Melancholic Features in Bipolar Depression and Response to Lamotrigine

A Pooled Analysis of Five Randomized Placebo-Controlled Trials

Evyn M. Peters, MD,* Yanbo Zhang, FRCPC,† Rohit Lodhi, FRANZCP,*
Hua Li, PhD,‡ and Lloyd Balbuena, PhD*

Abstract:

Background: A pilot study suggested lamotrigine may be more effective for bipolar depression with melancholic features. We tested this hypothesis in a pooled analysis of 5 randomized double-blind placebo-controlled trials of lamotrigine for acute bipolar depression.

Methods: The pooled sample consisted of 1072 adult outpatients. Depressive symptoms were assessed for 7 to 10 weeks with the Hamilton Depression Rating Scale and the Montgomery-Åsberg Depression Rating Scale. The outcome measure was end-trial response (score reduction ≥ 50%). Melancholic features were assessed with both the Structured Clinical Interview for DSM-IV and baseline depression scale items, according to DSM criteria.

Results: The item-based melancholic specifier was associated with numerically larger treatment effects, although subgroup-treatment interactions in logistic regression models did not reach statistical significance. The small subgroup of patients with severe psychomotor retardation also appeared to benefit from lamotrigine. However, the Structured Clinical Interview for DSM-IV melancholic specifier was not associated with larger treatment effects. Baseline depression severity was inconsistently associated with response, depending on which scale was used to define severity. The 2 melancholia variables had poor agreement despite having similar prevalences.

Conclusions: Our results do not clearly support the original hypothesis but do reinforce the importance of replicating secondary analyses of clinical trials with additional data.

Key Words: anticonvulsant, endogenous depression, mood disorder, psychomotor retardation

(J Clin Psychopharmacol 2021;41: 315–319)

P reviously we hypothesized lamotrigine may be more effective for melancholic depression.[1] The rationale was that lamotrigine inhibits presynaptic glutamate release,[2] which appears to blunt the cortisol stress response[3] and mitigate the deleterious neuropsychological effects of chronic corticosteroid exposure,[4] whereas melancholic features are associated with hypothalamic-pituitary-adrenal (HPA) axis hyperactivity[5,6] and cognitive dysfunction,[7] possibly mediated by glutamatergic excitotoxicity.[6] Furthermore, lamotrigine appears to improve psychomotor retardation[8]—a cardinal melancholic sign.[9,10]

In a pilot study, we reanalyzed 1 of 5 industry-sponsored trials of lamotrigine for acute bipolar depression after using baseline scale items to split the sample into melancholic and nonmelancholic subgroups.[1] There was more separation of lamotrigine from placebo in the melancholic subgroup, but not all results were significant, and we were unable to adequately control for baseline severity.[1] The latter was a notable limitation as Geddes et al[11] had reanalyzed all 5 trials and found a larger effect in patients with baseline depression scores above the sample mean.[11] Here, we conducted a pooled analysis of all 5 trials to replicate the pilot results in a larger sample with a standard melancholia assessment, to consider models with melancholia and depression severity included together, and to explore whether baseline psychomotor retardation alone was associated with larger treatment effects.

MATERIALS AND METHODS

Data Acquisition

Data were provided by the sponsor via clinicalstudydatarequest. com. We obtained ethical approval and preregistered before accessing the data. Each trial is summarized online (see Supplemental Table 1 http://links.lww.com/JCP/A732). They were all randomized, double-blind, placebo-controlled, parallel-group, monotherapy trials.[12] All data were pooled, except for the 50 mg/d arm, following Geddes et al.[11]

Participants

The pooled sample consisted of 1072 adult outpatients 18 years or older with bipolar disorder I or II currently experiencing an acute major depressive episode. Notable exclusion criteria were as follows: prior lamotrigine treatment; active mania or suicidality; rapid cycling; social phobia, panic disorder, obsessive-compulsive disorder, or bulimia nervosa in the last year; substance abuse or dependence in the last month or year, respectively; pregnancy/ breastfeeding; other psychotropic medication; recent psychotherapy; thyroid disease; and epilepsy.[12]

Instruments

The Structured Clinical Interview for DSM-IV (SCID)[13] was used to confirm study eligibility. Depressive symptoms were measured at baseline and weekly with the 17- and 31-item versions of the Hamilton Depression Rating Scale (HAMD)[14,15] and the Montgomery–Åsberg Depression Rating Scale (MADRS).[16] We used all 3 scales because the HAMD-17 and MADRS overrepresent melancholic symptoms, whereas the HAMD-31 includes atypical symptoms (eg, increased sleep, increased appetite).[14–16]

Outcome Measures

The outcome measures were end-trial response rates (score reduction ≥ 50% from baseline) on the HAMD-17, HAMD-31,

and MADRS. Response rates were chosen over change scores because they are less susceptible to floor effects,[17] and we anticipated melancholic depression would be associated with higher baseline scores. We did not prioritize 1 scale as the primary outcome but were looking for consistency across scales.

Melancholic Features

We used 2 methods to define melancholic depression. The first was the SCID melancholic specifier. Second, we used baseline HAMD and MADRS item scores to split the sample into melancholic and nonmelancholic subgroups according to the DSM criteria. Because there is no accepted threshold for diagnosing melancholic symptoms with item scores, we planned to adjust the diagnostic algorithm so that the prevalence roughly approximated that returned by the SCID (see Supplemental Appendix A http://links.lww.com/JCP/A733). This was done to avoid having to rely on prespecified criteria that turned out to be inappropriately selective or inclusive. Psychomotor retardation alone was assessed with HAMD item 8.

Analysis

Missing HAMD and MADRS scores were imputed by carrying forward the last nonmissing score to the end of the trial, before calculating response rates. This was done to make the analysis comparable to Geddes et al and the original trials.[11,12] The analysis then proceeded in 4 sequential steps. First, the scale-derived melancholia variable was created and compared to the SCID variable. Second, treatment effects were estimated by calculating response-rate differences (ie, risk differences) between treatment groups, with Wilson-Newcombe confidence intervals and number-needed-to-treat (NNT) statistics. This was done stratified by melancholic status and baseline depression severity. Third, logistic regression was used to test for treatment effect heterogeneity across melancholic subgroups, if suggested by the response rates. Models were stratified by melancholic status, with treatment condition predicting response, and then repeated unstratified with a melancholia-by-treatment interaction. We also included baseline severity as a covariate in the models.

No other covariates (eg, age, sex) were considered. Fourth, the sequence above was repeated with psychomotor retardation as the subgroup variable instead of melancholic status.

RESULTS

Melancholic Subgroups

The process for creating a scale-derived melancholia variable is described in Supplemental Appendix A http://links.lww.com/JCP/A733. The final criteria are presented in Table 1, with baseline factors stratified by both variables. Melancholic status was associated with higher baseline depression scores, as expected, but not age, sex, or bipolar type. The 2 variables had poor agreement (61.1%; $\kappa = 0.22$) despite having similar prevalences (see Table 1).

Response Rates

Response rates stratified by melancholic status and baseline depression scale scores (dichotomized at the mean) are presented in Table 2. As a point of reference, for the entire sample, the response-rate differences were 8.4 (NNT = 11.9), 7.0 (NNT = 14.2), and 8.9 (NNT = 11.3) with the HAMD-17, HAMD-31, and MADRS, respectively. SCID melancholia was associated with lower response rates in both treatment conditions such that the response-rate differences were comparable across subgroups, similar to those in the entire sample. In contrast, although the scale-derived melancholic subgroup also had lower lamotrigine response rates, the placebo response rates were even lower, resulting in somewhat larger treatment effects (though less pronounced with the HAMD-31).

Patients with higher HAMD-17 scores had larger response-rate differences across scales, but the opposite pattern was found with severity defined by the HAMD-31 and MADRS. Although scale-derived melancholia and HAMD-17 severity were both associated with numerically larger response-rate differences, this benefit appeared to be largely confined to patients meeting both criteria (see Fig. 1). However, we did not test this further because it was not part of the analysis plan.

TABLE 1. Baseline Descriptive Statistics Stratified by Melancholic Status

Variables	Melancholia (SCID)		Melancholia (Scale-Derived)	
	Present	**Absent**	**Present**	**Absent**
n (%)	518 (48.3)	554 (51.7)	479 (44.7)	593 (55.3)
LTG, n (%)	253 (48.8)	288 (52.0)	243 (50.7)	298 (50.3)
Female, n (%)	302 (58.3)	324 (58.5)	279 (58.3)	347 (58.5)
Age, M (SD)	39.5 (12.1)	38.5 (11.7)	39.0 (11.6)	39.0 (12.2)
Bipolar I, n (%)	371 (71.6)	396 (71.5)	344 (71.8)	423 (71.3)
HAMD-17, M (SD)	25.3 (3.88)	23.5 (3.57)	26.0 (3.76)	23.1 (3.37)
	$F = 66.2^*, R^2 = 0.06$		$F = 181.0^*, R^2 = 0.14$	
HAMD-31, M (SD)	36.9 (6.52)	34.1 (6.35)	37.5 (6.65)	33.9 (6.07)
	$F = 50.8^*, R^2 = 0.05$		$F = 86.7^*, R^2 = 0.07$	
MADRS, M (SD)	30.8 (5.45)	28.3 (5.82)	33.0 (4.29)	26.6 (5.26)
	$F = 52.6^*, R^2 = 0.05$		$F = 449.4^*, R^2 = 0.30$	

Note. The criteria used to establish melancholic status with scale items were anhedonia (MADRS item 8 ≥ 4) or nonreactive mood (MADRS items 1 or 2 ≥ 5), and at least 3 of: psychomotor disturbance (HAMD items 8 or 9 ≥ 1), guilt (HAMD item 2 ≥ 1), late insomnia (HAMD item 6 ≥ 1), or appetite/weight loss (HAMD items 12 or 16 = 2). One-way ANOVA was used to test depression score differences by melancholic status.

LTG, lamotrigine; HAMD, Hamilton Depression Rating Scale (17- and 31-item versions); MADRS, Montgomery–Åsberg Depression Rating Scale.

* $P < 0.001$.

TABLE 2. Response Rates Stratified by Melancholic Status and Baseline Depression Scale Scores Dichotomized at the Sample Mean

Subgroup	HAMD-17				HAMD-31				MADRS			
	LTG	PBO	Diff (95% CI)	NNT	LTG	PBO	Diff (95% CI)	NNT	LTG	PBO	Diff (95% CI)	NNT
SCID-MEL +	42.3	34.0	8.3 (−0.01 to 16.5)	12.0	42.7	34.3	8.4 (−0.01 to 16.6)	12.0	44.3	36.6	7.7 (−0.8 to 16.0)	13.0
SCID-MEL −	47.2	39.1	8.1 (−0.1 to 16.2)	12.3	48.3	42.9	5.4 (−2.9 to 13.6)	18.5	51.4	41.7	9.7 (1.3 to 17.8)	10.3
Scale-MEL +	42.4	30.1	12.3 (3.7 to 20.6)	8.1	42.0	33.1	8.9 (0.3 to 17.4)	11.2	46.5	33.9	12.6 (3.8 to 21.1)	7.9
Scale-MEL −	47.0	41.7	5.3 (−2.7 to 13.2)	18.9	48.7	43.1	5.6 (−2.4 to 13.5)	17.8	49.3	43.4	5.9 (−2.1 to 13.8)	16.8
HAMD-17 ≥ 25	43.6	30.0	13.7 (5.0 to 22.0)	7.3	43.6	30.4	13.2 (4.6 to 21.6)	7.6	45.7	31.2	14.5 (5.8 to 22.8)	6.9
HAMD-17 ≤ 24	46.0	41.8	4.1 (−3.8 to 12.0)	24.2	47.3	45.2	2.1 (−5.9 to 10.0)	48.1	50.0	45.6	4.4 (−3.6 to 12.4)	22.6
HAMD-31 ≥ 36	39.2	36.3	3.0 (−5.5 to 11.3)	33.8	42.0	38.6	3.3 (−5.2 to 11.8)	30.2	42.4	38.2	4.1 (−4.4 to 12.5)	24.4
HAMD-31 ≤ 35	50.0	36.8	13.2 (5.0 to 21.1)	7.6	49.0	38.6	10.4 (2.2 to 18.4)	9.6	53.1	40.0	13.1 (4.9 to 21.1)	7.6
MADRS ≥30	40.5	34.7	5.9 (−2.1 to 13.7)	17.0	39.5	36.1	3.4 (−4.6 to 11.3)	29.5	45.4	37.2	8.1 (0.04 to 16.1)	12.3
MADRS ≤29	50.0	38.5	11.5 (2.8 to 19.9)	8.7	52.8	41.2	11.6 (2.9 to 20.0)	8.7	51.2	41.2	10.0 (1.3 to 18.4)	10.0

Note. Response rates and differences were proportions converted to percentages.

Diff, risk difference; MEL, melancholic status (present or absent); HAMD, Hamilton Depression Rating Scale (17- and 31-item versions); MADRS, Montgomery–Åsberg Depression Rating Scale; LTG, lamotrigine; PBO, placebo; Diff, risk difference; NNT, number needed to treat; SCID, Structured Clinical Interview for DSM-IV; MEL, melancholic status (present or absent); Scale, scale-derived variable.

Logistic Regressions

Proceeding with the scale-derived variable, in stratified models, the main effects of lamotrigine were statistically significant in the melancholic subgroup (HAMD-17, OR = 1.71, SE = 0.33, $P = 0.01$; HAMD-31, OR = 1.47, SE = 0.28, $P = 0.04$; MADRS, OR = 1.70, SE = 0.32, $P = 0.01$) but not in the nonmelancholic subgroup (HAMD-17, OR = 1.24, SE = 0.21, $P = 0.20$; HAMD-31, OR = 1.25, SE = 0.21, $P = 0.17$; MADRS, OR = 1.27, SE = 0.21, $P = 0.15$). In full models, the melancholic-treatment interactions did not reach statistical significance (HAMD-17, OR = 1.38, SE = 0.35, $P = 0.21$; HAMD-31, OR = 1.17, SE = 0.29, $P = 0.54$; MADRS, OR = 1.33, SE = 0.33, $P = 0.25$) and did not change after adjusting for baseline severity (HAMD-17, OR = 1.36, SE = 0.35, $P = 0.23$; HAMD-31, OR = 1.17, $SE = 0.29$, $P = 0.54$; MADRS, OR = 1.34, SE = 0.33, $P = 0.25$).

Psychomotor Retardation

Response rates over psychomotor retardation scores are presented in Supplemental Appendix B http://links.lww.com/JCP/A734. Response-rate differences tended to decrease slightly as scores increased from 0 (no retardation, n = 201) to 1 (slight retardation, n = 420) to 2 (obvious retardation, n = 400). Only at scores 3 or greater (interview difficult/stupor, n = 51) were the response-rate differences larger (10.6–18.5, NNT = 5.4–9.5) compared with the rest of the sample (6.78–8.30, NNT = 12.0–14.7). We did not use this cutoff score to conduct additional analyses because the subgroup was so small.

Sensitivity Analyses

These were conducted post hoc to test the robustness of the treatment effect in the scale-derived melancholic subgroup. It remained significant sequentially excluding data from each study (except on HAMD-31 response) and was comparable between bipolar types (see Supplemental Appendix B http://links.lww.com/JCP/A734). There was modest agreement between scale-derived melancholia at baseline and a week earlier at the screening visit (78.4%; κ = 0.56). The screening-visit variable also agreed poorly with the SCID (59.4%; κ = 0.19) and was associated with significant treatment effects (OR = 1.57–1.58, all $P = 0.02$). An alternate scale-derived variable was created with a different algorithm (see Supplemental Appendix A http://links.lww.com/JCP/A733) to see if results were sensitive to the method by which the original

variable was specified. It had a similar prevalence (47.7%) and poor agreement with the SCID (61.9%; κ = 0.24), and the response rates were nearly identical, with significant effects in the melancholic subgroup only (see Supplemental Appendix B http://links.lww.com/JCP/A734).

DISCUSSION

We were able to replicate the pilot study result in that treatment effects were numerically larger in patients with baseline scale item scores consonant with the DSM melancholic specifier, although the subgroup-treatment interactions did not reach statistical significance. Overall, the hypothesis was not clearly supported—melancholic depression diagnosed with the SCID, a standard measure, was not associated with larger treatment effects, and the scale-derived variable had poor agreement with the SCID. Moreover, there did not appear to be a monotonic relationship between psychomotor retardation and the response-rate difference, although the small subgroup with the highest scores seemed to benefit from lamotrigine over placebo.

Our interpretation is that the scale-derived variable selected patients who were less likely to respond to placebo, resulting in somewhat larger treatment effects, but they did not necessarily meet the SCID melancholic specifier, and we cannot know if they would have been classified as melancholic with another diagnostic system. Interestingly, the effect of baseline HAMD-17 severity on response rates, previously reported by Geddes et al,[11] seemed to be more pronounced in the scale-derived melancholic subgroup. Considering the other depression scales, higher baseline scores per se did not appear to result in larger treatment effects. Therefore, the present results have expanded upon the Geddes et al meta-analysis by suggesting certain scale items are more relevant in this regard.

Several points are pertinent for future research. First, it cannot be assumed that a melancholic diagnosis made post hoc with scale items will necessarily agree with the SCID. A notable difference between measures is that the SCID considers symptoms during the entire depressive episode (which may have resolved by the time of assessment) whereas the scales measure currently active symptoms.[13–16] Second, sum scores from different depression scales may not always be associated with treatment outcomes in the same way. A substantial amount of research has examined whether antidepressant effects are moderated by baseline HAMD-17 sum scores,[18] although this study supports the notion that some
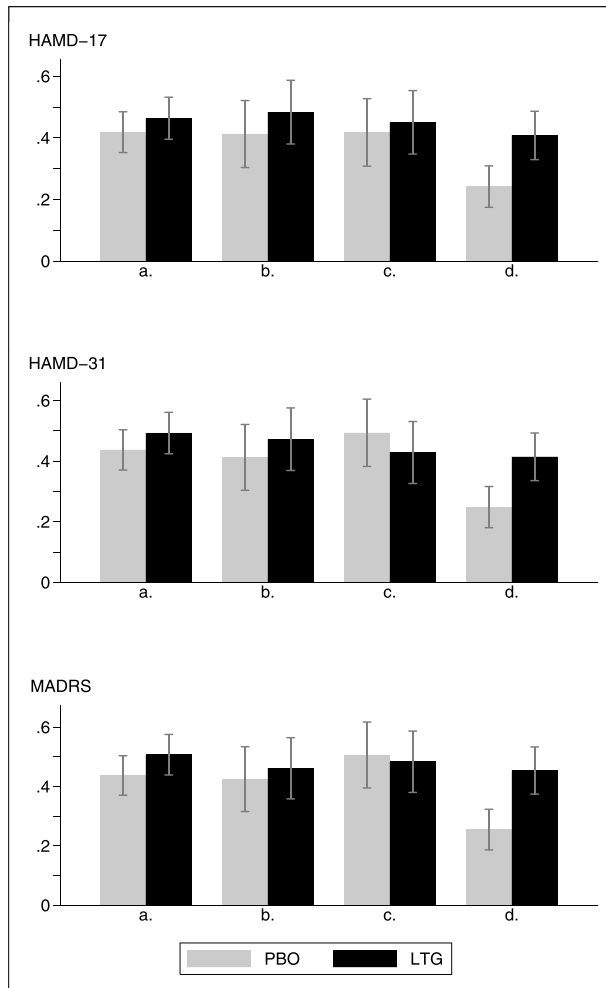
FIGURE 1. Response rates (y axis) are presented as proportions with 95% confidence intervals. These were calculated with 3 different rating scales: the HAMD-17, the HAMD-31, and the MADRS. Subgroups were (A) nonmelancholic and baseline HAMD-17 ≤ 24 (n = 422); (B) nonmelancholic and baseline HAMD-17 ≥ 25 (n = 171); (C) melancholic and baseline HAMD-17 ≤ 24 (n = 170); and (D) melancholic and baseline HAMD-17 ≥ 25 (n = 309). In the subgroup with scale-derived melancholia and HAMD-17 scores ≥ 25, the response-rate differences were between 16.6 (NNT = 6.0) and 19.9 (NNT = 5.0). PBO, placebo; LTG, lamotrigine.

items should be weighed more heavily.[19] In general, the results reinforce the importance of replicating secondary analyses of clinical trials with additional data, using a variety of methods, before drawing firm conclusions.

A major limitation is that there was only 1 standard assessment of melancholia that relied on the DSM criteria. Although common, these criteria have been criticized, often for over-diagnosing melancholia.[9,10] It is possible that other diagnostic measures would have produced different results. Psychomotor retardation was also assessed with a single item; this part of the analysis was exploratory and cannot speak to the importance of psychomotor retardation as a melancholic sign.[9,10] Another limitation is that serum lamotrigine levels were not available, so we cannot rule out the possibility that some patients were underdosed. It must also be stated that lamotrigine is not approved to treat acute bipolar depression. As with most industry-sponsored trials, the extensive list of exclusion criteria limits generalizability. It is also possible that differences

by melancholic status would have been more pronounced in an inpatient sample.[5] The hypothesis itself could be criticized because the rationale was that lamotrigine might benefit patients with HPA axis hyperactivity, but rather than testing this directly, we focused on a clinical diagnosis known to be associated with HPA axis hyperactivity. Having to infer the presence of biological abnormalities from clinical diagnoses, knowing these are unlikely to be present in every case, is a frustrating limitation that unfortunately also still pervades much of contemporary psychiatry.

REFERENCES

1. Peters EM, Bowen R, Balbuena L. Melancholic symptoms in bipolar II depression and responsiveness to lamotrigine in an exploratory pilot study. *J Clin Psychopharmacol.* 2018;38:509–512.

2. Ketter TA, Manji HK, Post RM. Potential mechanisms of action of lamotrigine in the treatment of bipolar disorders. *J Clin Psychopharmacol.* 2003;23:484–495.

3. Makatsori A, Duncko R, Moncek F, et al. Modulation of neuroendocrine response and non-verbal behavior during psychosocial stress in healthy volunteers by the glutamate release-inhibiting drug lamotrigine. *Neuroendocrinology.* 2004;79:34–42.

4. Brown ES, Wolfshohl J, Shad MU, et al. Attenuation of the effects of corticosteroids on declarative memory with lamotrigine. *Neuropsychopharmacology.* 2008;33:2376–2383.

5. Carroll BJ. The dexamethasone suppression test for melancholia. *Br J Psychiatry.* 1982;140:292–304.

6. Gold PW. The organization of the stress system and its dysregulation in depressive illness. *Mol Psychiatry.* 2015;20:32–47.

7. Bosaipo NB, Foss MP, Young AH, et al. Neuropsychological changes in melancholic and atypical depression: a systematic review. *Neurosci Biobehav Rev.* 2017;73:309–325.

8. Mitchell PB, Hadzi-Pavlovic D, Evoniuk G, et al. A factor analytic study in bipolar depression, and response to lamotrigine. *CNS Spectr.* 2013; 18:214–224.

9. Martino DJ, Szmulewicz AG, Valerio MP, et al. Melancholia: an attempt at definition based on a review of empirical data. *J Nerv Ment Dis.* 2019; 207:792–798.

10. Taylor MA, Fink M. Restoring melancholia in the classification of mood disorders. *J Affect Disord.* 2008;105:1–14.

11. Geddes JR, Calabrese JR, Goodwin GM. Lamotrigine for treatment of bipolar depression: independent meta-analysis and meta-regression of individual patient data from five randomised trials. *Br J Psychiatry.* 2009; 194:4–9.

12. Calabrese JR, Huffman RF, White RL, et al. Lamotrigine in the acute treatment of bipolar depression: results of five double-blind, placebo-controlled clinical trials. *Bipolar Disord*. 2008;10:323–333.

13. First M, Gibbon M, Spitzer RL, et al. *Structured clinical interview for DSM-IV Axis I Disorders Research Version (SCID-I)*. New York: Biometric Research, New York State Psychiatric Institute; 1996.

14. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56–62.

15. Williams J, Link M, Rosenthal N, et al. *Structured interview guide for the Hamilton Depression Rating Scale-Seasonal Affective Disorder Version (SIGH-SAD)*. New York: New York State Psychiatric Institute; 1988.

16. Montgomery SA, Åsberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134:382–389.

17. Vöhringer PA, Ghaemi SN. Solving the antidepressant efficacy question: effect sizes in major depressive disorder. *Clin Ther*. 2011;33: B49–B61.

18. Furukawa TA, Maruo K, Noma H, et al. Initial severity of major depression and efficacy of new generation antidepressants: individual participant data meta-analysis. *Acta Psychiatr Scand*. 2018;137:450–458.

19. Trivedi MH, South C, Jha MK, et al. A novel strategy to identify placebo responders: prediction index of clinical and biological markers in the EMBARC trial. *Psychother Psychosom*. 2018;87:285–295.