Correlations between patterns of activity and the response to treatment yield distinct signatures for different antidepressive treatments

Stefan Spulber[1], Frederik Elberling[1], Sandra Ceccatelli[1], Martin Gärde[2], Mikael Tiger[2], Johan Lundberg[2]

[1] - Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden.

[2] - Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institutet & Stockholm Health Care Services, Region Stockholm, Karolinska University Hospital, Stockholm, Sweden.

Abstract

Wrist actigraphy is a non-invasive technology to monitor activity over extended periods of time. Specific alterations in circadian patterns of activity have been described in several psychiatric disorders and are associated with symptom severity in major depressive disorder (MDD). The aim of this study was to investigate the correlations between activity patterns and the response to different MDD treatments. We used actigraphy recordings acquired for 3 independent clinical trials investigating the response to internet-delivered cognitive behavioral therapy (iCBT), escitalopram, or ketamine. We designed a feature extraction procedure applicable both to recordings spanning a single circadian cycle, and to average circadian profiles. First, we investigated the effects of MDD treatments independently from magnitude of response, and found that each treatment altered a specific subset of features. We next asked whether features of activity patterns before treatment correlate with the magnitude of response to MDD treatment, and whether the correlations are treatment-specific. We found distinct signatures in the correlations between longitudinal changes in individual features and the response to different treatments. Next, we trained linear regression ensembles to model the response to treatment using the patterns of activity during the depressive episode and ran in-depth analysis of coefficients to characterize the differences between treatments. We found that better response is associated with higher between-days variability for iCBT, and with lower between-days variability for ketamine, while lower within-day fragmentation was common to both treatments. Our data suggests that the analysis of circadian patterns of activity can potentially be used for predicting the response to different MDD treatments, however larger confirmatory studies are required for future clinical applications.

Introduction

Major depressive disorder (MDD) is a highly prevalent mental disorder with heterogenous biological background. There are many antidepressive interventions available, and all have comparable efficiency in reducing symptom severity (~50% at population level) [1]. Cognitive behavioral therapy (CBT) is well-established antidepressive non-pharmacological intervention [2]. Internet-delivered CBT (iCBT) has been shown to be as effective as face-to-face and has the advantage of using standardized treatment modules to ensure homogenous intervention suitable for longitudinal studies [3]. Antidepressant drugs act on specific neurotransmitter signaling pathways, including serotonin, dopamine, noradrenaline, and glutamate. Selective serotonin reuptake inhibitors (SSRI) are the most commonly prescribed antidepressants, and it typically takes several weeks for the antidepressive effects to stabilize [4]. The introduction of rapid acting antidepressants, such as ketamine in subanesthetic dose [5] lead to a paradigm shift in the treatment of depression [6]. The prediction of response to treatment is relevant particularly for the wellbeing of the patient as well as for avoiding unnecessary consumption of healthcare services. Several attempts have been described in the literature, including the change in reaction to facial expression early after starting the treatment [7], based on the patterns of change in scores for individual items after 4 weeks of treatment [8], or using combinations of genetic and non-genetic biomarkers [9]. While carrying undeniable advantages towards personalizing the MDD treatment, they either rely on treatment to be initiated before evaluation or require collecting biological samples for genotyping.

The recording of movement (actigraphy) provides a promising non-invasive technology for data collection, and accumulating evidence points to associations between specific alterations in activity and different psychiatric disorders [10–12]. For MDD, the most common alterations reported are lower total amount of activity and blunted circadian rhythms [13,14]. Similar associations have been described between activity levels and severity of depressive symptoms [15,16], and we have shown recently that the patterns of activity correlate with depression severity independent of actual levels of activity [17]. However, the potential correlations between activity patterns and the response to treatment have received less attention to date.

The aim of this study was to investigate the correlations between activity patterns and the response to different MDD treatment alternatives. To this end we designed a features extraction procedure including both sequence- and circadian profile-based features. We first explored the effects of different antidepressive treatments on circadian patterns of activity in

MDD subjects. Second, we explored the changes in patterns of activity in relation to the magnitude of response to treatment. Lastly, we trained multivariate linear regression ensembles to fit the response to treatment and analyzed the impact and implications of individual features.

Materials and methods

Data collection

We pooled together actigraphy recordings from 3 independent studies: (1) a study on serotonin transporter availability in patients given internet-based Cognitive Behavioral Therapy (iCBT) for the treatment of a major depressive episode [18] (Ethical Permit No. 2014/452-31, 2015/1177-32, Swedish Ethical Review Authority). Briefly, healthy controls (HC) and subjects suffering from depression were recruited by advertisements in local newspapers. Diagnosis of depression was established after full psychiatric assessment by a psychiatrist or by resident physician supervised by a senior psychiatrist. The study included patients with an ongoing major depressive episode according to DSM-IV criteria, with a history of at least one prior episode and that were not undergoing any psychopharmacological treatment for MDD. Eligible patients had a MADRS score between 18 and 35. HC had no history of psychiatric illness and matched the patients by age, sex, and intellectual ability [18]. (2) a study on the effects of escitalopram treatment on serotonin 1B receptor binding in MDD [Gärde M et al., manuscript in preparation] (Ethical Permit No. 2018/1403-31/1, Swedish Ethical Review Authority; pre-registration: AsPredicted.org #33267). (3) a study on the effects of ketamine on serotonin receptor 1B binding in patients with depression resistant to selective serotonin reuptake inhibitors (SSRI) [19] (Ethical Permit No. 2017/799-31, Swedish Ethical Review Authority; pre-registration: AsPredicted.org #17602). Briefly, the study included patients with an ongoing major depressive episode, with MADRS $\geq$ 20, resistant to SSRI treatment in an adequate dose for at least 4 weeks. Ongoing antidepressant treatment was discontinued and actigraphy data was collected after a washout period of at least 5 times the half-life of the SSRI. The patients were instructed to wear the actigraph continuously on the wrist of the non-dominant arm and not remove it unless for personal safety reasons. The recording started prior to the first ketamine infusion and continued for the duration of the open-label ketamine treatment program (0.5mg/kg, 2 infusions/week for 2 weeks) [19].
The timing of actigraphy recording and MADRS evaluations are depicted in Fig. 1A, and basic demographics, including the number of subjects and number of consecutive days of actigraphic recordings are depicted in Tables 1 and 2. Actigraphic recordings on patients

receiving iCBT as MDD treatment were acquired using GENEActiv Original wrist-worn actigraphs (ActivInsights, Cambs, UK). The devices use three-dimensional accelerometers (dynamic range up to 8*g*; 12-bit encoding, resolution 3.9 m*g*) at 30 Hz sampling rate to record wrist movement. The raw data was downloaded using proprietary software, then processed in Matlab® (The Mathworks, Natick, MD, USA), using a modified version of the code (https://github.com/DavidRConnell/geneactivReader), as described earlier [20]. Briefly, the Euclidean norm of change in acceleration vector was first smoothed using a rolling Gauss window spanning 30 consecutive datapoints (1s), then a high-pass filter was applied (threshold: 20 m*g* = 196 mm/s^2) before computing the sum of changes in acceleration vectors over 1 min epochs (1440 samples/24h). Actigraphic recordings on patients receiving escitalopram or ketamine as antidepressant treatment were acquired using Actiwatch 2 wrist-worn devices (Philips Respironics, Murrysville, PA, USA) set to record activity only integrated over 1 min epochs. The raw data was downloaded according to manufacturer's instructions (Actiware 6.x, Philips Respironics) then exported as text files. The text file was imported to Matlab® using a custom function designed to yield an output similar to the one generated by the import function for GENEActiv devices.

Quality control and inclusion criteria

The quality control was performed by the same observer, blind to group belonging. All recordings were first inspected visually using a standardized procedure designed to identify stretches of missing data, artifacts, and gross abnormal circadian patterns of activity (*e.g.*, shift work, or other consistent activity at night). Intervals containing suspected shift work (not reported at the time of recording), potential artifacts, or missing data, were cropped out. Only recordings spanning at least 3 consecutive days were included in further analyses.

Pre-processing and feature extraction

All processing of actigraphy data was performed in Matlab®. The data import procedures yielded sizeable differences in range of raw output (activity/min) between GENEActiv and Actiwatch 2 devices, but the coefficient of variability of individual days was similar across devices (see Supplementary Fig. S1). These differences were managed in the design of feature extraction procedures as follows: (1) no features relying on absolute magnitude values were included; and (2) features such as circadian peak and trough were calculated after mean normalization of raw data. We first cropped all recordings between the first and the last recorded midnight to yield an integer number of 24-h periods. For each subject we applied the feature extraction procedures on individual days and on circadian profiles calculated as the

average across corresponding time bins of registered individual days. The features calculated on individual days were used for estimation of variability across days.

The features we extracted describe the regularity, fragmentation, and complexity of circadian patterns of activity, and uses custom implementations of publicly available algorithms. The complete list of features and their calculations is available in Supplementary Table 1. The following features were extracted: circadian period; circadian peak and trough; relative amplitude [13,20,21]; scaling exponents [22]; intradaily variability (IV); interdaily stability (IS) [23]; day-to-day variability, and propensity to sustain activity.

Circadian period was estimated using the Lomb-Scargle algorithm optimized for Matlab implementation [24]. The Lomb-Scargle periodogram was preferred over the commonly used Sokolove-Bushell algorithm [25] because the latter has been shown to yield period estimates biased towards periods below 24 h [26]. The circadian period was calculated over the entire recording using an oversampling factor of 10 to yield resolution of the estimated in the range of minutes. The scaling exponent for detrended fluctuation analysis was calculated for the magnitude of measured activity in 1-min bins using boxes equally spaced on a logarithmic scale between 4 min (4 consecutive samples) and 24 h (1440 consecutive samples) as described by Hu et al. [22]. The scaling exponent is a feature of the intrinsic regulatory mechanisms controlling the rest/activity patterns. It has not been shown to be sensitive to extrinsic factors the subject is exposed to in normal daily activity, but is altered as a result of disease [10,22,27]. Intradaily variability estimates the fragmentation of activity patterns by calculating the ratio between mean squared differences between consecutive time intervals and the mean squared difference from global mean activity per interval; it increases as the frequency and the magnitude of transitions between rest and active intervals increase, and decreases as the active and inactive intervals consolidate [21]. Interdaily stability evaluates the coupling between activity patterns and circadian entrainers and is calculated as the ratio between variability around circadian profile and global variability. High values indicate consistent activity patterns across days, consistent with strong coupling between activity and circadian entrainers. The relative amplitude of circadian rhythms of activity (RA) estimates the robustness of average circadian rhythms [13,28]. The range of RA is bounded between 0 (no circadian rhythms) and 1 (robust circadian rhythms, with consistent timing of consolidated rest interval longer than 5 h across days). The day-to-day variability comprised 3 features as follows: circadian profile variance between consecutive days (SeqVar), calculated as Euclidean distance between consecutive days, normalized to the total number of samples per day; variation from average circadian profile (ProfileVar), calculated as the Euclidean

distance between each day and the average profile, normalized to the total number of samples per day; and the normalized difference between consecutive days (SeqProfileVar), calculated as the ratio between mean difference between circadian profiles of consecutive days and mean deviation from average circadian profile. The propensity to sustain activity (SustainProp) was calculated as the slope of likelihood to sustain or increase activity in the next minute against current level of activity. The distribution was calculated for minutes with activity count>10 (which eliminated the range specific for sleep, ~30% of datapoints/day) in 20 equally spaced bins covering the range up to the 99th percentile of active minutes. This is a hybrid measure applying a sequence-based analysis on the distribution of activity counts/min (assumed to have exponential distribution). The likelihood to further increase activity drops with increasing the activity counts in the current minute, therefore the slope is negative, and approaches 0 at the right tail of the distribution. A shallow slope indicates the subject is unlikely (not willing) to sustain even low levels of activity. For scrambled data (preserved distribution, but random sequence), the slope is about -0.3 (likelihood to sustain activity decreases by 30% when the activity count increases 10-fold).

Feature selection

There is virtually no consensus around the analysis of actigraphy and parameter selection for feature extraction. Therefore, the initial feature space included a number of features with different degrees of similarity, as illustrated by the matrix of correlations in Supplementary Fig. S2. To mitigate the impact of multicollinearity, one can choose between dimensionality reduction techniques (*e.g.*, PCA, UMAP), and feature selection. The former has the advantage of identifying latent structures in the feature space and accurately accounting for underlying patterns in downstream calculations, but the analysis of individual feature contribution is not straightforward. The latter facilitates the intuitive interpretation of the final result, but carries the risk of missing out on underlying latent structures. We opted for feature selection to take advantage of the interpretability of the results. This is particularly relevant for Bayesian model averaging, where posterior inclusion probability is artificially deflated by including highly correlated features. The initial feature space included 90 features, distributed in 4 overlapping clusters as follows: circadian profiles; mean daily features; daily variability around the average; and sequential daily variability. Day-to-day variability (calculated as the ratio between mean squared sequential differences in circadian profile and the mean squared difference from average profile) could only be calculated on the full recording and was assigned to the first cluster. We used a hybrid heuristic and data driven approach for feature

selection, namely we focused on reducing the clusters of conceptually related features which were highly correlated in healthy controls and MDD patients before treatment (see Supplementary Fig. S3). The correlation patterns yielded by features calculated on circadian profiles largely matched the correlations among features calculated on single days. In addition, the correlations between variability around the average and sequential variability were virtually perfect, rendering the 2 sets redundant. The final feature space consisted of 28 features, including 12 from the circadian profile cluster and their correspondents in daily variability around the average; propensity to sustain activity and its variance across days; day-to-day variability of circadian profile; and subject age (Supplementary Table 2).

Effects of antidepressive treatments and correlations with response to treatment

The design of analysis uses several underlying assumptions. (1) The effect of antidepressant treatment is assumed to have reached steady-state at the time of recording, and no feature exhibits significant drifting. This was verified by inspection of trends in individual features for each subject (not shown). (2) MADRS evaluation reflects an average state over spanning the entire recording period. The self-assessment MADRS scale requires the subject to evaluate the severity of symptoms over the last 3 days, and it is therefore justifiable to assume irrelevant changes over a period of 5-7 days around the time of evaluation. (3) The pharmacological effects are stable on both mood and circadian regulation of activity. Classical antidepressive treatments take a couple of weeks to reach full effect on mood (*i.e.*, iCBT and escitalopram), which is then stable for sufficiently long time for the purpose of actigraphy recording. Fast-acting antidepressants (*i.e.*, ketamine), induce measurable effects on mood within a couple of hours after administration. The regimen of ketamine administration in the cohort analyzed here was designed to ensure stable drug effect covering 2 weeks in the open-label phase of the study.

The response to treatment was calculated as relative change from baseline MADRS. The MADRS scores used for the estimation of response to treatment were the last measurement available before starting the treatment (baseline), and the last measurement available after treatment initiation (see also Fig. 1C, D for details).

Ensemble training and Bayesian model averaging

We implemented an ensemble approach to fit the response to treatment. The training was performed independently for 3 groups: iCBT, ketamine, and any treatment (iCBT, ketamine, or escitalopram). We have generated the initial ensemble by independent homogenous training using a systematic bootstrapping (with replacement) scheme for selecting up to 6

features/model (systematic testing of all possible combinations of features). This approach ensured a minimum of 2 subjects/feature for the iCBT cohort (N=12). The number of subjects available in the escitalopram cohort (N=7) was too low to train meaningful models, therefore the response to escitalopram was not modeled independently. After training, we applied an adaptation of Occam's window algorithm [29,30]. First, we included all models satisfying the following criteria: VIF<5 (to avoid multicollinearity issues); and adjusted R-square>0.2 (adjusted R-square>0.1 for "any treatment" ensemble). Next, we excluded all models receiving less support from the data than their simpler submodels. More complex models (using *n* features) which are less accurate than other any of the less complex models trained on the same subset of features (using any combination of *n-1* features) were excluded. This procedure penalizes more complex models if their accuracy is not superior to simpler models. The pruned ensembles were then sorted by increasing RMSE (decreasing accuracy) and the aggregated output of each ensemble was calculated as the cumulative Bayesian average of individual models for each subject, using the inverse model RMSE as weights [29].
We applied a Bayesian approach for the analysis of coefficients in the pruned ensembles. First, we estimated the prior inclusion probability for an individual feature as the proportion of models including a specific feature relative to the total number of models possible to train under the constraint of maximum 6 features/model to 0.2035 (see Supplementary Material). The posterior inclusion probability (PIP) for each feature was calculated as the frequency of occurrence in the pruned ensembles, and was used for defining levels of evidence strength as follows: PIP>0.2035 identifies features with frequency of occurrence increased as compared to initial probability (enriched), and indicates medium to strong evidence of correlation; in contrast, PIP < prior inclusion probability denotes features depleted after pruning, and indicates weak evidence of correlation. The effect size for individual features was estimated as the average of standardized coefficients across all models in the pruned ensemble. Lastly, we calculated the coefficient of variation (CV=standard deviation/average) for standardized coefficients to estimate the stability of individual features (context-independence).

Results

We have focused our analyses mostly on visualizing differences and correlations within the group of MDD subjects, and this approach is justified by the fact that MDD is a disease with largely unknown etiology and heterogenous manifestations: according to the diagnostic criteria listed by DSM-5 [31], 227 possible combinations of symptoms can lead to a

diagnostic of MDD. Preliminary analyses did not identify significant differences between male and female subjects in individual features, symptom severity, or response to MDD treatment. The data available in the three cohorts allowed partitioning in several ways: the data collected from patients before treatment were pooled into a single group (MDD, N=35), which was used as reference for all subsequent comparisons involving subjects on antidepressive treatments; similarly, all subjects with MADRSpost ≤ 10 were assigned to a single group "remission" (N=12, including iCBT, N=3; escitalopram, N=2; ketamine, N=7).

Individual features are differentially affected by MDD treatments

First, we asked whether MDD treatments had any effects on activity patterns, independently from the magnitude of response. This approach assumed that antidepressant drugs may impact circadian regulation of activity independent of antidepressant effects (reviewed in [32,33]). We did not find significant differences between MDD patients before treatment and healthy controls (Fig. 2A). Individual MDD treatments displayed significant effects on a wide range of features, with distinct, non-overlapping signatures across treatments (Fig. 2A). The effects can be summarized as follows: following iCBT, the subjects display increased daily variability in circadian profile. Under escitalopram treatment, the patterns of activity are less fragmented and more consistent across days, and the subjects display a higher propensity to sustain activity as compared to untreated MDD subjects. Under ketamine treatment, the patterns of activity are more fragmented and variable across days, but less complex as compared to untreated MDD subjects. For subjects in remission, regardless of the treatment, we found no difference from MDD or healthy controls.

Next, we asked whether the magnitude of response to treatment correlated with longitudinal changes in individual features of activity (Fig. 2B). In subjects treated by iCBT, the magnitude of response to treatment did not correlate with changes in individual features, while drug treatments displayed sparse significant correlations with individual features. In subjects treated with escitalopram, better response correlated with increase in circadian peak of activity, decrease in fragmentation of activity, and decrease in propensity to sustain activity. The response to ketamine treatment correlated with a shift in location of circadian peak of activity towards earlier occurrence. Better response to treatment in general correlated with a shift in location of circadian peak of activity towards earlier occurrence, and a stabilization of circadian rhythms (illustrated by decreased daily variability of IS30).

MDD treatments have distinct signatures in correlations between magnitude of response and individual features

We asked whether features of activity patterns before treatment correlate with the magnitude of response to treatment, and whether the correlations are treatment-specific.

When analyzed separately, we found distinct signatures for each treatment (Fig. 3A). For iCBT, better response was associated with higher age, steeper propensity to sustain activity, higher variability in location of circadian trough, and lower and less variable stability of circadian patterns of activity. For escitalopram we did not find correlation surviving the FDR correction, but we observed a trend in correlation between response to treatment and lower variability in fragmentation of activity patterns. Better response to ketamine was associated with decreased variability in magnitude of circadian peak of activity. When all cohorts were pooled, better response to treatment was associated with more stable patterns of activity, as illustrated by decreased variability in location of circadian peak as well as in scaling exponent (Fig. 3A).

We trained 3 independent ensembles of multiple linear regression models to fit the response to treatment for iCBT and ketamine separately, and used the pooled dataset for fitting the response to treatment irrespective of treatment group ("any treatment"). The procedure yielded 41768, 18178, and 23223 models for iCBT, ketamine, and any treatment, respectively (Fig. 3B). To analyze the performance of each ensemble, we sorted the models by decreasing fitting accuracy (increasing RMSE) and calculated the accuracy of ensemble fitting (aggregated output) in a cumulative fashion (Fig. 3C). The aggregated accuracy of all ensembles outperformed all independent models, and the accuracy changed with increasing the number of models included in a non-monotonic manner: we observed an initial improvement in accuracy, followed by degradation with different rates across ensembles (0.8 vs. 0.2 for iCBT and 0.7 for ketamine) (Fig. 3C). Notably, the ensemble fitting the response to any treatment had an accuracy considerably lower than the ensembles fitting the response to either iCBT or ketamine. This is illustrated by the minimum and maximum aggregated RMSE (0.1377 and 0.1588 vs. 0.0085 and 0.0496 for iCBT and 0.0572 and 0.1092 for ketamine), as well as by the slope of cumulative aggregated accuracy (Fig. 3C). For a MADRS score of 30 before treatment, this translates into an expected error in prediction of MADRS score after treatment between 0.25 and 1.5 for iCBT; 1.7 and 3.3 for ketamine; and 4.1 and 4.8 for any treatment.

To probe the biological relevance of modeling the response to treatment, we performed a Bayesian analysis of coefficients. The contribution of individual features in ensembles presumably changes depending on the context (*i.e.*, all other features included in the model), and the importance of variables in the model can be assessed using several approaches [34]. We described the contribution of individual features using the following parameters: PIP (*i.e.*, frequency of inclusion in the pruned ensemble); mean standardized coefficient (to estimate effect magnitude); and coefficient of variance (to estimate the context-dependence) (Fig. 4A; see also Supplementary Fig. S4). These three measurements are not necessarily orthogonal, but their joint analysis highlights the most relevant features and supports the interpretation of correlation coefficients. For each ensemble we found enriched features with low variability, with partial overlap across ensembles (Fig. 4B). The mean standardized coefficients for individual features varied considerably across ensembles, and filtering the feature space by CV or PIP yielded distinct signatures across ensembles (Fig. 4B; see also Supplementary Fig. S5). We further explored the differences between the ensembles modeling the response to iCBT or ketamine by plotting the standardized coefficients for individual features (Fig. 4C). The distribution of enriched features across the quadrants highlighted the distinct contributions of relevant features to modeling the response to treatment. Two out of 5 features enriched in both ensembles displayed opposite correlations (*i.e.*, the variability in location of circadian trough and the variability in circadian profile), while the rest had similar correlations (negative correlation for age and variability in circadian peak amplitude, and positive correlation for propensity to sustain activity). In addition, the differentially enriched features also appeared to have large effects only in the ensemble they were enriched in. Given the bias in the selection of subjects for the ketamine study (insufficient response to SSRI [19]), we investigated the potential differences between the two populations, and found that variability in circadian profile was higher, and the scaling exponent for short intervals was lower in ketamine-treated subjects than in iCBT-treated subjects. To summarize, the differentially enriched features suggested the following correlations: better response to iCBT correlated with weaker circadian entrainment and earlier circadian peak of activity; better response to ketamine correlated with more fragmented and less variable circadian patterns of activity, as well as shorter circadian period. In addition, better response to either iCBT or ketamine correlated with increasing age and steeper propensity to sustain activity. The features enriched in the ensemble fitting the response to any treatment were largely a combination of the features enriched for fitting the response to iCBT or ketamine, while age, propensity to sustain activity and variability in circadian peak amplitude had similar

correlations with the response to any treatment as for iCBT and ketamine (Supplementary Fig. S5B).

Discussion

We focused on developing a sound pipeline for data pre-processing, feature extraction and feature selection for actigraphy as objective measurement during a depressive episode. We found sparse correlations between the response to specific MDD treatment and individual features, and we observed distinct signatures in the patterns of correlations for specific treatments. We modeled the response to treatment using the features of circadian patterns of activity before treatment as input to multiple linear regression ensembles and applied Bayesian model averaging to account for model uncertainty. The aggregated performance varied widely among ensembles, the accuracy observed for iCBT and ketamine being considerably higher than for "any treatment". The Bayesian analysis of coefficients yielded distinct correlates of response to treatment for iCBT and ketamine.

MDD is a disease with largely unknown etiology and heterogenous manifestations, therefore comparisons against healthy controls may yield biased results due to restrictive selection criteria. Remarkably, a large population-based study found that circadian rhythm features (relative amplitude of activity intensity or skin temperature) do not display sex-dependent differences, and have small effect on predicting the diagnosis for mood disorders [35]. Similarly, we did not find sex-related differences between groups, nor did we find significant differences between MDD subjects and healthy controls. Therefore we pooled male and female subjects and focused instead on within-population effects in the MDD group. The ketamine cohort is fundamentally different from the iCBT and escitalopram cohorts because the lack of satisfactory response to ongoing SSRI treatment was a selection criterion [19]. We found only sparse differences among cohorts before treatment, which suggests that the overall heterogeneity of the MDD population obscures the differences between subgroups. As expected, we found distinct signatures for each treatment which may be attributed to the pharmacodynamic differences. Interestingly, they apparently converge on increasing the regularity of circadian rhythms, although we observed virtually no overlap between features changing due to treatment and the features displaying variations proportional to the magnitude of response for specific treatments. We found that changes in 2 specific features correlate with the magnitude of response when all cohorts were pooled together (timing of circadian peak of activity, and variability in circadian entrainment), which suggests that more robust circadian

entrainment of activity (stable alignment to the 24h light-dark cycle) correlates with better response to MDD treatment. This may indicate convergent effects of different treatments, and points to circadian entrainment of activity being a relevant endpoint for evaluation of response to treatment in MDD subjects. We did not find differences between MDD subjects and healthy controls, before treatment or in remission after treatment, and therefore cannot conclusively assess whether symptom remission is associated with an apparent normalization of circadian patterns of activity.

The intrinsic heterogeneity of unipolar depression is reflected in the diverse pharmacodynamic profile of antidepressant drugs available. However, the response rate to monotherapy is largely similar among different classes (53.3% in iSPOT-D [36]; 48.6% in STAR*D [37]), and response to a certain drug class does not exclude response to a different class. We hypothesized that the response to treatment, estimated as relative change from baseline, can be modeled using features of circadian patterns of activity before treatment initiation. This approach has two potential pitfalls: (1) it does not account for variations in individual items in the scale; and (2) it ignores all intermediate states between before and after treatment. Regarding the former, it has been shown that the underlying variability is rather limited [38–41], which makes the total score a robust estimator of current state. In addition, we have shown earlier that the patterns of activity can be reliably used for modeling symptom severity before treatment, *i.e.*, without potential pharmacological interference on circadian patterns of activity [17]. The dynamics of response to treatment has been shown to be accurately approximated by an exponential decay [9], however the timeline differs considerably among treatments. For iCBT, the response appears linear (see [18]), and the response is most reliably evaluated at the end of the course (10 weeks). The response to SSRIs has been classified as either fast (stable response is reached within 2 weeks after treatment onset), or slow (response stabilizes after about 4-6 weeks) [42], while the drop in symptom severity can be documented within hours after a single dose of ketamine administration and is stable for a couple of days [5,43]. We assumed that symptom severity was stable before treatment and had reached a plateau after treatment at the time of recording. Provided the interval between measurements is appropriate for the actual treatment (*i.e.*, 10-12 weeks for iCBT; ~8 weeks for SSRIs; and ~2 weeks for ketamine), our modeling focused on the difference between 2 plateaus and made no assumption on the timeline of transition between the two states.

Approaching individual putative pathological mechanisms (*e.g.*, “serotonin deficiency”) is suitable for drug development but may not be satisfactory for multifactorial disorders such as major depression. The crosstalk between signaling pathways is illustrated by the recent finding that serotonin receptor 1B binding in ventral striatum before ketamine administration correlates with both symptom severity, and the response to treatment [19]. We believe that the investigation of correlations between patterns of activity, as expression of current status, and the response to treatment, as change driven by therapeutic intervention, is a meaningful approach. We have shown recently that patterns of activity can predict symptom severity [17], and in this study we focus on the identification of most relevant correlates for response to treatment. The linear regression ensemble parallel training increases the generalizability of the findings, while Bayesian model averaging accounts for model uncertainty and provides the tools for posterior analysis of coefficients [29,44]. The added value of Bayesian analysis of coefficients is to support the high-level interpretation of correlations between patterns of activity and response to treatment. The differences between iCBT and ketamine when modeling the response to treatment extend beyond the distinct signatures in terms of relevant features. For iCBT, weaker circadian entrainment, and earlier circadian peak of activity correlate with better response. For ketamine, more fragmented and less variable circadian patterns of activity, and shorter circadian period correlate with better response. In addition, higher age and steeper propensity to sustain activity correlated to better response to either treatment. Notably, the features differentially enriched for either iCBT and ketamine are related to circadian distribution of activity. Circadian entrainment of activity is dependent of the central clock, located in the suprachiasmatic nucleus (SCN), and the vast majority of neurotransmitter signaling pathways on which antidepressants act have been shown to impact circadian rhythms (recently reviewed in [43,45,46]). Briefly, serotonin tonically inhibits the direct input from retinal ganglionic neurons, and blocking serotonin reuptake weakens photic entrainment of the SCN (reviewed in [47]). Ketamine has been shown to directly interfere with clock gene transcriptional activity in primary fibroblasts [48], and appears to have widespread effects on circadian clock machinery [43], suggesting that ketamine treatment facilitates circadian entrainment of activity. Our results indicates that MDD patients with higher variability in magnitude of circadian peak of activity before treatment experience better antidepressant effects, which is consistent with the stabilization of circadian rhythms following ketamine treatment. In addition, we found that better response correlates with a shift towards earlier occurrence of circadian peak, in agreement with previous reports [49]. Regarding iCBT as non-pharmacological intervention, the patient is supposed to gradually

change behavior in a voluntary manner, therefore alterations in activity patterns can be expected after completing the treatment.

The accuracy of aggregated output indicates that useful models may potentially be developed for specific treatments (*i.e.*, iCBT or ketamine), but the correlations with response to any treatment have limited practical applicability. Our data suggest potential prospective applications for predicting the response to MDD treatment. However, we cannot assess the generalizability of our findings because we did not test the performance against an independent dataset.

In conclusion, we identified distinct signatures in the correlations between patterns of activity and response to specific MDD treatments. We further found that the stabilization of circadian entrainment correlates with the magnitude of response independently from the antidepressive treatment followed. Our data suggests that the analysis of circadian patterns of activity may potentially be used for predicting the response to treatment in MDD patients, but larger confirmatory studies are required to support clinical applications.

Acknowledgments

The authors would like to thank Prof. Matteo Bottai (Division of Biostatistics, Karolinska Institutet) for inspiring discussions and suggestions for planning the analyses. Hardware and support for actigraphy data collection was provided by the Ass. Prof. Viktor Kaldo (Department of Psychology, Karolinska Institutet) and Philips (Philips Healthcare, Sleep and Respiratory Care, Murrysville, PA, USA). Funding to support the work was provided by Swedish Society of Medicine, Fredrik and Ingrid Thuring Foundation, Region Stockholm (ALF project 20190429, and clinical research appointment), The Swedish Brain Foundation (Hjärnfonden), and The Söderström König Foundation (SLS-746501) (MT); Swedish Research Council (2013-09304), Region Stockholm (ALF project 20170192 and higher clinical research appointment) (JL); Swedish Research Council (2019-01191), The Swedish Brain Foundation (Hjärnfonden, FO2016-0116), and Torsten Söderberg Foundation (M59/16) (SC); and Karolinska Institutet research grants (MT, SC, JL). The funding agencies and sponsors did not have any influence on the conceptualization, design, data collection, analyses, decision to publish, or preparation of the manuscript.

Conflicts of Interest

The research leading up to the present work generated US Patent No. 10,731,216 – Methods and compositions for biomarkers of depression and pharmacoresponse (SC and SS inventors). The patent is owned by NorthernLight Diagnostics AB (SC and SS co-founders), a company developing decision support tools for mental health care. The company did not provide funding and was not involved in the present study.

References

1. Khan A, Faucett J, Lichtenberg P, Kirsch I, Brown WA. A Systematic Review of Comparative Efficacy of Treatments and Controls for Depression. https://doi.org/10.1371/journal.pone.0041778.
2. Weitz ES, Hollon SD, Twisk J, van Straten A, H Huibers MJ, David D, et al. Baseline Depression Severity as Moderator of Depression Outcomes Between Cognitive Behavioral Therapy vs Pharmacotherapy An Individual Patient Data Meta-analysis Supplemental content at jamapsychiatry.com Original Investigation | META-ANALYSIS. JAMA Psychiatry. 2015;72:1102–1109.
3. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. 2017. 2017. https://doi.org/10.1080/16506073.2017.1401115.
4. Machado-Vieira R, Baumann J, Wheeler-Castillo C, Latov D, Henter ID, Salvadore G, et al. The Timing of Antidepressant Effects: A Comparison of Diverse Pharmacological and Somatic Treatments. Pharmaceuticals. 2010;3:19.
5. Berman RM, Cappiello A, Anand A, Oren DA, Heninger GR, Charney DS, et al. Antidepressant effects of ketamine in depressed patients. Biol Psychiatry. 2000;47:351–354.
6. Krystal JH, Abdallah CG, Sanacora G, Charney DS, Duman RS. Ketamine: A Paradigm Shift for Depression Research and Treatment. Neuron. 2019;101:774–778.
7. Browning M, Kingslake J, Dourish CT, Goodwin GM, Harmer CJ, Dawson GR. Predicting treatment response to antidepressant medication using early changes in emotional processing. European Neuropsychopharmacology. 2019;29:66–75.
8. Athreya AP, Brückl T, Binder EB, John Rush A, Biernacka J, Frye MA, et al. Prediction of short-term antidepressant response using probabilistic graphical models with replication across multiple drugs and treatment settings. Neuropsychopharmacology. 2021;46:1272–1282.
9. Taliaz D, Spinrad A, Barzilay R, Barnett-Itzhaki Z, Averbuch D, Teltsh O, et al. Optimizing prediction of response to antidepressant medications using machine learning and integrated genetic, clinical, and demographic data. Transl Psychiatry. 2021;11.
10. Chapman JJ, Roberts JA, Nguyen VT, Breakspear M. Quantification of free-living activity patterns using accelerometry in adults with mental illness. Sci Rep. 2017;7:1–12.
11. Tazawa Y, Wada M, Mitsukura Y, Takamiya A, Kitazawa M, Yoshimura M, et al. Actigraphy for evaluation of mood disorders: A systematic review and meta-analysis. J Affect Disord. 2019;253:257–269.
12. Pan W, Song Y, Kwak S, Yoshida S, Yamamoto Y. Quantitative evaluation of the use of actigraphy for neurological and psychiatric disorders. Behavioural Neurology. 2014;2014.

13. Lyall LM, Wyse CA, Graham N, Ferguson A, Lyall DM, Cullen B, et al. Association of disrupted circadian rhythmicity with mood disorders, subjective wellbeing, and cognitive function: a cross-sectional study of 91 105 participants from the UK Biobank. Lancet Psychiatry. 2018;5:507–514.
14. Morres ID, Hatzigeorgiadis A, Krommidas C, Comoutos N, Sideri E, Ploumpidis D, et al. Objectively measured physical activity and depressive symptoms in adult outpatients diagnosed with major depression. Clinical perspectives. Psychiatry Res. 2019;280:112489.
15. Morres ID, Hatzigeorgiadis A, Stathi A, Comoutos N, Arpin-Cribbie C, Krommidas C, et al. Aerobic exercise for adult patients with major depressive disorder in mental health services: A systematic review and meta-analysis. Depress Anxiety. 2019;36:39–53.
16. Helgadóttir B, Forsell Y, Ekblom Ö. Physical Activity Patterns of People Affected by Depressive and Anxiety Disorders as Measured by Accelerometers: A Cross-Sectional Study. PLoS One. 2015;10:e0115894.
17. Spulber S, Elberling F, Svensson J, Tiger M, Ceccatelli S, Lundberg J. Patterns of activity correlate with symptom severity in major depressive disorder patients. Transl Psychiatry. 2022;12:226.
18. Svensson JE, Svanborg C, Plavén-Sigray P, Kaldo V, Halldin C, Schain M, et al. Serotonin transporter availability increases in patients recovering from a depressive episode. Transl Psychiatry. 2021;11:264.
19. Tiger M, Veldman ER, Ekman CJ, Halldin C, Svenningsson P, Lundberg J. A randomized placebo-controlled PET study of ketamine´s effect on serotonin1B receptor binding in patients with SSRI-resistant depression. Transl Psychiatry. 2020;10:159.
20. Ekholm B, Spulber S, Adler M. A randomized controlled study of weighted chain blankets for insomnia in psychiatric disorders. Journal of Clinical Sleep Medicine. 2020;16:1567–1577.
21. Gonçalves B, Adamowicz T, Louzada F, Moreno C, Araujo J. A fresh look at the use of nonparametric analysis in actimetry. Sleep Med Rev. 2015;20:84–91.
22. Hu K, Ivanov PCh, Chen Z, Hilton MF, Stanley HE, Shea SA. Non-random fluctuations and multi-scale dynamics regulation of human activity. Neuroscience. 2004;149:508–517.
23. Gonçalves BSB, Cavalcanti P, Tavares GR, Campos TF, Araujo JF. Nonparametric methods in actigraphy: An update. Sleep Science. 2014;7:158–164.
24. Saragiotis C. Lomb normalized periodogram. 2021.
25. Sokolove PG, Bushell WN. The chi square periodogram: its utility for analysis of circadian rhythms. J Theor Biol. 1978;72:131–160.
26. Tackenberg MC, Hughey JJ. The risks of using the chi-square periodogram to estimate the period of biological rhythms. PLoS Comput Biol. 2021;17:1–16.
27. Fasmer OB, Hauge E, Berle JØ, Dilsaver S, Oedegaard KJ. Distribution of active and resting periods in the motor activity of patients with depression and schizophrenia. Psychiatry Investig. 2016;13:112–120.
28. Edgar N, McClung CA. Major depressive disorder: a loss of circadian synchrony? Bioessays. 2013;35:940–944.
29. Raftery AE, Madigan D, Hoeting JA. Bayesian Model Aeraging for Linear Regression Models. J Am Stat Assoc. 1997;92:179–191.
30. Madigan D, Raftery AE. Model Selection and Accounting for Model Uncertainty in Graphical Models Using Occam’s Window. J Am Stat Assoc. 1994;89:1535–1546.
31. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association; 2013.

32. Kiehn J-T, Faltraco F, Palm D, Thome J, Oster H. Circadian Clocks in the Regulation of Neurotransmitter Systems. Pharmacopsychiatry. 2019. 30 May 2019. https://doi.org/10.1055/a-1027-7055.
33. Silva S, Bicker J, Falcão A, Fortuna A. Antidepressants and circadian rhythm: Exploring their bidirectional interaction for the treatment of depression. Pharmaceutics. 2021;13.
34. Grömping U. Variable importance in regression models. Wiley Interdiscip Rev Comput Stat. 2015;7:137–152.
35. Brooks TG, Lahens NF, Grant GR, Sheline YI, FitzGerald GA, Skarke C. Phenome-Wide Association Study of Actigraphy in the UK Biobank. MedRxiv. 2021:2021.12.09.21267558.
36. Saveanu R, Etkin A, Duchemin A-M, Goldstein-Piekarski A, Gyurak A, Debattista C, et al. The International Study to Predict Optimized Treatment in Depression (iSPOT-D): Outcomes from the acute phase of antidepressant treatment. J Psychiatr Res. 2015;61:1–12.
37. A. John Rush MD, Madhukar H. Trivedi MD, Stephen R. Wisniewski PhD, Andrew A. Nierenberg MD, Jonathan W. Stewart MD, Diane Warden, Ph.D. MBA, et al. Acute and Longer-Term Outcomes in Depressed Outpatients Requiring One or Several Treatment Steps: A STAR*D Report. American Journal of Psychiatry. 2006;163:1905–1917.
38. Rabinowitz J, Rabinowitz AA. Outlier-response pattern checks to improve measurement with the Montgomery-Asberg depression rating scale (MADRS). J Affect Disord. 2022;299:444–448.
39. Rabinowitz J, Schooler NR, Brown B, Dalsgaard M, Engelhardt N, Friedberger G, et al. Consistency checks to improve measurement with the Montgomery-Asberg Depression Rating Scale (MADRS). J Affect Disord. 2019;256:143–147.
40. Simmonds-Buckley M, Catarino A, Delgadillo | Jaime, Delgadillo J. Depression subtypes and their response to cognitive behavioral therapy: A latent transition analysis. Depress Anxiety. 2021;38:907–916.
41. Catarino A, Fawcett JM, Ewbank MP, Bateup S, Cummins R, Tablan V, et al. Refining our understanding of depressive states and state transitions in response to cognitive behavioural therapy using latent Markov modelling. Psychol Med. 2022;52:332–341.
42. Fiori LM, Orri M, Aouabed Z, Théroux JF, Lin R, Nagy C, et al. Treatment-emergent and trajectory-based peripheral gene expression markers of antidepressant response. Translational Psychiatry 2021 11:1. 2021;11:1–8.
43. Sato S, Bunney B, Mendoza-Viveros L, Bunney W, Borrelli E, Sassone-Corsi P, et al. Rapid-acting antidepressants and the circadian clock. Neuropsychopharmacology. 2022;47:805–816.
44. Astuti AB, Iriawan N, Kuswanto H. Model Components Selection in Bayesian Model Averaging Using Occam's Window for Microarray Data. Journal of Scientific Modeling & Computation. 2014;1.
45. Lee R, McGee A, Fernandez F-X. Systematic review of drugs that modify the circadian system's phase-shifting responses to light exposure. Neuropsychopharmacology. 2022;47:866–879.
46. Daut RA, Fonken LK. Circadian regulation of depression: A role for serotonin. Front Neuroendocrinol. 2019;54:100746.
47. Ciarleglio CM, Resuehr HES, McMahon DG. Interactions of the serotonin and circadian systems: nature and nurture in rhythms and blues. Neuroscience. 2011;197:8–16.

48. Bellet MM, Vawter MP, Bunney BG, Bunney WE, Sassone-Corsi P. Ketamine Influences CLOCK:BMAL1 Function Leading to Altered Circadian Gene Expression. PLoS One. 2011;6:23982.
49. Duncan WC, Slonena E, Hejazi NS, Brutsche N, Yu KC, Park L, et al. Motor-Activity Markers of Circadian Timekeeping Are Related to Ketamine's Rapid Antidepressant Properties. Biol Psychiatry. 2017;82:361–369.

Table 1. Description of patient cohorts and antidepressant treatments

Treatment	pre	post	both[1]	Treatment scheme	Ref
HC	19[2] (14)	N/A	N/A	N/A	[x]
iCBT	12 (9)	11 (9)	8 (6)	10 self-administered, internet-delivered sessions (1 session/week)	[x]
Escitalopram	7 (3)	6 (2)	5 (2)	50 mg/ day	[y]
Ketamine	16 (8)	15 (6)	12 (6)	5 mg i.v. infusion, 2/week	[z]

Numbers in parentheses indicate the number of female subjects in each group.
[1] – used for paired comparisons, post vs. pre

[2] – 1 subject was recorded 3 times; 1 subject was recorded 2 times; 16 independent subjects

Table 2. Group re-allocation and descriptive statistics of data available.

Group	N subjects	N days	days/ M	days/ F
HC	19 (14)	232 (173)	11.8	12.4
MDD (pre)[1]	35 (20)	277 (172)	6.2	8.6
iCBT (post)	11 (9)	108 (96)	6	10.7
Escitalopram (post)	8 (2)	62 (8)	10.3	4
Ketamine (post)	15 (6)	141 (72)	7.7	12

Numbers in parentheses indicate the number of days actigraphy recorded in female subjects. See also Fig 1 for details on timing of recording relative to treatment administration.
[1] – includes all recordings before treatment from all treatment groups

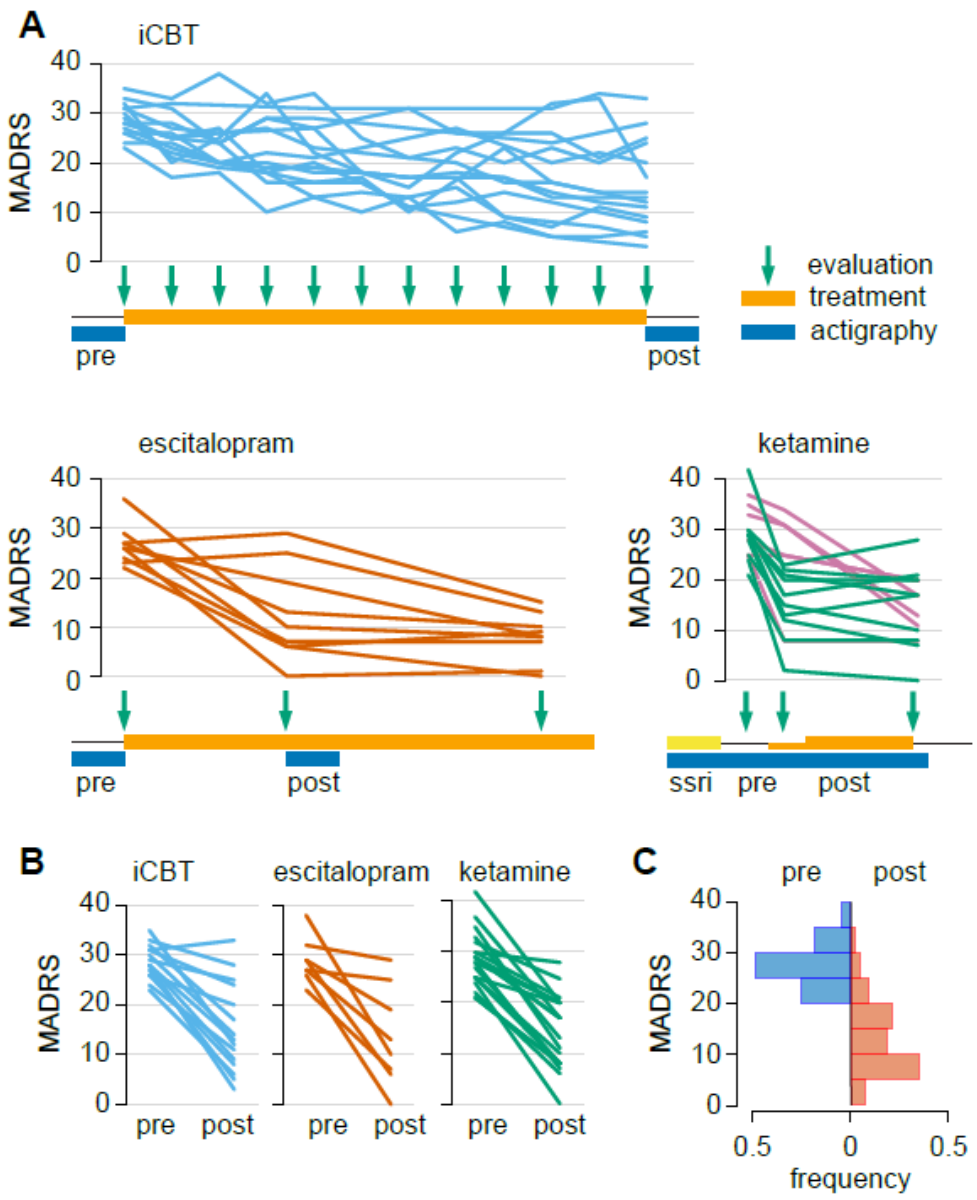

Fig. 1 Description of source data. (A) Timeline for data collection for all cohorts included. Only subjects for which actigraphy recordings were available are shown. (B) Change in MADRS score in individual subjects for each cohort. (C) Distribution of MADRS scores before and after treatment in pooled cohorts. The subjects with MADRSpost < 10 are included in the "remission" group in subsequent analyses.

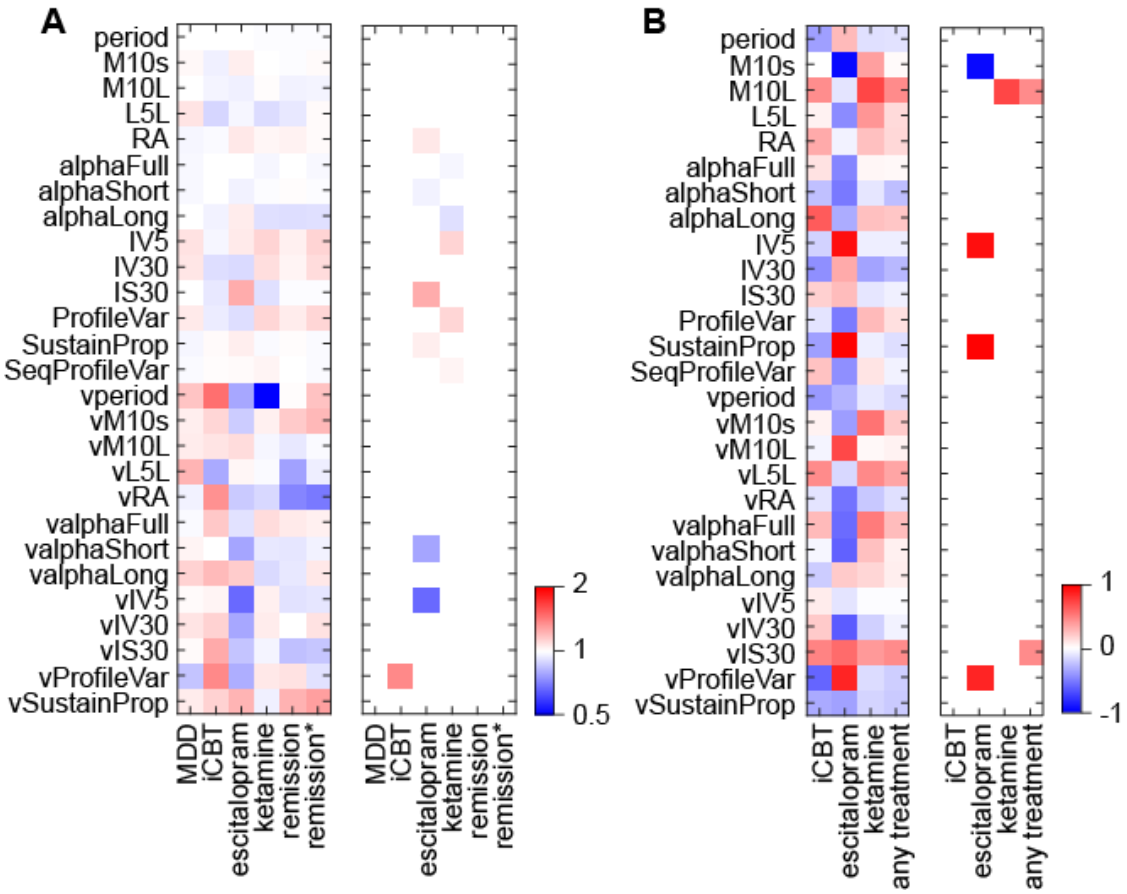

Fig. 2 Changes in patterns of activity associated with ongoing antidepressive treatment. (A) Between-group differences in individual features. Each treatment group is compared against MDD subjects before treatment as reference. MDD subjects before treatment are compared against healthy controls. The heatmap in the left panel displays the ratios between each group average and their respective reference group average. Significant differences surviving FDR correction are displayed in the right panel. Note the distinct signatures of individual treatments. Subjects in remission (independent of antidepressant treatment) display no significant differences from MDD subjects before treatment (remission) or from healthy controls (remission*). (B) Correlations between changes in individual features and the magnitude of response to treatment. Daily variability of IS (vIS30) and timing of circadian peak (M10L) appear to correlate with the response to treatment independently from actual antidepressive treatment.

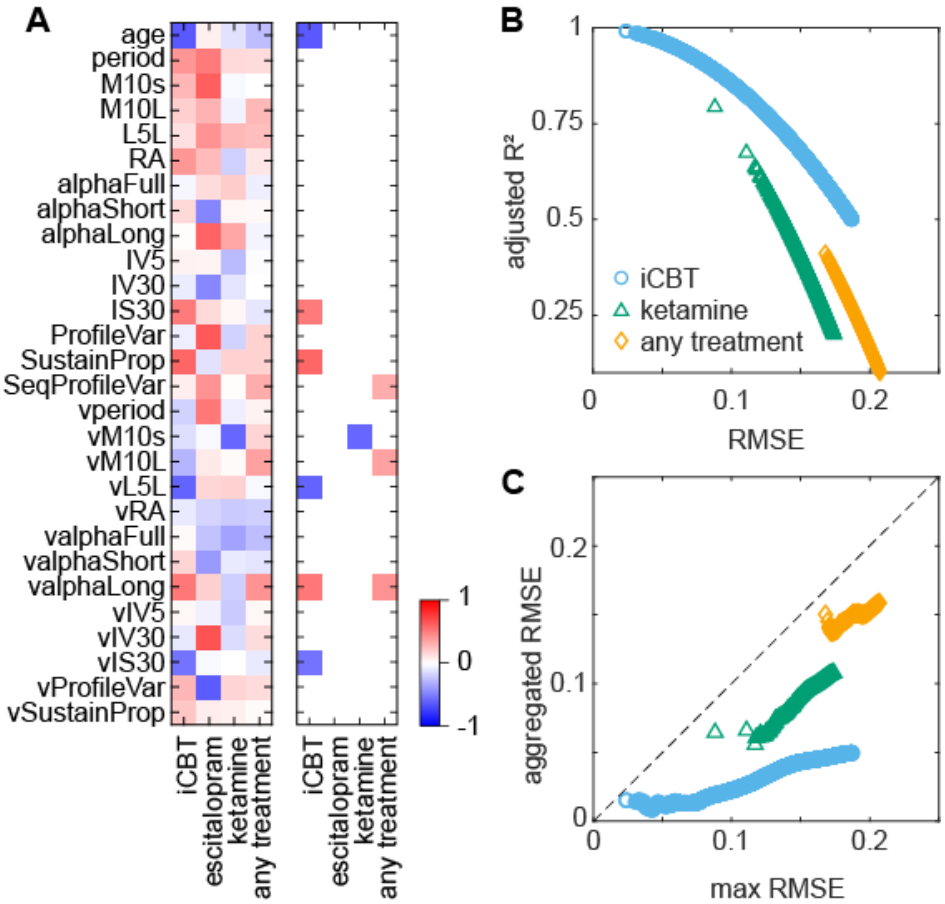

Fig. 3 Correlations between activity features before treatment and the response to different antidepressant treatments. (A) Correlations with magnitude of response to treatment. Left panel: heatmap of raw correlation coefficients. Right panel: correlation coefficients surviving FDR correction. Note that response to treatment regardless of the nature of antidepressant treatment does not correlate with any individual feature, and that individual treatments have distinct signatures in terms of individual features correlating with the magnitude of the response. (B) Parallel training of ensembles of multiple regression models for response to treatment. Illustration of evaluation of performance for individual models. (C) Evaluation of performance of pruned ensembles. Aggregated output is calculated as cumulative average of individual model output, after sorting the ensembles by decreasing accuracy (increasing RMSE).

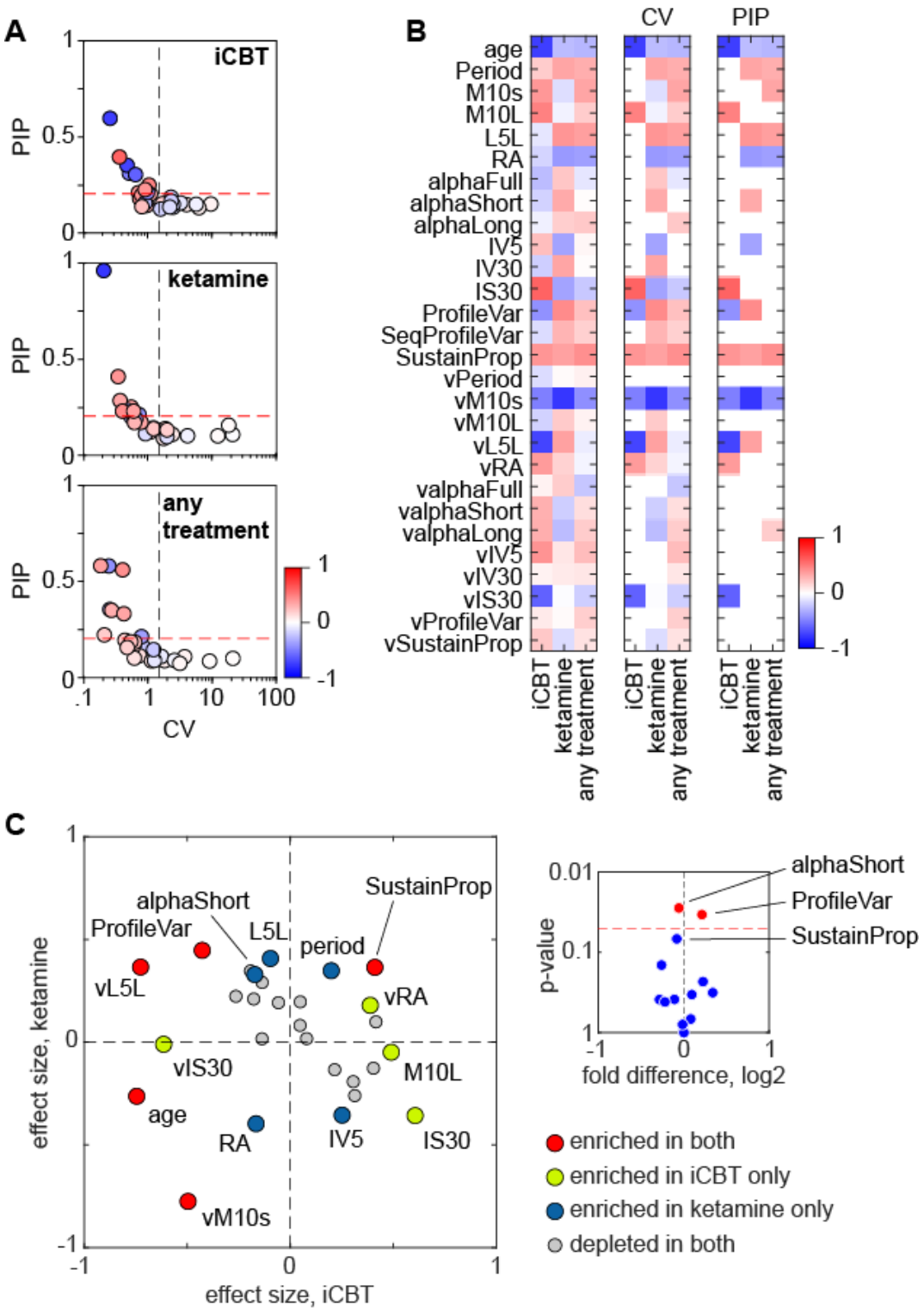

Fig. 4 Bayesian analysis of coefficients. (A) Analysis of occurrence and variability of individual features in each ensemble. Thresholds for CV (arbitrarily set to 1.5) and PIP (set to prior inclusion probability, 0.2035). Features in the top-left quadrant have stable coefficients

and have medium to strong evidence of correlation with the magnitude of response to treatment. (B) Average standardized coefficients for all ensembles to illustrate the differences in ensemble composition. The heatmap in the first panel shows all average standardized coefficients. The second and third panel display only the coefficients with CV < 1.5 and PIP > 0.2035, respectively (see (A) for reference). Note the distinct, non-overlapping signatures for each ensemble. (C) Direct comparison between effect sizes of individual features in ensembles modeling the response to iCBT and ketamine. Enriched features highlighted and color-coded based on differential enrichment in either ensemble. Features mapping in top-right and bottom-left quadrants indicate similar effects, while features mapping in top-left and bottom-right quadrants indicate opposite effects between ensembles. In addition, features mapping close to either axis indicate large difference in effect size between ensembles. Comparison between the two populations for enriched features displayed in the insert. Significant differences surviving FDR correction found in alphaShort and ProfileVar, both located in the top-left quadrant, indicating opposite correlations.