MATTERS ARISING
https://doi.org/10.1038/s41587-020-0738-2

nature biotechnology

Check for updates

Reply to: EEG-based model and antidepressant response

Wei Wu[1,2,3], Diego A. Pizzagalli[4], Madhukar H. Trivedi[5,6] and Amit Etkin[1,2,3] ✉

REPLYING TO: G. Nilsonne & F. E. Harrell Jr *Nature Biotechnology* https://doi.org/10.1038/s41587-020-00768-5 (2020)

Recently, we published a paper reporting a pattern of brain activity as assessed by resting electroencephalography (EEG) that was specifically predictive of outcome with the antidepressant sertraline in comparison to placebo[1]. In response, Nilsonne and Harrell raise two criticisms in their Matters Arising[2]: first, that the baseline Hamilton Depression Rating Scale ($HAMD_{17}$) score was not adjusted for in the treatment outcome and, as a result, the sertraline signature from the EMBARC (Establishing Moderators and Biosignatures of Antidepressant Response in Clinical Care) trial might only be predicting the baseline $HAMD_{17}$ score and, second, that the EMBARC sertraline signature could only weakly predict the treatment outcome in the second depression dataset. In this reply, we provide a detailed response to address these criticisms, demonstrating that they are inaccurate owing to a misunderstanding of the original analyses or that the claims are not supported by the data. Collectively, we believe that these criticisms do not alter any conclusions in the paper.

With regard to the first critique, as reported in the original paper, SELSER (Sparse EEG Latent Space Regression) was not able to identify an EEG pattern related to baseline $HAMD_{17}$, but only the change in $HAMD_{17}$ associated with treatment. Moreover, extensive prior work has suggested that any prediction of outcome from baseline HAMD measures is controversial and of limited effect size[3,4]. Nonetheless, to rule out this possibility in our analyses, we followed precisely the approach outlined by Nilsonne and Harrell[2]. Namely, we used the restricted cubic spline fit approach[5] to regress out the baseline $HAMD_{17}$ score from the post-treatment $HAMD_{17}$ score for the sertraline arm in EMBARC. The MATLAB code[6] is as follows:

```
[bhat f sse1 knots] = rcspline(preHAMD, postHAMD, 'prc6');
adjustedPostHAMD = postHAMD – f(preHAMD);
```

We then correlated the EMBARC sertraline signature with the adjusted post-treatment $HAMD_{17}$ score. Specifically, we would expect a negative correlation because a greater change in EEG-predicted $HAMD_{17}$ score pre-treatment versus post-treatment should translate into a lower post-treatment $HAMD_{17}$ score. As expected, the correlation coefficient was −0.49 (Fig. 1), ruling out the possibility that the sertraline signature is predictive of the baseline $HAMD_{17}$ score only.

Moreover, although the restricted cubic spline fit approach is a useful technique to account for the baseline $HAMD_{17}$ score, it has its own limitations when used in an individual-level prediction context, which Nilsonne and Harrell[2] may have overlooked. Because the estimation of the spline fit is performed on the group level, each patient's outcome (that is, adjusted post-treatment $HAMD_{17}$ score) depends on the $HAMD_{17}$ scores of other patients included in the spline fit estimation. This is problematic, as the outcome for the same individual would vary across the loop of the cross-validation, depending on which individuals are included in the training set to estimate the spline fit function.

It is also important to note that the change in score post-treatment in comparison to pre-treatment is a very common endpoint accepted by the US Food and Drug Administration for approval of new medications or devices (see https://www.fda.gov/media/121348/download (the sponsor briefing document for brexanolone injection) and https://www.fda.gov/media/121376/download (the sponsor briefing document for esketamine)).

With regard to the second critique, centered on Fig. 4 in our paper, Nilsonne and Harrell[2] appear to misunderstand the intention of the analysis of the second depression dataset. Figure 4 in our paper examined whether the strength of the sertraline EEG signature relates to additional treatment response phenotypes, based on retrospective information about treatment response. This type of retrospective information, assessed on a standardized instrument, is the hallmark definition of treatment-resistant depression as a phenotype within the broader clinical category of depression. Thus, we were not predicting future depression response, and neither longitudinal treatment outcome data nor $HAMD_{17}$ scores were available in this dataset. As stated in the paper, patients in the second depression dataset were drawn from a naturalistic depression study. The goal was to test the generalizability of the sertraline EEG signature in depression, rather than a direct replication of the EMBARC study (as no such data are available in the field). As such, Nilsonne and Harrell's argument for a prediction analysis here is incorrect and derives from their misunderstanding, rather than any reflection on the predictive utility of the EEG signature.

The question of whether the EMBARC sertraline signature is related in some way to historical information on treatment responsiveness/resistance can be addressed by applying the EMBARC sertraline signature to the second depression dataset and then comparing the values of the signature in partial-response and treatment-resistant groups (defined on the basis of the Antidepressant Treatment Response Questionnaire (ATRQ)[7] as described in our paper[1]) under the classic hypothesis-testing framework. This is what we showed in Fig. 4 in our paper[1]. In other words, predictive accuracy is irrelevant for the nature of the comparison made and the scientific question addressed. We also never claimed

[1]Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA. [2]Wu Tsai Neuroscience Institute, Stanford University, Stanford, CA, USA. [3]Alto Neuroscience, Inc., Los Altos, CA, USA. [4]Department of Psychiatry, Harvard Medical School and McLean Hospital, Belmont, MA, USA. [5]Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX, USA. [6]O'Donnell Brain Institute, University of Texas Southwestern Medical Center, Dallas, TX, USA. ✉e-mail: amitetkin@altoneuroscience.com

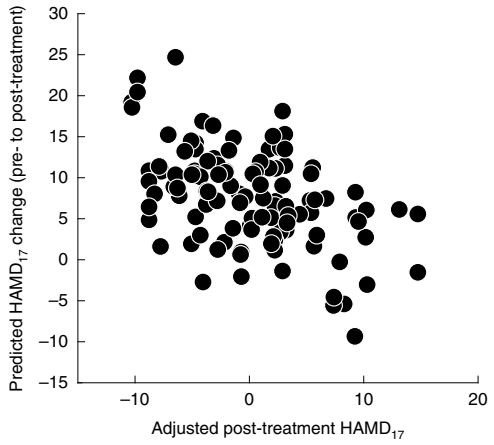

Fig. 1 | Scatterplot of the predicted change in $HAMD_{17}$ score and the adjusted post-treatment $HAMD_{17}$ score. The adjusted post-treatment $HAMD_{17}$ score was obtained by regressing out the baseline $HAMD_{17}$ score via the six-knot restricted cubic spline fit method.

that the EMBARC sertraline signature could distinguish between historically treatment-resistant patients and partial responders (defined on the basis of ATRQ score at the baseline visit) with high accuracy on an individual level for the second depression dataset.

It is likewise also important to point out two relevant findings that Nilsonne and Harrell[2] seem to overlook. First, we reported in the paper that in the historical response dataset there was a negative correlation between the number of failed trials and the magnitude of the $HAMD_{17}$ score improvement predicted by the EMBARC sertraline signature. This further underscores the clinical relevance of the EEG signature.

Second, we used a range of cross-validation methods to test the robustness of the EMBARC sertraline signature, including twofold, fourfold and tenfold cross-validation, as well as leave-study-site-out analysis. This last analysis is a strong demonstration that the findings are indeed robust for sertraline and specific for sertraline in comparison to placebo. Ultimately, as we also mentioned in the original paper, independent replication will be important, but EMBARC itself was a massive undertaking and took 10 years from study inception to this outcome. As such, collection of a future dataset will be no small feat.

Moreover, albeit indirectly, our paper provides convergent evidence across multiple datasets to support the generalizability of the SELSER-optimized predictive signature beyond EMBARC. This evidence includes the close relationship between the predictive signature and treatment responsiveness/resistance in a second depression sample and the convergence between the rsEEG and task-fMRI predictive signatures when applied to a third depression sample.

In summary, we have clarified the following: first, that the EMBARC sertraline signature is predictive of treatment outcome even after the baseline $HAMD_{17}$ score is regressed out of the post-treatment $HAMD_{17}$ score via the restricted cubic spline fit approach as requested by the Matters Arising[2] and, second, that the assertion that we were predicting responders versus non-responders in Fig. 4 of our paper is a misunderstanding of the intention of the analysis performed on the second depression dataset, which was to determine whether the EMBARC sertraline signature is related to historical information on treatment responsiveness/resistance.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-020-0738-2.

Received: 19 March 2020; Accepted: 16 October 2020;
Published online: 14 December 2020

References

1. Wu, W. et al. An electroencephalographic signature predicts antidepressant response in major depression. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-019-0397-3 (2020).
2. Nilsonne, G. & Harrell, F. Jr. EEG-based model and antidepressant response. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-020-00768-5 (2020).
3. Chekroud, A. M. et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. *Lancet Psychiatry* **3**, 243–250 (2016).
4. De Vries, Y. A. et al. Influence of baseline severity on antidepressant efficacy for anxiety disorders: meta-analysis and meta-regression. *Br. J. Psychiatry* **208**, 515–521 (2016).
5. Harrell Jr, F. E. & Slaughter, J. C. *Biostatistics for Biomedical Research* http://hbiostat.org/doc/bbr.pdf (2020).
6. Bantis, L. Restricted cubic spline. MATLAB Centreal File Exchange. https://www.mathworks.com/matlabcentral/fileexchange/41241-restricted-cubic-spline (2020).
7. Chandler, G. M. et al. Validation of the Massachusetts General Hospital antidepressant treatment history questionnaire (ATRQ). *CNS Neurosci. Therapeutics* **16**, 322–325 (2020).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020

Author contributions

W.W. and A.E. contributed to the analysis and interpretation of the data and the drafting and revision of the manuscript. D.A.P. and M.H.T. contributed to the drafting and revision of the manuscript.

Competing interests

W.W. receives salary and equity from Alto Neuroscience. Over the past 3 years, D.A.P. received consulting fees from Akili Interactive Labs, BlackThorn Therapeutics, Boehringer Ingelheim, Compass Pathway, Otsuka and Takeda Pharmaceuticals as well as an honorarium from Alkermes. In addition, he has received stock options from BlackThorn Therapeutics. No funding from these entities was used to support the current work, and all views expressed are solely those of the authors. M.H.T. has served as an adviser or consultant for Abbott Laboratories, Abdi Ibrahim, Akzo (Organon Pharmaceuticals), Alkermes, AstraZeneca, Axon Advisors, Bristol-Myers Squibb, Cephalon, Cerecor, CME Institute of Physicians, Concert Pharmaceuticals, Eli Lilly, Evotec, Fabre Kramer Pharmaceuticals, Forest Pharmaceuticals, GlaxoSmithKline, Janssen Global Services, Janssen Pharmaceutica Products, Johnson & Johnson PRD, Libby, Lundbeck, Meade Johnson, MedAvante, Medtronic, Merck, Mitsubishi Tanabe Pharma Development America, Naurex, Neuronetics, Otsuka Pharmaceuticals, Pamlab, Parke-Davis Pharmaceuticals, Pfizer, PgxHealth, Phoenix Marketing Solutions, Rexahn Pharmaceuticals, Ridge Diagnostics, Roche Products, Sepracor, Shire Development, Sierra, SK Life and Science, Sunovion, Takeda, Tal Medical/Puretech Venture, Targacept, Transcept, VantagePoint, Vivus and Wyeth-Ayerst Laboratories; he has received grants or research support from the Agency for Healthcare Research and Quality, Cyberonics, NARSAD, NIDA and NIMH. A.E. receives salary and equity from Alto Neuroscience and options for equity from Mindstrong Health, Akili Interactive Labs and Sizung for unrelated work.

Additional information

Correspondence and requests for materials should be addressed to A.E.

Reprints and permissions information is available at www.nature.com/reprints.