14293-6.
8. Perlis RH. A clinical risk stratification tool for predicting treatment resistance in major depressive disorder. *Biol Psychiatry* 2013; **74**(1): 7-14.
9. Fournier JC, DeRubeis RJ, Shelton RC, Hollon SD, Amsterdam JD, Gallop R. Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *J Consult Clin Psychol* 2009; **77**(4): 775-87.
10. DeRubeis RJ, Cohen ZD, Forand NR, Fournier JC, Gelfand LA, Lorenzo-Luaces L. The Personalized Advantage Index: translating research on prediction into individualized treatment recommendations. A demonstration. *PLoS ONE* 2014; **9**(1): e83875.
11. Barber JP, Muenz LR. The role of avoidance and obsessiveness in matching patients to cognitive and interpersonal psychotherapy: empirical findings from the treatment for depression collaborative research program. *J Consult Clin Psychol* 1996; **64**(5): 951-8.
12. Frank E, Cassano GB, Rucci P, et al. Predictors and moderators of time to remission of major depression with interpersonal psychotherapy and SSRI pharmacotherapy. *Psychol Med* 2011; **41**(1): 151-62.
13. Andreescu C, Mulsant BH, Houck PR, et al. Empirically derived decision trees for the treatment of late-life depression. *Am J Psychiatry* 2008; **165**(7): 855-62.
14. Jarrett RB, Minhajuddin A, Kangas JL, Friedman ES, Callan JA, Thase ME. Acute phase cognitive therapy for recurrent major depressive disorder: who drops out and how much do patient skills influence response? *Behav Res Ther* 2013; **51**(4-5): 221-30.
15. Fournier JC, DeRubeis RJ, Hollon SD, et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 2010; **303**(1): 47-53.
16. Driessen E, Cuijpers P, Hollon SD, Dekker JJ. Does pretreatment severity moderate the efficacy of psychological treatment of adult outpatient depression? A meta-analysis. *J Consult Clin Psychol* 2010; **78**(5): 668-80.
17. Weitz ES, Hollon SD, Twisk J, et al. Baseline depression severity as moderator of depression outcomes between cognitive behavioral therapy vs pharmacotherapy: An individual patient data meta-analysis. *JAMA psychiatry* 2015; **72**(11): 1102-9.
18. Rush AJ, Wisniewski SR, Warden D, et al. Selecting among second-step antidepressant medication monotherapies: predictive value of clinical, demographic, or first-step treatment features. *Arch Gen Psychiatry* 2008; **65**(8): 870-80.
19. Ninan PT, Rush AJ, Crits-Christoph P, et al. Symptomatic and syndromal anxiety in chronic

forms of major depression: effect of nefazodone, cognitive behavioral analysis system of psychotherapy, and their combination. *J Clin Psychiatry* 2002; **63**(5): 434-41.
20. Steidtmann D, Manber R, Blasey C, et al. Detecting critical decision points in psychotherapy and psychotherapy + medication for chronic depression. *J Consult Clin Psychol* 2013; **81**(5): 783-92.
21. Lipkovich I, Dmitrienko A, Denne J, Enas G. Subgroup identification based on differential effect search--a recursive partitioning method for establishing response to treatment in patient subpopulations. *Stat Med* 2011; **30**(21): 2601-21.
22. Lipkovich I, Dmitrienko A. Strategies for identifying predictive biomarkers and subgroups with enhanced treatment effect in clinical trials using SIDES. *J Biopharm Stat* 2014; **24**(1): 130-53.
23. Lipkovich I, Dmitrienko A, B. R. D' Agostino S. Tutorial in biostatistics: data-driven subgroup identification and analysis in clinical trials. *Stat Med* 2017; **36**(1): 136-96.
24. Papakostas GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind, randomized clinical trials in MDD. *Eur Neuropsychopharmacol* 2009; **19**(1): 34-40.
25. Furukawa TA, Cipriani A, Atkinson LZ, et al. Revisiting placebo response rates in antidepressant trials: a systematic review of published and unpublished double-blind studies. *Lancet Psychiatry* 2016.