Metaforest Algorithm Insights: Predictors of Nocebo Response in ADHD

Mireia Porta[1]*, Maggie Barcheni[2]*, David Ramírez-Saco[1,3], Ruth Cunill[4], Magí Farré[2,5], Marc Saez[6, 7], Beatriz López[6], Xavier Castells[1]

* Mireia Porta and Maggie Barcheni contributed equally.

[1] TransLab Research Group; Department of Medical Sciences; Universitat de Girona, Girona, Spain

[2] Pharmacology, Toxicology and Therapeutics Department, Universitat Autònoma de Barcelona, Barcelona, Spain

[3] Department of Clinical Pharmacology, Vall d'Hebron Barcelona Hospital Campus, Barcelona, Spain

[4] Parc Sanitari Sant Joan de Déu-Numància, Parc Sanitari Sant Joan de Déu, Barcelona, Spain

[5] Department of Clinical Pharmacology, Hospital Universitari Germans Trias i Pujol, Badalona, Spain

[6] Research Group on Statistics, Econometrics and Health (GRECS), University of Girona, Spain

[7] CIBER of Epidemiology and Public Health (CIBERESP), Madrid, Spain

The authors declare no conflicts of interest.

Corresponding Author:

Maggie Barcheni

mbarshini@gmail.com

Pharmacology, Toxicology and Therapeutics Department, Universitat Autònoma de Barcelona, Barcelona, Spain

ABSTRACT

Background: Predicting the nocebo response in randomized controlled trials (RCTs) is crucial as it can help minimize its influence and improve the evaluation of the side effects of interventions for ADHD. The aim of this study is to determine the effect of covariates related to study design, intervention, and patients' characteristics on the nocebo response in patients with Attention Deficit Hyperactivity Disorder (ADHD) using Metaforest, and, ultimately, to investigate Metaforest's performance in predicting nocebo response in ADHD RCTs.

Methods: This study is a secondary analysis of a previously published systematic review [1]. Nocebo response was defined as the proportion of patients experiencing at least one AE while receiving placebo. We used Metaforest for investigating patient-, intervention, and study design-related nocebo response moderators in ADHD RCTs.

Results: One hundred and five studies were included. Overall, 55.4% of patients experienced at least one AE while receiving placebo. However, between-study variability on nocebo response was very high with nocebo response ranging from 4.2% to 90.2% leading to high statistical heterogeneity (I^2 = 88.3%). Older patients showed a higher nocebo response. The moderating effects of the year of publication, treatment length and gender were also significant. The predictive performance of the model was low-moderate ($R^2\ test$ = 0, 1922; MSE = 0, 0408)

Conclusions: Age was the most important nocebo response modifier, followed by year of publication, treatment length and gender. Metaforest lacked the capability to predict nocebo responses in future studies.\

Keywords: **Metaforest, Nocebo Response, ADHD,** placebo, heterogeneity

INTRODUCTION

Importance of the Study:

- Understanding Nocebo Response in ADHD
- Optimizing Clinical Trial Methodologies
- Identifying Influential Factors (age, publication year, treatment duration, gender)
- Addressing Wide Response Variability
- Challenges in Predicting Nocebo Responses
- Improving Patient Care and Outcomes

The Plain Language Summary (PLS):

This research focuses on understanding and predicting the nocebo response—when patients experience side effects from placebos—in individuals with Attention Deficit Hyperactivity Disorder (ADHD) participating in clinical trials. By analyzing various factors like study design, treatment specifics, and patient traits, the study employs a specialized approach called Metaforest to explore how these variables influence the nocebo response. The investigation encompasses 105 ADHD studies, revealing that 55.4% of patients encountered side effects while receiving placebos. However, the extent of variation across studies, ranging from 4.2% to 90.2%, leads to significant difficulty in drawing precise conclusions. The study identifies age as the primary influencer of the nocebo response, with older patients reporting more side effects. Additionally, publication year, treatment duration, and gender emerged as significant factors affecting the response. Although efforts were made to predict nocebo responses using Metaforest, the model's predictive accuracy was only moderately effective. To sum up, this research enhances our comprehension of how distinct elements shape nocebo responses in ADHD patients, highlighting age, publication year, treatment duration, and gender as key influencers, even though accurate prediction of these responses using the Metaforest method presents challenges.

1. Nocebo response

The nocebo response, akin to the placebo effect, is grounded in well-established theoretical mechanisms such as expectancy—preconceived beliefs that can elicit a response to an inactive product as anticipated—and conditioning—where previous exposure to an active product creates memories that can induce similar effects with an inert substance [1, 2]. Despite extensive focus on the placebo response, the nocebo response has not been as thoroughly examined in clinical trials and medical practice [3], understanding and mitigating the nocebo response are crucial for improving patient outcomes and treatment adherence. The current literature on the nocebo response in psychiatric disorders is scarce and tends to focus on depressive ones, describing significant nocebo response in randomized clinical trials (RCTs) [4], which compromises adherence and efficacy of current treatments in clinical practice [5]. There has been limited research on bipolar disorder [6], schizophrenia [7], or other mental disorders. The state of the art in Attention Deficit Hyperactivity Disorder (ADHD) is even less established. One meta-analysis [8] studied the nocebo response in ADHD but presented several limitations, namely: the nocebo response was a secondary outcome, since it was defined as the proportion of patient dropping out due to adverse events (AEs), its scope was restricted to moderate-severe AEs, and the sources of between-study variability in the nocebo response were not assessed.

The study by Ramírez-Saco et al. (2022) [1] was the first to primarily investigate nocebo response in ADHD patients. The covariates examined included age, treatment duration, method of AE collection (systematic vs. non-systematic), and type of ADHD treatment (stimulant vs. non-stimulant). These factors were used to explore the between-study variability in nocebo response. This study identified a notable nocebo response in ADHD, with 55.5% of patients in ADHD RCTs experiencing at least one adverse event while on placebo. This result aligns with findings in other psychiatric conditions such as depression [9, 4] and is somewhat lower than those reported in bipolar disorder [6] and schizophrenia [7]. It exceeds previous estimates of the nocebo response in ADHD [8].

Predicting the nocebo response in RCTs is essential to reduce its impact and enhance the assessment of side effects for ADHD interventions. Ramírez-Saco et al. (2022) [1] explored the

influence of covariates related to study design, interventions, and patient characteristics on the variability in nocebo response across studies. This investigation used meta-regression to provide regression coefficients that indicate the strength and direction of relationships between study-level variables and effect sizes. Positive correlations were identified with age, treatment duration, and methods for collecting adverse events, whereas type of drug and psychotherapy showed negative correlations. However, this analysis accounted for only 40% of the variability in nocebo response across ADHD studies, suggesting the need to explore alternative methods beyond meta-regression. Metaforest is a relatively new machine learning method that adapts the random forest algorithm to meta-analysis and enables investigating the sources of variation causing observed effects to differ from one RCT to another [10]. Metaforest combines the output of multiple decision trees within the context of meta-analysis to produce a single result. For classification tasks, the output of the random forest is the class selected by the majority of trees. For regression tasks, the mean or average prediction of the individual trees is returned. Random decision forests correct for the tendency of decision trees to overfit to their training set, and they generally outperform individual decision trees. Metaforest offers several advantages over meta-regression, the statistical method most commonly used to investigate the sources of between-study variability: 1) it is robust to overfitting, 2) it captures non-linear relationships, 3) it performs variable selection, reducing the chances of selecting irrelevant variables as relevant, and 4) it handles interactions between moderators. Although random forests lack the intuitive appeal of single trees, they provide a useful metric for interpreting the overall importance of moderators, or "variable importance". Metaforest has previously been used in climate research [11] and studies of early life stress in animal models [12], but its use in medicine is rare [13].

This study aims to determine the effect of covariates related to study design, intervention, and patients' characteristics on the nocebo response in patients with ADHD using Metaforest, and, ultimately, to investigate Metaforest's performance in predicting nocebo response in ADHD RCTs. Our study differs from Faraone et al. (2021) [8], which focused on patient dropouts due to AEs. We examine the broader occurrence of adverse events and employ a novel analytical approach using the Metaforest algorithm. This allows us to better identify patient and study characteristics influencing nocebo response, providing a deeper understanding of between-study variability.

METHODS

This article reports a secondary analysis of the systematic review conducted by Ramírez-Saco et al. (2022), which followed a pre-specified protocol registered in PROSPERO (CRD42021242733). The methodology applied aligns with the original protocol. [1].

Study Design and Inclusion / Exclusion Criteria

We performed a meta-analysis using the Metaforest technique on RCTs that evaluated the efficacy and safety of various pharmacological treatments for ADHD, as defined by the DSM-III-R, DSM-IV, DSM-IV-TR, or DSM-5, without age restrictions. Of the 350 studies originally identified, 105 trials were selected based on pre-specified criteria. We excluded withdrawal RCTs, studies with a placebo lead-in phase, and those focusing on non-core ADHD symptoms. The included studies had a double-blind phase lasting at least one week.**Source of Data**

Data was sourced from the Minerva Database on January 2, 2021. The Minerva Database [14] contains detailed information on all RCTs that have assessed the efficacy and safety of pharmacological treatments for ADHD. This database is updated weekly using data from Medline, CENTRAL, and PsycINFO. By January 2021, the Minerva Database included information from over 350 RCTs, documented in more than 700 scientific publications, regulatory documents, industry reports, and clinical trial registries. These RCTs were identified through systematic search methods. While the last search was completed in January 2021, the comprehensiveness of the database ensures that the data remains relevant and robust. Minerva captures approximately 2,250 variables per study, covering study design, patient and intervention characteristics, physiological data, and risk of bias assessed with the Cochrane Collaboration tool (The Cochrane Collaboration. Retrieved from https://handbook-5-1.cochrane.org/) [15]. The Cochrane Collaboration tool assesses the risk of bias for each study by categorizing it as "low" (indicating all domains have low risk), "high" (indicating at least one domain has high risk), or "unclear" (risk is uncertain). Previous studies have demonstrated the utility of the Minerva Database [16,17, 1].

Study authors and pharmaceutical companies were contacted via email to request any missing data. If some information remained missing despite these efforts, it was addressed using Multiple Imputation by Chained Equations (MICE) [18]. MICE is a technique for handling missing data by making assumptions about the missing data mechanism, such as data being missing at random.

Study of Variables

The dependent variable of interest was the nocebo response, defined as the proportion of placebo recipients who experienced adverse events (AEs). Various covariates were collected, including:The **mean age of participants** was reported for each randomized controlled trial (RCT), along with the **gender distribution**, focusing on the proportion of men. The **ethnicity** of participants was also documented, specifically highlighting the proportion of Caucasians in each study.

To assess the baseline condition of the participants, the **baseline ADHD severity** was measured using DSM-based ADHD rating scales. Additionally, the dataset indicated whether participants were **treatment-naive**, meaning they had not received any prior treatment for ADHD. The **type of drug** administered was classified as either a **psychostimulant** or a **non-stimulant**, and the **treatment regimen** was described as either **fixed** or **flexible dose**. The **duration of the intervention** in weeks was also recorded for each trial.

The use of **concomitant psychotherapy** was noted, identifying whether participants received any form of psychotherapy alongside their medication. The **legal status of the drug**—whether it was **FDA-approved** or not—was another key factor. Additionally, the number of **study sites** involved in each RCT was reported, and whether a **placebo lead-in phase** was included was also documented. The dataset also included the **probability of receiving placebo**, expressed as the percentage of patients assigned to placebo treatment relative to the total number of participants.

The **study design** was categorized as either **parallel** or **cross-over**, and the inclusion of participants with **comorbid conditions** was also recorded. The method of **adverse event collection** was described as either **open** or **systematic**, depending on the reporting process used in each study. Furthermore, the **sponsor** of the study was identified as either **commercial** or **non-commercial**, and the **year of publication** was noted to provide a temporal context.

Finally, the dataset indicated whether the trial included participants from the **USA** and classified the **risk of bias** in each study as either **high**, **low**, or **unclear**.

Study Procedures and Statistical Analysis

The nocebo response was calculated as the number of patients receiving placebo who experienced AEs during the double-blind phase divided by the number of patients allocated to placebo.

Before performing Metaforest, missing data were imputed. Afterwards, the presence of multicollinearity was examined using the generalized variance inflation factor [19].

Metaforest was performed following established methods [20]. First, the study database was randomly split into a training set (70%) and a test set (30%). The Metaforest model was trained using the "training dataset," and the accuracy of its predictions for nocebo response was evaluated using the "test dataset".

To expedite the computational process, the optimal number of trees for the Metaforest model was determined, ensuring convergence of the model.

To prevent overfitting, a pre-selection of moderators was conducted by means of Metaforest-recursive feature elimination algorithm. This method involved replicating the analysis 100 times, eliminating the moderator with the most negative importance in each replication, until a set of moderators with only positive importance was obtained. The importance of the moderators was assessed by randomly permuting the values of a moderator, nulling its relationship with the output, and evaluating its impact on the predictive performance of the model. Permutation importance considers the contribution of a variable individually, as well as in interactions with other variables. Therefore, if a variable has high permutation importance but does not appear to have a significant main effect on the outcome, it suggests that the variable may be influential in interactions with other variables.

The model's hyperparameters were adjusted experimentally. First, the number of candidate variables considered at each tree node for splitting the branches was varied from 2 to 10. Second, the minimum number of cases required to remain within a branch after splitting was also varied from 2 to 10. These two parameters already exist in Random Forest, but in Metaforest, a third parameter is introduced, which is the method for RCT weighing in the analysis. The weighing method can be uniform, fixed, or random. Optimal values for these hyperparameters were obtained from k-fold cross-validation over the training dataset [21]. In this study, 10-folds were considered. All possible combinations of parameter adjustments were tested, and the combination with the lowest root mean squared error (RMSE) value was selected. The hyperparameters tuning process was performed using the caret package.

The predictive performance of the model was evaluated using two indicators: R^2, and MSE. These indicators were calculated based on the predictions made by the trees using the test set.

RESULTS

Patient, Intervention and Study Characteristics

A total of 105 RCTs met the inclusion/exclusion criteria of this study, involving 8,743 patients who received placebo. Missing data imputation showed a similar distribution to the observed ones, and no multicollinearity among the covariates was found.

Patients had a mean age of 23.6 years and were mostly males and Caucasians. Overall, patients exhibited moderate-severe ADHD symptom severity at baseline.

Of the 105 studies, 41% focused on stimulant medications, while 59% investigated non-stimulants. Most studies investigated approved drugs for ADHD treatment, with over half of them employing a flexible dose regimen and not providing concomitant psychotherapy. The average duration of treatment was 9.0 weeks.

In terms of study characteristics, most studies were multicentric, did not employ a placebo lead-in phase prior to randomization, used systematic methods for collecting AEs such as questionnaires or checklists, and had a parallel design. On average, the probability of receiving placebo was 39.9%.

In relation with "other covariates", most studies were conducted in the USA and were published in 2006 or later. A little over a quarter of RCTs were considered to have a high risk of bias, with a high dropout rate amongst patients receiving placebo during the RCTs being the most common cause of bias.table 1.

Characteristic	**Mean (SD) or %**
Age	23.6 (12.1) years
Male	63.6%
Caucasian	70.2%
Children (≤12 years)	26 studies
Adolescents (13-17)	10 studies
Adults (≥18 years)	42 studies

Table 1: Patient, Intervention, and Study Design Characteristics

Nocebo Response Moderators' Analysis

Overall, 55.4% of patients experienced at least one AE while receiving placebo. However, between-study variability was very high with nocebo response ranging from 4.2% to 90.2% (Electronic Supplementary Material (ESM) 1) leading to high statistical heterogeneity (I^2 = 88.3%).

Metaforest analysis using the training dataset converged to approximately 2000 trees (ESM 2).

The analysis of the effect of covariates on nocebo response showed no influence of the number of sites, legal status of the drug, probability of receiving placebo, type of drug, sponsor, treatment regimen, psychotherapy, risk of bias, treatment naivety as an inclusion criterion, comorbidity as an inclusion criterion, and placebo lead-in (ESM 3). Based on this preselection, age, year of publication, method for collecting AEs, gender, ethnicity, study design, region, baseline ADHD severity, and treatment length remained as potential nocebo response moderators. While age, gender, region, and method for collecting AEs correlated positively with

nocebo response, study design and year of publication showed a negative association and baseline ADHD severity, ethnicity and treatment length showed other association. The predictive performance of the model was found to be low-moderate ($R^2\ test$ = 0, 1922; MSE = 0, 0408)

Considering the suboptimal predictive performance of the model and the primary goal of the study, which was to identify the impact of covariates on the nocebo response, we carried out a *post hoc* analysis employing the full dataset. This *post hoc* analysis adopted a more conventional statistical methodology, as opposed to a machine learning approach, with the principal focus on investigating the impact of covariates on the nocebo response, without taking into account the model's performance. This analysis converged at 1,000 trees (ESM 4) and revealed that nocebo response had the same associations as the analysis using training dataset (Figure 1).When compared to meta-regression (ESM 5), it was found that type of drug and psychotherapy did not moderate nocebo response, while study design, gender, ethnicity, baseline ADHD severity, publication date, and region did.

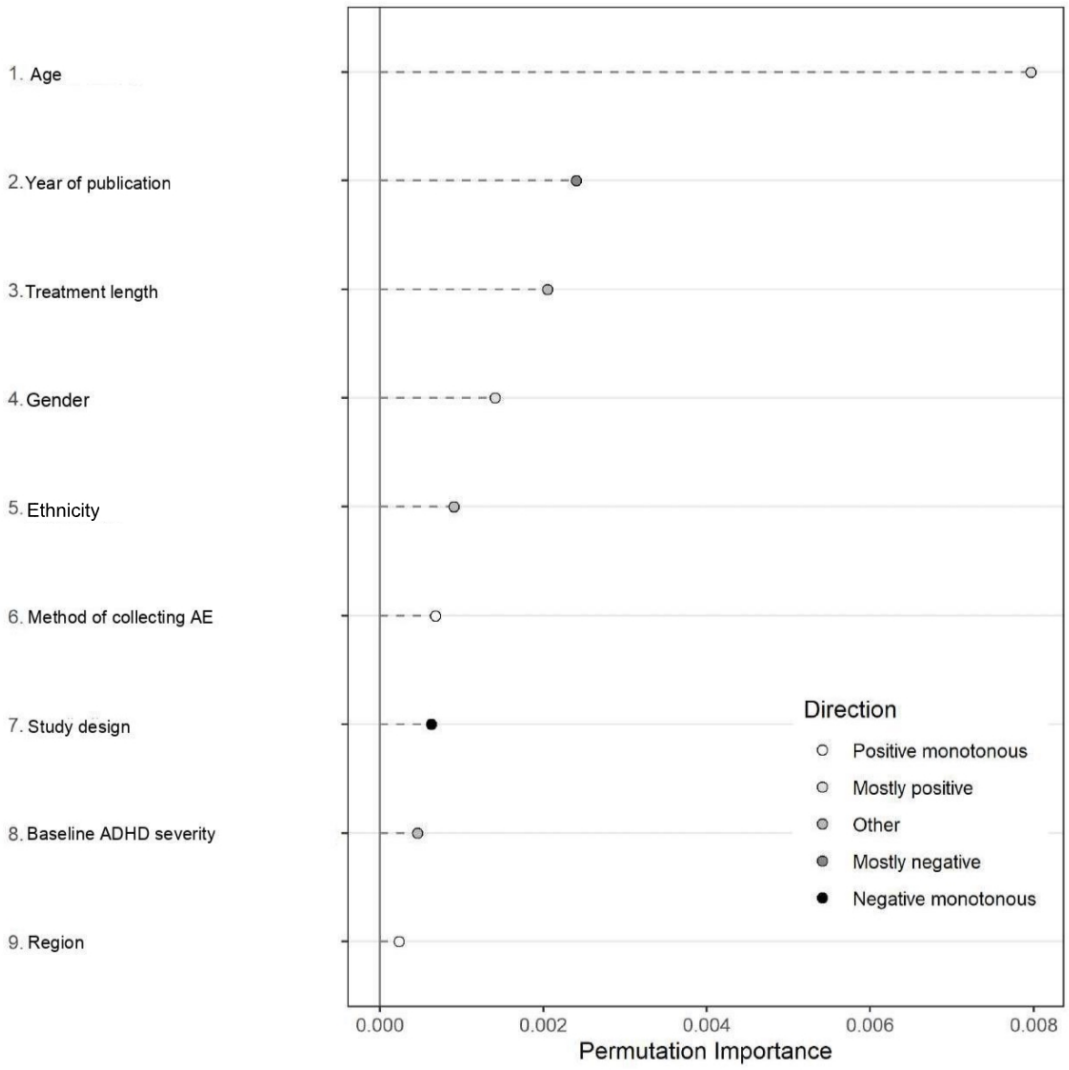


Figure 1: Summary of the variable importance and the direction of its effect on nocebo response using the entire dataset.

DISCUSSION

We investigated the effect of patient-, intervention-, and study design-related covariates on between-study variability in nocebo response in ADHD using Metaforest and evaluated the model's performance. A positive association was found with age, region, gender, and method for collecting AEs, while study design and year of publication had a negative association. Other type of association was found with treatment length, ethnicity and baseline ADHD severity. However, it was found that the constructed model exhibited inadequate performance, suggesting a lack of predictive capability. We repeated the analysis using the entire dataset. This *post hoc* analysis represents a more traditional statistical approach, focused on exploring the effects of covariates on the nocebo response in the analyzed dataset rather than on the model's generalizability to other data sets. This *post hoc* analysis yielded similar findings to the primary one.

The Metaforest analysis identified age as the single most important moderating variable. Older patients exhibited a higher nocebo response, which could potentially be because the expectation of experiencing AEs may rise with age. An alternative explanation could be that AEs may go unnoticed in children, emphasizing the importance of accurately recording AEs in this population.

The moderating effects of the year of publication, treatment length, and gender were also significant. It is worth noting that the year of publication has also been identified as an important moderator of the placebo response in ADHD [17] and in schizophrenia [22], depression [23], obsessive compulsive disorder [24], and bipolar disorder [25]. It could be that, as awareness about the safety of ADHD treatments has grown in the public eye has grown, patients' and clinicians' expectations of treatment safety may have increased, leading to reduced nocebo response over time. On the other hand, conflicting evidence exists regarding the influence of gender on placebo and nocebo responses in psychiatric and neurological diseases [26]: many meta-analyses detected higher nocebo response in women, while it was lower in other conditions such as neuropathic pain [27]. Metaforest found treatment length to have a non-linear relationship with nocebo response, which deserves further investigation.

The influence of ethnicity, method of collecting AEs, study design, baseline ADHD severity and region on nocebo response were of minor relevance. However, it must be emphasized that the method of collecting AEs emerged as a critical moderator of the nocebo response when meta-regression was used to analyze the between-study variability on nocebo response [1]. This finding underlines that Metaforest and meta-regression can produce different results even when used to analyze the same data set. A methodological implication can be inferred from our research: it may be advisable to use both methods to identify potential moderators in future studies. This could aid minimizing false positive associations by only considering as effect moderators those covariates found to be associated with the dependent outcome using both methods. An alternate approach might involve initially using Metaforest to select the covariates to be later included in the meta-regression analysis, as the latter method provides coefficients for each covariate, thus offering a quantitative explanation of the results.

Random forests provide a measure of explained variance analogous to R^2 in meta-regression. In our study, R^2=0,1922, while in (Ramírez-Saco et al., 2022) [1] study, R^2 =0.4. When comparing these two values we can conclude that meta-regression explained the variability better than Metaforest.

Limitations and strengths

Despite the large number of trials analyzed, given the fact that between-study variability is also high (I^2 = 88.3%) more trials are required to fully identify the sources of such variability. Unfortunately, there are no established methods for determining the necessary sample size for Metaforest analysis. To bolster statistical power, the analysis was conducted using the entire dataset; however, this approach primarily offers a descriptive overview of the current literature with limited applicability to future studies.

Additionally, while we have incorporated most covariates examined in other research, some significant factors may have been overlooked, such as illness duration [28], which is often not reported in ADHD trials. Ecological bias refers to the potential misinterpretation of results when conclusions about individuals are drawn from group-level data. In this study, the meta-forest analysis relies on aggregated data, meaning that the findings should be interpreted with caution, particularly when considering individual-level effects.[29]. Furthermore, certain categories (e.g., region or comorbidities) show skewed distributions. To address class imbalances, we consolidated studies into broader categories such as “Not US” or “comorbidity as an inclusion criterion.” This approach simplifies statistical analysis but results in a loss of detailed information. Fifthly, regarding the type of drug, non-stimulants comprise a heterogeneous group (multiple drugs whose common link is the absence of psychostimulant effects), and caution must be exercised when interpreting any findings.

The Metaforest model constructed failed to demonstrate sufficient predictive potential, thereby indicating its inability to generalize findings from training to testing datasets and subsequently undermining the prediction of nocebo responses in future ADHD RCTs. This could imply a low external validity of our study. It is essential to highlight that while most investigations into the moderating effects of covariates do not assess model performance, such testing should be deemed a standard practice as it provides insights into the external validity of the studies.

Despite these limitations, our study's strengths stand out. We pioneered the use of Metaforest to analyze between-study variability in nocebo responses, expanding the method's applicability. Our findings, distinct from those of meta-regression, underscore the value of employing both methods in future investigations.

Conclusions

In conclusion, our study investigated the impact of various patient-, intervention-, and study design-related covariates on between-study variability in nocebo response in ADHD using Metaforest, and evaluated the model’s performance. We found that age, region, gender, and method for collecting adverse events (AEs) had a positive association, while study design and year of publication had a negative association with nocebo response. Other variables such as treatment length, ethnicity, and baseline ADHD severity showed different types of associations. However, our model demonstrated inadequate performance, highlighting a lack of predictive capability. A post hoc analysis confirmed these findings, indicating that while Metaforest offers

valuable insights, it may not always outperform traditional meta-regression in explaining variability.

Age emerged as the most significant moderator, with older patients showing a higher nocebo response, possibly due to increased awareness of AEs. The significance of publication year, treatment length, and gender as moderators was also noted, aligning with findings in other psychiatric and neurological conditions. Despite the minor relevance of other factors, our study underscores the importance of accurately recording AEs, especially in children.

Our research highlights the methodological value of using both Metaforest and meta-regression to identify potential moderators, as each method may yield different results. This dual approach can help minimize false positives and provide a more comprehensive understanding of the factors influencing nocebo response.

While the main limitation was the insufficient number of included RCTs, affecting the statistical power, and potential ecological bias, the strengths of our study lie in its innovative application of Metaforest to explore between-study variability in nocebo responses. This study suggests that integrating both Metaforest and meta-regression analyses can enhance the robustness of future research in this field.

AUTHORS CONTRIBUTION

The authors confirm their contribution to the paper as follows: study conception and desig: MP, XC; data analysis and interpretation of results: DR, RC, MF, MS, BL; draft manuscript: MB. All authors reviewed the results and approved the final version of the manuscript.

CONSENT FOR PUBLICATION

Not applicable

CONFLICT OF INTEREST

The author(s) declare no conflict of interest, financial or otherwise.

FUNDING

None.

ACKNOWLEDGMENT

Declared none.

Electronic supplementary material

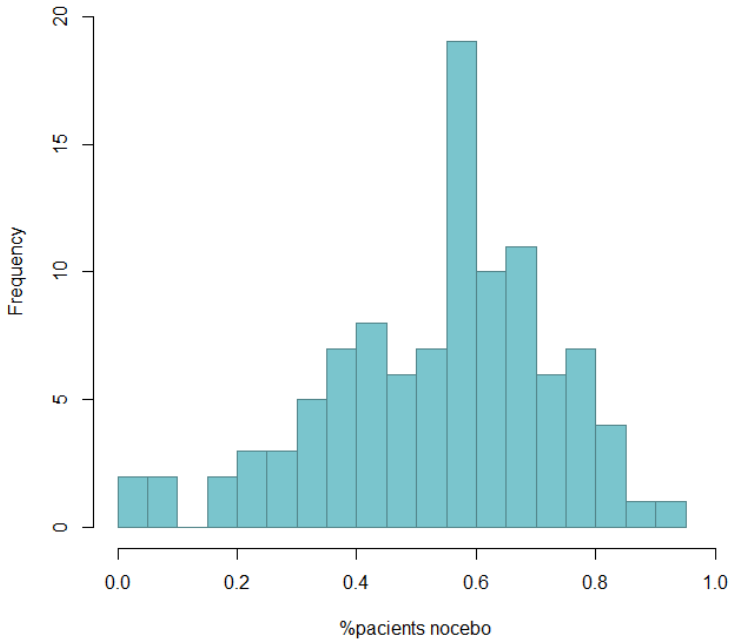
ESM 1. Nocebo response distribution

ESM 2. Convergence graph using the training dataset

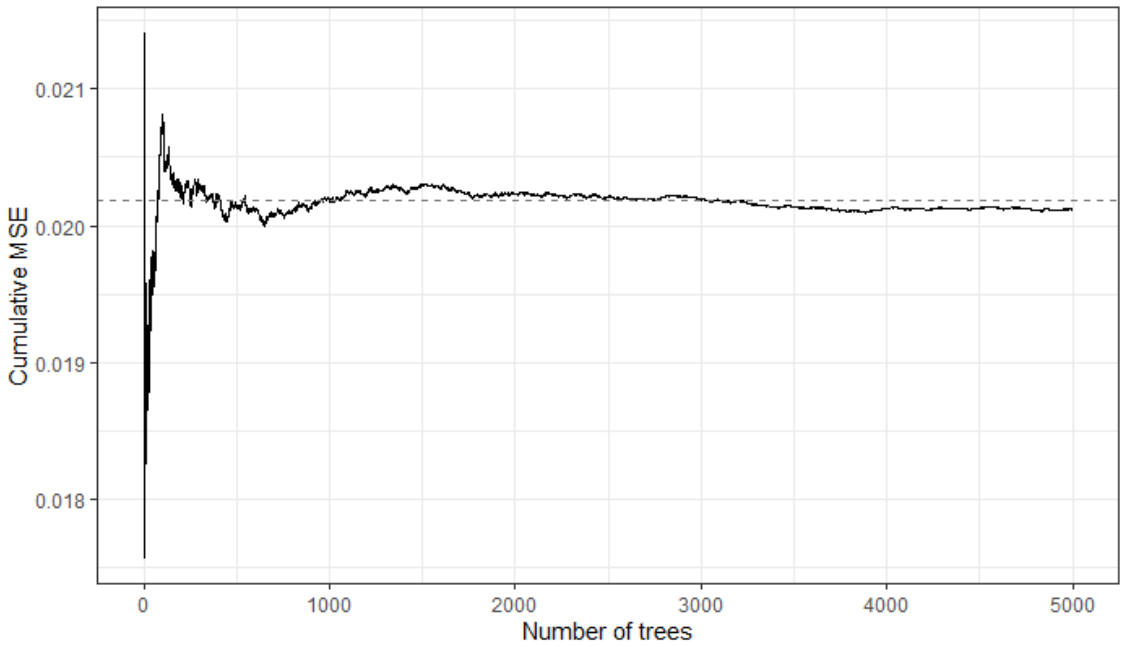
ESM 3: Variable importance using the training dataset

ESM 4. Convergence plot using the entire dataset

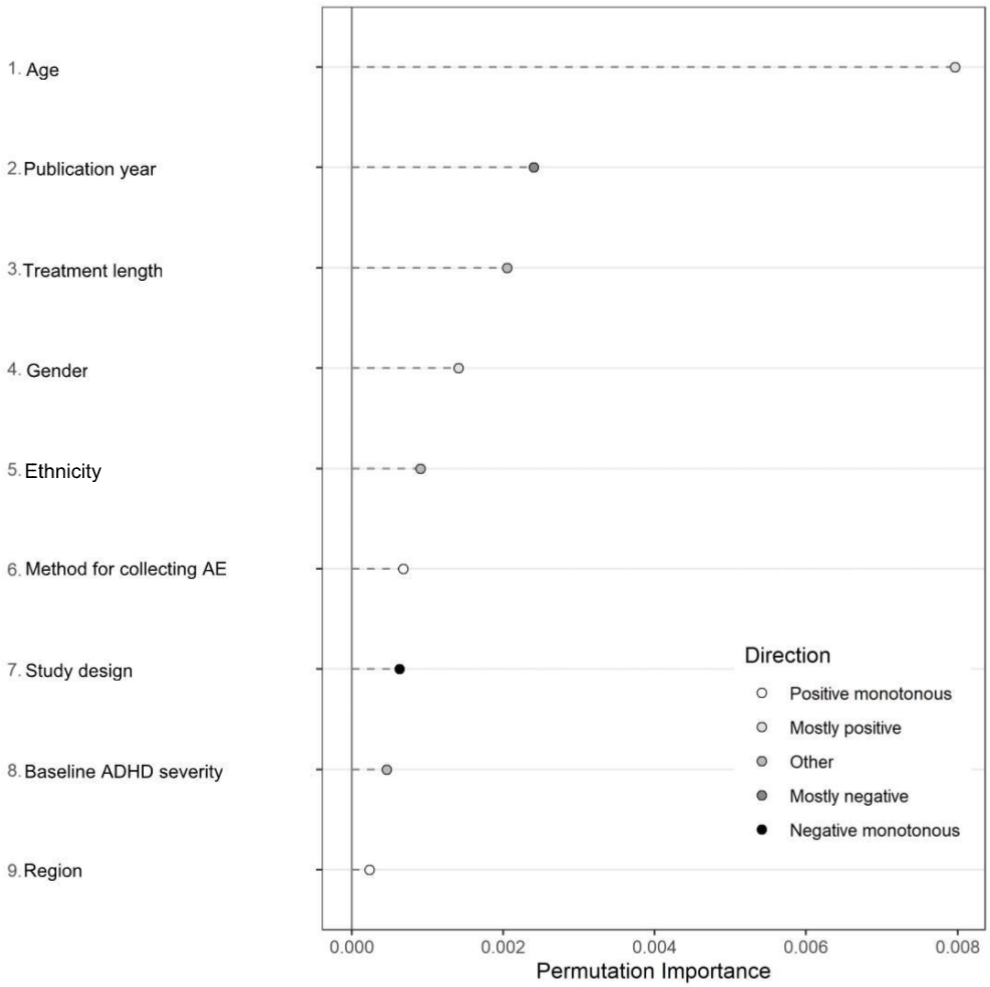
ESM 5. Covariates identified as nocebo response moderators by Metaforest and meta-regression



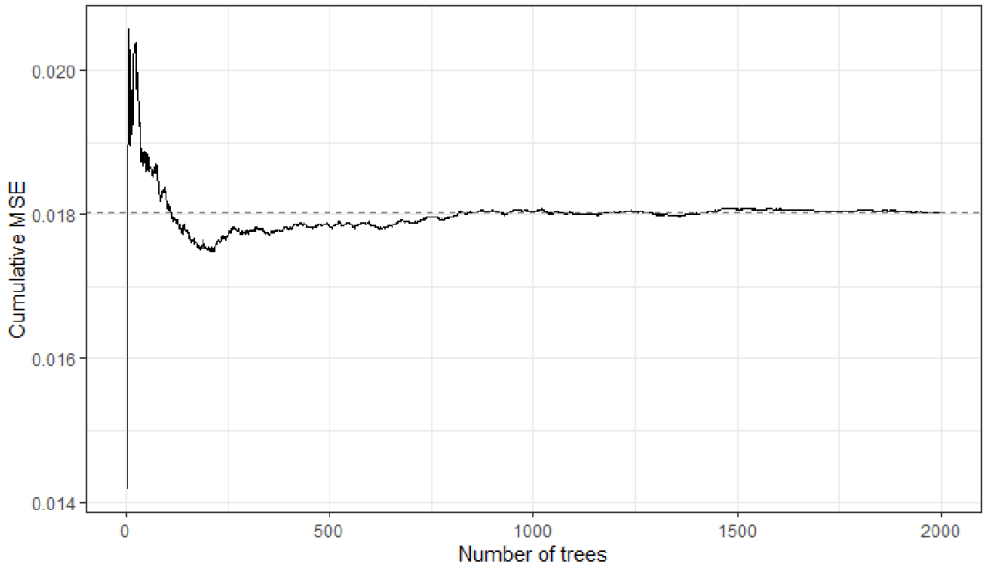
ESM 1. Nocebo response distribution



ESM 2. Convergence graph using the training dataset



ESM 3: Variable importance using the training dataset



ESM 4. Convergence plot using the entire dataset

	Metaforest	Meta-regression
Age	Yes	Yes
Gender (% men)	Yes	No
Ethnicity (% Caucasian)	Yes (other)	No
Baseline ADHD severity	Yes (other)	No
Naivety as inclusion criterion	No	No
Type of drug (psychostimulant)	No	Yes
Treatment regimen	No	No
Treatment length (weeks)	Yes (other)	Yes
Psychotherapy	No	Yes
Legal status of drug (approved for ADHD)	No	No
Number of study sites	No	No
Placebo lead-in phase	No	No
Probability of receiving placebo	No	No
Study design (parallel)	Yes	No
Comorbidity as inclusion criterion	No	No
Method for collecting AEs (proactive)	Yes	Yes
Sponsor (commercial)	No	No
Publication date (year)	Yes	No
Region	Yes	No
Risk of bias	No	No

ESM 5: Covariates identified as nocebo response moderators by Metaforest and meta-regression (from Ramírez-Saco et al., 2022)

REFERENCES

1. 1. Ramírez-Saco, D., Barcheni, M., Cunill, R., Sáez, M., Farré, M., & Castells, X. (2022). Nocebo Response in Attention Deficit Hyperactivity Disorder: Meta-Analysis and Meta-Regression of 105 Randomized Clinical Trials. *Journal of Attention Disorders*. https://doi.org/10.1177/10870547221075845

2. Benedetti, F., Lanotte, M., Lopiano, L., & Colloca, L. (2007). When words are painful: Unraveling the mechanisms of the nocebo effect. *Neuroscience*, *147*(2), 260–271. https://doi.org/10.1016/J.NEUROSCIENCE.2007.02.020

3. Benedetti, F., Frisaldi, E., & Piedimonte, A. (2019). The need to investigate nocebo effects in more detail. *World Psychiatry*. https://doi.org/https://doi.org/10.1002/wps.20627

4. Mitsikostas, D. D., Mantonakis, L., & Chalarakis, N. (2014). Nocebo in clinical trials for depression: A meta-analysis. *Psychiatry Research*, *215*(1), 82–86. https://doi.org/10.1016/J.PSYCHRES.2013.10.019

5. Rojas-Mirquez, J. C., Jose, M., Rodriguez-Zuñiga, M., Bonilla-Escobar, F. J., Andres Garcia-Perdomo, H., Petkov, M., Becerra, L., Borsook, D., & Linnman, C. (2014). *BEHAVIORAL NEUROSCIENCE Nocebo effect in randomized clinical trials of antidepressants in children and adolescents: systematic review and meta-analysis*. https://doi.org/10.3389/fnbeh.2014.00375

6. Dodd, S., J Walker, A., J M Brnabic, A., Hong, N., Burns, A., & Berk, M. (2019). Incidence and characteristics of the nocebo response from meta-analyses of the placebo arms of clinical trials of olanzapine for bipolar disorder. *Bipolar Disorders*.

7. Palermo, S., Giovannelli, F., Bartoli, M., & Amanzio, M. (2019). Are patients with schizophrenia spectrum disorders more prone to manifest nocebo-like-effects? A meta-analysis of adverse events in placebo groups of double-blind antipsychotic trials. In *Frontiers in Pharmacology* (Vol. 10). Frontiers Media S.A. https://doi.org/10.3389/fphar.2019.00502

8. Faraone, S. v., Newcorn, J. H., Cipriani, A., Brandeis, D., Kaiser, A., Hohmann, S., Haege, A., & Cortese, S. (2022). Placebo and nocebo responses in randomised, controlled trials of medications for ADHD: a systematic review and meta-analysis. In *Molecular Psychiatry* (Vol. 27, Issue 1, pp. 212–219). Springer Nature. https://doi.org/10.1038/s41380-021-01134-w

9. Meister, R., Jansen, A., Härter, M., Nestoriuc, Y., & Kriston, L. (2017). Placebo and nocebo reactions in randomized trials of pharmacological treatments for persistent depressive disorder. A meta-regression analysis. *Journal of Affective Disorders*, *215*, 288–298. https://doi.org/10.1016/J.JAD.2017.03.024

10. Van Lissa C (2017) MetaForest: Exploring heterogeneity in meta-analysis using random forests. :1– 60

11. Terrer C et al. (2019) Nitrogen and phosphorus constrain the CO2 fertilization of global plant biomass. Nat Clim Chang 9:684–689 Available at: https://doi.org/10.1038/s41558-019-0545-2

12. Bonapersona, V., Kentrop, J., van Lissa, C. J., van der Veen, R., Joëls, M., & Sarabdjitsingh, R. A. (2019). The behavioral phenotype of early life adversity: A 3-level meta-analysis of rodent studies. *Neuroscience & Biobehavioral Reviews*, *102*, 299–307. https://doi.org/10.1016/J.NEUBIOREV.2019.04.021

13. Castells X, Saez M, Barcheni M, Cunill R, Serrano D, López B, van Lissa CJ. Placebo Response and Its Predictors in Attention Deficit Hyperactivity Disorder: A Meta-Analysis and Comparison of Meta-Regression and MetaForest. *Int J Neuropsychopharmacol. 2022 Jan 12;25(1):26-35*. https:/doi: 10.1093/ijnp/pyab054.

14. Minerva Database (2020) Minerva Database. Minerva Database Available at: https://www.minervadatabase.org/en/ [Accessed August 15, 2020]

15. The Cochrane Collaboration. Retrieved from https://handbook-5-1.cochrane.org/

16. Castells, X., Baykova, E., Mayoral, S., Cunill, R., & Serrano, D. (2020). P.054 Gender bias in randomized, controlled trials of pharmacological interventions for attention deficit hyperactivity disorder. *European Neuropsychopharmacology*, *40*, S36–S37. https://doi.org/10.1016/J.EURONEURO.2020.09.052

17. Castells, X., Ramon, M., Cunill, R., Olivé, C., & Serrano, D. (2021). Relationship Between Treatment Duration and Efficacy of Pharmacological Treatment for ADHD: A Meta-Analysis and Meta-Regression of 87 Randomized Controlled Clinical Trials. *Journal of Attention Disorders*, *25*(10), 1352–1361. https://doi.org/10.1177/1087054720903372

18. Shah, A. D., Bartlett, J. W., Carpenter, J., Nicholas, O., & Hemingway, H. (2014). Practice of Epidemiology Comparison of Random Forest and Parametric Imputation Models for Imputing Missing Data Using MICE: A CALIBER Study. *American Journal of Epidemiology*, *179*(6), 764–774. https://doi.org/10.1093/aje/kwt312

19. Fox J, Monette G (1992) Generalized Collinearity Diagnostics. J Am Stat Assoc 87:178–183 Available at: http://www.jstor.org/stable/2290467

20. Van Lissa CJ (2020) Small sample meta-analyses: Exploring heterogeneity using metaforest. In: Small Sample Size Solutions: A Guide for Applied Researchers and Practitioners (Van De Schoot RM, ed). CRC Press.

21. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: Data mining, inference, and prediction. Berlin: Springer.

22. Leucht S, Chaimani A, Leucht C, Huhn M, Mavridis D, Helfer B, Samara M, Cipriani A, Geddes JR, Salanti G, Davis JM (2018) 60 years of placebo-controlled antipsychotic drug trials in acute schizophrenia: meta-regression of predictors of placebo response. Schizophr Res 201:315–323.

23. Undurraga J, Baldessarini RJ (2012) Randomized, placebo-controlled trials of antidepressants for acute major depression: thirty-year meta-analytic review. Neuropsychopharmacology 37:851–864.

24. Kotzalidis GD, Del Casale A, Simmaco M, Pancheri L, Brugnoli R, Paolini M, Gualtieri I, Ferracuti S, Savoja V, Cuomo I, De Chiara L, Mosca A, Sani G, Girardi P, Pompili M, Rapinesi C, On Behalf Of The Sapienza Group For The Study Of The Placebo Effect In Psychiatric Disorders (2018) Placebo effect in Obsessive-Compulsive Disorder (OCD). Placebo response and placebo responders in OCD: the trend over time. Curr Neuropharmacol 17:741–774.

25. Sysko R, Walsh BT. A systematic review of placebo response in studies of bipolar mania. J Clin Psychiatry. 2007 Aug;68(8):1213-7. doi: 10.4088/jcp.v68n0807

26. Enck P, Klosterhalfen S. Does Sex/Gender Play a Role in Placebo and Nocebo Effects? Conflicting Evidence From Clinical Trials and Experimental Studies. *Front Neurosci. 2019 Mar 4;13:160.* https://doi: 10.3389/fnins.2019.00160.

27. Papadopoulos D, Mitsikostas DD. A meta-analytic approach to estimating nocebo effects in neuropathic pain trials. *J Neurol. 2012 Mar;259(3):436-47*. https://doi: 10.1007/s00415-011-6197-4. Epub 2011 Aug 3.

28. Agid, O., Siu, C. O., Potkin, S. G., Kapur, S., Watsky, E., Vanderburg, D., Robert Zipursky, M. B., & Remington, G. (2013). Article Meta-Regression Analysis of Placebo Response in Antipsychotic Trials. In *Am J Psychiatry* (Vol. 170).

29. Greenland S, Morgenstern H (1989) Ecological bias, confounding, and effect modification. Int J Epidemiol 18:269–274 Available at: https://pubmed.ncbi.nlm.nih.gov/2656561/ [Accessed February 9, 2021].