Distinct predictors of short- versus long-term depression outcomes following electroconvulsive therapy
Supplemental Material

Brian J. Mickey*, Yarden Ginsburg, Erica Jensen, Daniel F. Maixner
Department of Psychiatry, Huntsman Mental Health Institute, University of Utah, Salt Lake City, Utah 84108, USA
Department of Psychiatry, University of Michigan, Ann Arbor, Michigan 48109, USA
*brian.mickey@utah.edu

Supplemental Methods

Statistical analyses were performed using R statistical software (version 3.6.0) in the R Studio environment (version 1.2.1335). During preliminary analyses, distributions of predictor variables and outcome variables were inspected; continuous variables with high skew were log- or square-root-transformed to improve normality. Next, descriptive statistics and Spearman correlations were calculated for predictor variables to characterize those variables and identify inter-correlations.

We hypothesized that five baseline subject characteristics would predict short- and long-term outcomes after ECT: duration of the index episode, medication resistance, psychotic features, catatonic features, and greater age. For each of these predictors, two-sided $p < 0.05$ (one-sided $p < 0.025$) was considered significant and two-sided $p < 0.10$ (one-sided $p < 0.05$) was considered marginal. For exploratory analyses of other baseline features, Bonferroni correction was applied to two-sided p-values to account for tests of 55 predictor variables (threshold $p = 0.05/55 = 0.0009$). Categorical features represented by fewer than 5 subjects were excluded from analysis.

The primary short-term outcome measure was dichotomous ECT response (responder versus non-responder). To test each baseline variable as a potential predictor of ECT response, we applied logistic regression ("glm" function, "binomial" family) with the sociodemographic or clinical feature as a predictor variable and ECT response as the outcome variable. Regression coefficients were converted back to proportions using the inverse logit function. We also computed effect sizes for each predictor: odds ratio for dichotomous predictors; Hedges' standardized mean difference for continuous predictors. The secondary short-term outcome variable was post-treatment MADRS. To evaluate each baseline feature as a potential predictor of the post-ECT MADRS score, we used linear regression ("lm" function) including baseline MADRS score and the feature of interest as predictor variables.

Long-term follow-up outcomes were analyzed with linear mixed models ("lmer" function, "lme4" package, version 1.1-21) with longitudinal PHQ-9 score as the response variable. We chose a linear mixed model to accommodate missing data and repeated correlated measures over time. Subject intercept was modeled as a random-effects predictor. Baseline PHQ-9 score and time (6, 12, 18, or 24 months) were included as fixed-effect predictors in all models. ECT response was included in some models, as described in the Results. To evaluate each sociodemographic or clinical feature, the feature and its interaction with acute ECT response were added to the model as fixed effects to identify associations with longitudinal outcome. Wald χ^2 tests and corresponding p-values were calculated from fitted models using the "Anova" function ("car" package, version 3.0–10).

Supplemental Results

Participant characteristics

Of 162 patients who screened as eligible, 114 (70%) enrolled in the study and received ECT (Supplemental Figure S1). Seventeen eligible patients were not asked to participate because they were considered poor candidates due to severe catatonia, psychosis, or cognitive impairment; an additional 18 were asked to participate but declined. Compared to the 114 enrolled subjects who received ECT, those that refused or were poor candidates were similar in sex distribution ($p > 0.2$, Fisher exact tests) but were older (mean age 66 vs 50 years, $p < 0.0001$, two-sample t tests). Therefore, our sample is likely to be younger and less severe, on average, than the population of patients seen in this clinical setting.

Predictors of drop-out during the index ECT series

The 9 ECT non-completers differed from the 105 completers on several baseline clinical features. Non-completers were more likely to have a primary diagnosis of bipolar disorder (6/9 versus 20/105, $p = 0.004$, Fisher exact test), comorbid post-traumatic stress disorder (2/9 versus 2/92, 13 missing values, $p = 0.04$, Fisher exact test), or

comorbid musculoskeletal disease (4/7 versus 18/99, 8 missing values, p = 0.03, Fisher exact test). In addition, a maternal history of mood disorder was more common among non-completers (9/9 versus 35/92, 13 missing values, p = 0.0003, Fisher exact test). Non-completers did not differ from completers on any baseline sociodemographic variable or on other clinical features in Table S1 (all p > 0.05).

Predictors of short-term outcome

Some intercorrelations were evident among the five hypothesized predictors. MGH stage (medication resistance) and episode duration were positively associated (Spearman rho = 0.60, $p = 3 \times 10^{-12}$). Catatonic features were associated with greater age (t = 4.6, p = 0.003, two-sample t test) and with psychotic features (odds ratio = 8.5, p = 0.024, Fisher exact test). Other intercorrelations were not significant (p > 0.05).

Associations of baseline demographic and clinical features with acute ECT response are shown in Supplemental Table S1. Maudsley stage, MGH stage, and episode duration were intercorrelated (Spearman rho = 0.61 to 0.84). Consistent with findings for MGH stage and episode duration, we found a marginally significant association (p = 0.053, logistic regression) of acute ECT response with lower Maudsley stage – a composite measure that incorporates illness severity, episode duration, and medication resistance. Similarly, more improvement in MADRS total score was found among subjects with lower Maudsley stage (p = 0.002, linear regression).

In exploratory analyses, we examined a wide range of other demographic and clinical baseline variables as potential predictors of dichotomous ECT response (Supplemental Table S1). Subjects with a maternal family history of mood disorder were less likely to respond to ECT (18/35 vs 42/57; p = 0.032, logistic regression), but no other baseline characteristic was statistically significant (p > 0.05). Exploratory analyses of ECT-related change in MADRS score revealed greater improvement among right-handed subjects (p = 0.004, linear regression) and those with agoraphobia (p = 0.047, linear regression), but there was no statistically significant association with family history or other baseline variables (p > 0.05). No baseline features remained significant after applying a conservative threshold of p = 0.0009 to account for multiple testing.

Response to the acute ECT series was associated with the number of ECT treatment sessions (p = 0.001, logistic regression). The mean number of sessions was 9.8 among responders and 11.9 among non-responders. This finding is consistent with the typical practice of delivering the minimum number of ECT treatments needed to achieve clinical response (i.e., continuing the series only until the maximum benefit is reached). Acute ECT response was not associated with initial electrode configuration (p = 0.33, logistic regression). Response rates were 62% versus 71% for subjects treated initially with bitemporal versus right unilateral ECT.

Predictors of long-term outcome

Based on the findings that better long-term PHQ-9 scores were associated with shorter episode duration, psychotic features, and catatonic features at baseline (p = 0.044, 0.021, 0.021, linear mixed models), and that PHQ-9 scores tended to improve more in older subjects (p = 0.067), we performed exploratory analyses using pairs of these four baseline features as predictors in different linear mixed models. We found that psychotic features persisted as an independent predictor of long-term outcomes when any of the other three predictors was included in the model (all p < 0.05). A similar pattern was apparent for catatonic features as an independent predictor (all p < 0.06). On the other hand, the predictive capacity of baseline episode duration tended to weaken when other predictors were included in the model (all p > 0.05).

We evaluated baseline characteristics as predictors of long-term PHQ-9 scores while controlling for acute ECT response. Those analyses showed that the effects of episode duration and psychotic features diminished (p = 0.088 and p = 0.060, linear mixed models), suggesting that the long-term influences of those characteristics were at least partially mediated through acute response to ECT. (Catatonia could not be evaluated similarly because all catatonic patients responded to ECT.) In contrast, married status and greater baseline social support continued to predict long-term PHQ-9 scores (p = 0.000034 and 0.0068, linear mixed models). Similarly, more severe illness at baseline (CGI-S, MADRS, and GAF) continued to predict better PHQ-9 scores longitudinally while controlling for acute ECT response (p = 0.001, 0.0005, and 0.047, linear mixed models). These findings indicate that the long-term influence of these social and severity predictors was independent of the initial therapeutic effects of ECT. The influence of these predictors was apparently not mediated by ongoing treatment, since intensity of treatment was not associated with PHQ-9 scores.

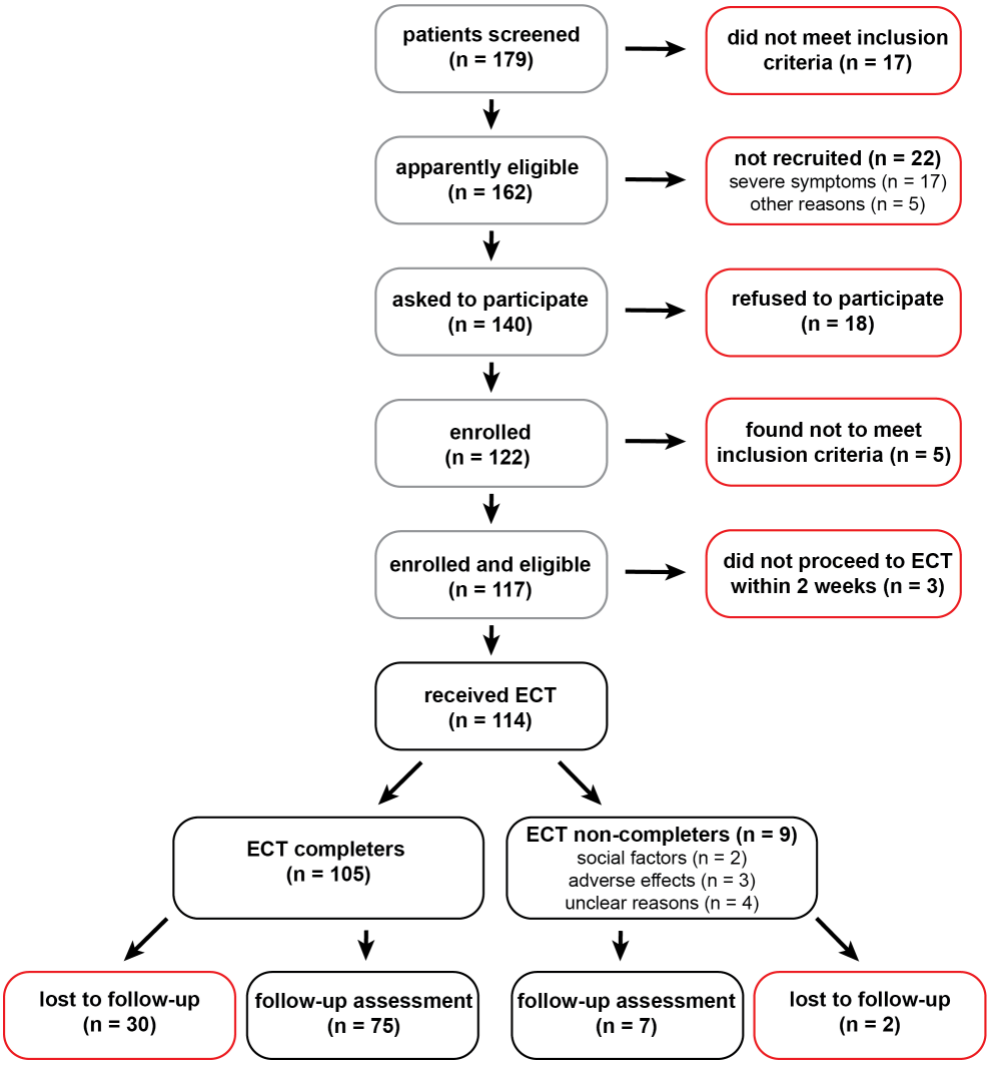
Using a linear mixed model that included baseline PHQ-9 score and acute ECT response as predictors, longitudinal PHQ-9 scores were lower among participants who were married (p = 0.000034). PHQ-9 scores were also lower among those with higher income (p = 0.00045). Marital status and income were strongly associated ($p = 6 \times 10^{-9}$,

two-sample t test). Income was not a significant predictor within married or unmarried subgroups ($p > 0.20$), so we concluded that marital status drove the association.

Self-reported social support at baseline (as measured with the WHOQOL-BREF social domain score) also predicted significantly lower longitudinal PHQ-9 scores ($p = 0.0068$). Social support was greater among married compared to unmarried subjects ($p = 0.03$, two-sample t test). When both predictors were included in a linear mixed model, marital status remained significant ($p = 0.0008$) and social support trended significant ($p = 0.057$), suggesting the effects of marital status and social support were independent to some extent. Psychological and environmental quality of life (WHOQOL-BREF domain scores) were nominally significant predictors of longitudinal PHQ-9 scores ($p = 0.024$ and 0.037, respectively), but when marital status was added to the respective models, neither remained significant (both $p > 0.10$).

Several clinician-rated measures of baseline depression severity predicted PHQ-9 scores longitudinally after controlling for baseline PHQ-9 score and acute ECT response. In contrast to baseline PHQ-9 score which predicted *higher* long-term PHQ-9 scores, baseline clinician-rated severity measures (MADRS, CGI-S, and GAF) predicted *lower* long-term PHQ-9 scores ($p = 0.0005$, 0.001, and 0.047, respectively). A non-significant effect in the same direction was found for HDRS ($p = 0.18$). MADRS, CGI-S, GAF, and HDRS were intercorrelated (absolute value of Spearman rho ranged 0.35 to 0.73, $p < 0.001$). Co-morbid generalized anxiety disorder (GAD) was also associated with better subsequent PHQ-9 scores ($p = 0.023$) but the effect weakened when MADRS was included in the model ($p = 0.097$), suggesting that GAD was simply marking more severe illness.

Supplemental Figure S1. Participant flow diagram.



Supplemental Table S1. Analysis of baseline predictors of acute response to ECT

Variable	Effect Size Metric	Effect Size Estimate [a]	95% confidence	interval	P [b]
Sociodemographic					
Age	Hedges' g	0.360	-0.057	0.778	0.07
Female	odds ratio	0.943	0.376	2.321	1
White	odds ratio	1.000	0.152	5.054	1
Married	odds ratio	0.753	0.307	1.834	0.54
Living alone	odds ratio	1.000	0.364	2.893	1
Education	Hedges' g	-0.036	-0.450	0.379	0.85
Disability	odds ratio	0.584	0.234	1.456	0.21
Income	Hedges' g	0.265	-0.169	0.700	0.21
DSM features					
Bipolar	odds ratio	1.206	0.384	4.247	0.80
Psychotic	odds ratio	3.381	0.688	32.938	0.13
Melancholic	odds ratio	1.512	0.502	4.401	0.45
Catatonic	odds ratio	infinite	0.464	infinite	0.17
Severity measures					
CGI severity	Hedges' g	0.259	-0.157	0.675	0.19
GAF	Hedges' g	-0.303	-0.720	0.113	0.14
MADRS	Hedges' g	0.057	-0.372	0.487	0.78
HDRS, 17-item	Hedges' g	0.209	-0.230	0.648	0.37
PHQ-9	Hedges' g	0.043	-0.415	0.502	0.86
QIDS-SR	Hedges' g	-0.141	-0.603	0.322	0.51
WSAS	Hedges' g	-0.206	-0.671	0.259	0.37
Chronicity and resistance					
Episode duration [c]	Hedges' g	-0.397	-0.815	0.021	0.06
Onset age [c]	Hedges' g	0.094	-0.322	0.510	0.64
Single episode	odds ratio	0.703	0.200	2.646	0.56
Number of episodes 2-10	odds ratio	1.225	0.498	3.036	0.68
Number of episodes >10	odds ratio	0.978	0.383	2.565	1
MGH stage	Hedges' g	-0.512	-0.938	-0.086	0.02
Maudsley stage	Hedges' g	-0.409	-0.827	0.009	0.04
Past improvement with ECT	odds ratio	2.603	0.834	9.792	0.09
Other clinical features					
GAD-7	Hedges' g	0.231	-0.232	0.693	0.32
ACE scale [d]	Hedges' g	-0.109	-0.580	0.362	0.63
WHOQOL-BREF Physical	Hedges' g	-0.097	-0.600	0.407	0.67
WHOQOL-BREF Psychological	Hedges' g	0.440	-0.069	0.948	0.08
WHOQOL-BREF Social	Hedges' g	0.000	-0.505	0.505	1
WHOQOL-BREF Environmental	Hedges' g	-0.290	-0.795	0.216	0.24
Maternal mood disorder	odds ratio	0.382	0.142	1.008	0.04
Paternal mood disorder	odds ratio	0.673	0.233	1.981	0.46
either parent	odds ratio	0.433	0.161	1.125	0.08
both parents	odds ratio	0.476	0.127	1.779	0.23
Right-handed	odds ratio	3.369	0.736	17.513	0.08
Psychiatric comorbidities					
Generalized anxiety disorder	odds ratio	0.643	0.244	1.658	0.38
Agoraphobia	odds ratio	0.598	0.228	1.540	0.28
Panic disorder	odds ratio	0.973	0.289	3.587	1
Social phobia	odds ratio	0.835	0.313	2.270	0.82
Specific phobia	odds ratio	1.658	0.274	17.784	0.71
Obsessive compulsive disorder	odds ratio	0.167	0.003	2.186	0.12
Substance use disorder, current	odds ratio	0.832	0.215	3.562	0.76
Substance use disorder, lifetime	odds ratio	1.209	0.471	3.131	0.83
Medical comorbidities					
Cardiac	odds ratio	4.411	0.551	203.395	0.27
Hypertension	odds ratio	1.261	0.469	3.602	0.66
Vascular	odds ratio	0.504	0.035	7.234	0.60
Respiratory	odds ratio	1.168	0.411	3.563	0.81
Eye, ear, nose, throat, larynx	odds ratio	2.776	0.313	133.243	0.67
Upper gastrointestinal	odds ratio	2.525	0.632	14.780	0.26
Lower gastrointestinal	odds ratio	infinite	0.327	infinite	0.30
Other genitourinary	odds ratio	1.161	0.294	5.577	1
Musculo-skeletal-integumentary	odds ratio	0.745	0.232	2.543	0.59
Neurological	odds ratio	0.780	0.233	2.874	0.77
Endocrine-metabolic	odds ratio	1.674	0.627	4.767	0.37
CIRS total score [d]	Hedges' g	0.227	-0.214	0.668	0.27
Body mass index [c]	Hedges' g	0.007	-0.407	0.422	0.97

[a] Positive Hedges' g means predictor is greater among responders. Odds ratio > 1 means predictor is more common among responders. [b] Calculated using Welch two-sample t test (continuous predictors) or Fisher exact test (dichotomous predictors). [c] Log transformed. [d] Square-root transformed.