Journal of Psychiatric Research 103 (2018) 46–53

Contents lists available at ScienceDirect

Journal of Psychiatric Research

journal homepage: www.elsevier.com/locate/jpsychires

Cortisol trajectory, melancholia, and response to electroconvulsive therapy

Brian J. Mickey[a,b,*], Yarden Ginsburg[b], Adam F. Sitzmann[b], Clara Grayhack[b], Srijan Sen[b], Clemens Kirschbaum[c], Daniel F. Maixner[b], James L. Abelson[b]

[a] Department of Psychiatry, University of Utah School of Medicine, Salt Lake City, USA
[b] Department of Psychiatry, University of Michigan Medical School, Ann Arbor, USA
[c] Department of Psychology, Technische Universität Dresden, Dresden, Germany

ARTICLE INFO

Keywords:
Depression
Cortisol
Electroconvulsive therapy
Biomarker
Melancholia

ABSTRACT

While biomarkers have been used to define pathophysiological types and to optimize treatment in many areas of medicine, in psychiatry such biomarkers remain elusive. Based on previously described abnormalities of hypothalamic-pituitary-adrenal (HPA) axis function in severe forms of depression, we hypothesized that the temporal trajectory of basal cortisol levels would vary among individuals with depression due to heterogeneity in pathophysiology, and that cortisol trajectories that reflect elevated or increasing HPA activity would predict better response to electroconvulsive therapy (ECT). To test that hypothesis, we sampled scalp hair from 39 subjects with treatment-resistant depression just before ECT. Cortisol trajectory over the 12 weeks preceding ECT was reconstructed from cortisol concentrations in sequential hair segments. Cortisol trajectories varied widely between individuals, and exploratory analyses of clinical features revealed associations with melancholia and global severity. ECT non-responders showed a decreasing trajectory (mean change −25%, 95%-CI = [–1%,–43%]) during the 8 weeks preceding ECT (group-by-time interaction, p = 0.004). The association between cortisol trajectory and subsequent ECT response was independent of clinical features. A classification algorithm showed that cortisol trajectory predicted ECT response with 80% accuracy, suggesting that this biomarker might be developed into a clinically useful test for ECT-responsive depression. In conclusion, cortisol trajectory mapped onto symptoms of melancholia and independently predicted response to ECT in this severely depressed sample. These findings deserve to be replicated in a larger sample. Cortisol trajectory holds promise as a reliable, noninvasive, inexpensive biomarker for psychiatric disorders.

1. Introduction

Biomarkers have been used to define disease types and select treatments in many areas of medicine, but such biomarkers remain elusive in psychiatry. For example, clinicians who treat patients with depression strive to match each patient with potential treatments – including psychosocial interventions, medications, and brain stimulation therapies – but no objective biomarkers have yet proven useful for clinical decision-making (Strawbridge et al., 2017). For patients with severe depression, a common decision point is whether to proceed with electroconvulsive therapy (ECT). ECT is uniquely effective for medication-resistant depression but also associated with significant side effects. Furthermore, of the estimated 100,000 patients treated annually in the US (Hermann et al., 1995), roughly one-third do not respond to ECT (Haq et al., 2015; Lisanby, 2007). The effectiveness of ECT can vary with factors such as correct diagnosis, depression severity, concurrent medication, and how the ECT is administered, but for many patients the reasons for non-response remain obscure. Accurate prediction of individual response to ECT could therefore reduce exposure of many patients to an ineffective treatment, while encouraging the appropriate use of this intervention for patients with ECT-responsive depression.

These potential benefits have kindled interest in identifying clinical features and biomarkers that predict ECT outcomes. Recent meta-analyses by our group and others showed that clinical features – including depressive episode duration, recent medication failure, age, and psychotic features – predicted ECT response, hinting that a patient's recent history might be informative for predicting subsequent outcomes (Haq et al., 2015; van Diermen et al., 2018). Yet, the relatively weak associations of pre-treatment clinical features with subsequent ECT response indicate that clinical features alone have limited predictive power (Haq et al., 2015; van Diermen et al., 2018). Evidence for pre-treatment plasma and genetic biomarkers that predict ECT response has been mixed, as reviewed recently (Pinna et al., 2016). Several studies

* Corresponding author. 501 Chipeta Way, Salt Lake City, UT 84108, USA.
E-mail address: brian.mickey@utah.edu (B.J. Mickey).

https://doi.org/10.1016/j.jpsychires.2018.05.007
Received 17 January 2018; Received in revised form 19 March 2018; Accepted 7 May 2018

have reported association of ECT response with pre-treatment neuroimaging features. For example, two recent magnetic resonance imaging studies applied machine learning and demonstrated the capacity to predict ECT response with accuracies of 68–89% (Redlich et al., 2016; Wade et al., 2016). These findings suggest that biomarkers – perhaps in combination with clinical features – might aid clinicians in deciding whether to attempt ECT.

Dysregulation of the hypothalamic-pituitary-adrenal (HPA) axis and cortisol secretion have long been associated with chronic stress and related psychiatric disorders (Holsboer and Ising, 2010) including the major mood disorders, as recently reviewed (Belvederi Murri et al., 2016; Gold, 2015; Zorn et al., 2017). Furthermore, several studies have associated HPA phenotypes with treatment outcomes, suggesting that such phenotypes might be useful in defining pathophysiological types of depression. For example, excessive HPA activity predicted poorer response to psychotherapy (Fischer et al., 2017); abnormal cortisol response to the combined dexamethasone suppression corticotropin-releasing hormone stimulation test predicted better response to antidepressant medication (Binder et al., 2009); and higher post-dexamethasone salivary cortisol was associated with better response to ECT (Vukadin et al., 2011). While these results are promising, it remains unclear whether HPA measures could predict outcomes in a clinically useful way.

Hair cortisol provides a unique window into the recent history of an individual's HPA function. Recent studies indicate that hair cortisol measurements are reliable across laboratories (Russell et al., 2015) and sensitive to stress (Dettenborn et al., 2012; Stalder et al., 2017; Wester and van Rossum, 2015). Because scalp hair grows at a rate of ~0.25 cm/week (Loussouarn et al., 2005), cortisol concentrations in each 1-cm segment of hair reflect the average cortisol level over a specified 4-week period. Cortisol measured from sequential hair segments can therefore be used to reconstruct an individual's cortisol trajectory over recent weeks. We hypothesized that cortisol trajectory would vary among patients with treatment-resistant depression due to heterogeneity in underlying pathophysiology, and that cortisol trajectories reflecting elevated or increasing HPA activity would predict better response to ECT. To test this hypothesis, we measured hair cortisol concentrations in depressed patients treated with ECT.

2. Materials and methods

2.1. Study design and participants

In this prospective, observational study of treatment-resistant depression, we recruited adult inpatients and outpatients who were referred for ECT at a single academic center. The study was approved by the local institutional review board and carried out in accordance with the Declaration of Helsinki. All subjects provided written informed consent. Eligible participants had medication-resistant, moderate-to-severe depression (DSM-IV/5 major depressive disorder or bipolar disorder) for at least 2 months. As detailed in the Supplemental Methods, medication resistance was defined by failure of at least one medication trial of adequate dose and duration within the current episode; all participants had a lifetime history of multiple medication failures. Forty-two eligible participants consented, provided hair samples, and completed ECT. Three subjects were excluded, leaving 39 available for analysis. See Supplemental Methods for details.

2.2. Baseline assessments

Each patient was assessed by two board-certified psychiatrists and a masters-level psychologist. Assessment included the Mini International Neuropsychiatric Interview Plus, version 5.0.0 (Sheehan et al., 1998) and the Structured Interview Guide for the Hamilton Depression Rating Scale (Williams et al., 1988). Based on evidence that classic melancholia is not adequately characterized by DSM-IV/5 criteria (Parker et al., 2010), we also rated each subject on a series of 11 signs and symptoms traditionally associated with melancholia (Gold, 2015; Parker et al., 2010; Taylor and Fink, 2006) and calculated a *melancholia index* (number of signs/symptoms of melancholia divided by 11; range 0–1). See Supplemental Methods for details.

2.3. Treatment and post-treatment assessments

During this observational study, ECT treatments were administered as clinically indicated using bitemporal or right-unilateral, brief-pulse stimulation. The primary outcome measure was post-treatment "responder" status, a dichotomous variable. Responders were defined as 1 or 2 ("very much" or "much" improved) and non-responders as ≥3 on the Clinical Global Impression (CGI) improvement scale. CGI ratings were based on depressive symptom ratings, patient and caregiver reports, and team consensus, and were performed before cortisol was assayed to prevent bias. Responder status was chosen as the primary outcome over other measures (e.g., change in a depression rating scale) because it represents meaningful benefit that informs dichotomous clinical decision-making. More specifically, ECT would likely not have been attempted with non-responders had the outcome been known, and a non-responder would be unlikely to attempt ECT again in the future. See Supplemental Methods for details.

2.4. Hair cortisol quantification

Hair samples were collected just before the ECT series using established procedures (see Supplemental Methods). Each subject's cortisol trajectory was reconstructed from this single hair sample. The most proximal 3-cm portion of each hair sample was cut into three 1-cm segments (5.5 mg per segment), washed, pulverized, and alcohol extracted (Kirschbaum et al., 2009). Cortisol concentrations (picograms of cortisol per mg of hair) were quantified with a commercial chemoluminescence-based immunoassay (Dettenborn et al., 2012). Because the distribution of cortisol concentration values was skewed and wide-ranging, values were $\log_{10}$ transformed, and changes were expressed as fold-change (i.e., the ratio of final over initial concentration).

2.5. Statistical analysis

Analyses were performed using R statistical software (version 3.2.4). Hair cortisol concentration was modeled as a repeated measure using a linear mixed model (*lmer* function, *lme 4* package, version 1.1–12) with subject as the random effect. To test the primary hypothesis that cortisol trajectory was associated with subsequent response to ECT, response group, time (i.e., hair segment), group-by-time, and age were modeled as fixed effects. Similarly, in exploratory analyses of each clinical feature, the feature, time, feature-by-time, and age terms were included as fixed effects. Wald χ^2 tests and corresponding p-values were calculated from fitted models using the *Anova* function (*car* package, version 2.1–2).

To evaluate the predictive potential of pre-treatment cortisol trajectory, we applied the K-nearest neighbors (KNN) algorithm (*knn* function, *class* package, version 7.3–14), a supervised, non-parametric, statistical learning tool (James et al., 2015). The KNN classifier was chosen over alternative approaches such as logistic regression or linear discriminant analysis because it makes no assumptions about the shape of the decision boundary, and it typically performs better than other methods when the number of features is small (James et al., 2015). Ten-fold cross-validation was applied to prevent overfitting. Confidence intervals were calculated using random permutation (n = 1000). The optimal value of K was determined algorithmically as the value that produced maximal mean accuracy; for this data set, $K = 3$ was optimal, $K = 4$ was next best (accuracy 72%), and other values of K performed poorly (accuracy < 70%).