Journal of Psychiatric Research 78 (2016) 94–102

Contents lists available at ScienceDirect

Journal of Psychiatric Research

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

Combining clinical variables to optimize prediction of antidepressant treatment outcomes

Raquel Iniesta [a, *], Karim Malki [a], Wolfgang Maier [b], Marcella Rietschel [c], Ole Mors [d], Joanna Hauser [e], Neven Henigsberg [f], Mojca Zvezdana Dernovsek [g], Daniel Souery [h], Daniel Stahl [i], Richard Dobson [a], Katherine J. Aitchison [a, j], Anne Farmer [a], Cathryn M. Lewis [a], Peter McGuffin [a], Rudolf Uher [a, k]

[a] Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, 16 De Crespigny Park, Denmark Hill, London, SE5 8AF, UK
[b] Department of Psychiatry, University of Bonn, Regina-Pacis-Weg 3, 53113 Bonn, Germany
[c] Central Institute of Mental Health, Division of Genetic Epidemiology in Psychiatry, Square J5, 68159, Mannheim, Germany
[d] Research Department P, Aarhus University Hospital, Norrebrogade 44, Aarhus C, DK-8000, Risskov, Denmark
[e] Laboratory of Psychiatric Genetics, Department of Psychiatry, Poznan University of Medical Sciences, Collegium Maius, Fredry 10, 61-701, Poznań, Poland
[f] Croatian Institute for Brain Research, Medical School, University of Zagreb, Salata 3, 10 000, Zagreb, Croatia
[g] University Psychiatric Clinic and the Medical Faculty, University of Ljubljana, Kongresni trg 12, 1000, Ljubljana, Slovenia
[h] Laboratoire de Psychologie Médicale, Université Libre de Bruxelles and Psy Pluriel – Centre Européen de Psychologie Médicale, Av Jack Pastur 47a, 1180, Uccle, Belgium
[i] Institute of Psychiatry, Psychology and Neuroscience, Kings College London, 16 De Crespigny Park, London, SE5 8AF, UK
[j] Department of Psychiatry and Medical Genetics, University of Alberta, 116 St and 85 Ave, Edmonton, AB, T6G 2R3, Canada
[k] Dalhousie University Department of Psychiatry, 5909 Veterans' Memorial Drive, Halifax, B3H 2E2, Nova Scotia, Canada

ARTICLE INFO

Article history:
Received 22 October 2015
Received in revised form
12 March 2016
Accepted 30 March 2016

Keywords:
Depression
Outcome
Antidepressant
Prediction
Machine learning
Statistical learning

ABSTRACT

The outcome of treatment with antidepressants varies markedly across people with the same diagnosis. A clinically significant prediction of outcomes could spare the frustration of trial and error approach and improve the outcomes of major depressive disorder through individualized treatment selection. It is likely that a combination of multiple predictors is needed to achieve such prediction. We used elastic net regularized regression to optimize prediction of symptom improvement and remission during treatment with escitalopram or nortriptyline and to identify contributing predictors from a range of demographic and clinical variables in 793 adults with major depressive disorder. A combination of demographic and clinical variables, with strong contributions from symptoms of depressed mood, reduced interest, decreased activity, indecisiveness, pessimism and anxiety significantly predicted treatment outcomes, explaining 5–10% of variance in symptom improvement with escitalopram. Similar combinations of variables predicted remission with area under the curve 0.72, explaining approximately 15% of variance (pseudo R^2) in who achieves remission, with strong contributions from body mass index, appetite, interest-activity symptom dimension and anxious-somatizing depression subtype. Escitalopram-specific outcome prediction was more accurate than generic outcome prediction, and reached effect sizes that were near or above a previously established benchmark for clinical significance. Outcome prediction on the nortriptyline arm did not significantly differ from chance. These results suggest that easily obtained demographic and clinical variables can predict therapeutic response to escitalopram with clinically meaningful accuracy, suggesting a potential for individualized prescription of this antidepressant drug.

1. Introduction

Major depressive disorder is a common condition, responsible for a substantial proportion of disability world-wide (Whiteford et al., 2013). Although a number of pharmacological and

* Corresponding author.
E-mail address: raquel.iniesta@kcl.ac.uk (R. Iniesta).

http://dx.doi.org/10.1016/j.jpsychires.2016.03.016

psychological treatment options are available, the outcomes are unsatisfactory. While some individuals experience dramatic improvements, most do not benefit sufficiently from the first treatment and have to undergo multiple treatment trials. Each trial takes weeks, with delays causing frustration, prolonging disability and risking adverse outcomes, including suicide. The unsatisfactory state of depression therapeutics has led to the consensus that diagnosis of depression may not be sufficient for treatment selection and additional information needs to be considered to estimate which treatment is likely to work for whom (Kupfer et al., 2012).

There is little evidence to guide clinicians in selecting a treatment for a given individual (Simon and Perlis, 2010). A single piece of information is unlikely to predict treatment outcome with an accuracy that is meaningful in clinical practice. Therefore, multiple factors may have to be considered to make the best prediction of outcomes at the individual level. The need for prediction at individual level has prompted the use of new methods, such as machine learning and statistical learning (Hastie et al., 2009). Unlike traditional statistics that focus on testing whether a single variable makes a statistically significant contribution, learning methods consider all available information across a number of variables to make the best prediction for an individual. The accuracy of prediction can then be compared to a standard benchmark to evaluate whether it is likely to be clinically significant (Uher et al., 2012d), i.e. whether it makes a meaningful difference to a particular individual.

Individualized treatment selection could be useful if it is based on predictors that are easily obtained (e.g. questionnaires and rating scales) and if it can differentially predict outcomes with alternative treatments. Two prior studies suggest that meaningful prediction of treatment outcomes from easy-to-obtain variables is achievable. A study of the STAR*D cohort found that 48 demographic and clinical variables robustly predicted treatment success with a clinically significant effect size (area under the curve 0.71, 11.4% variance explained) (Perlis, 2013). The prediction was robust in stringent validation test. A second study found that the relative benefits of cognitive-behavioural therapy and antidepressant medication can be predicted from eight demographic and clinical variables in a way that makes a meaningful difference in outcomes for 60% of 154 participants (DeRubeis et al., 2014). While both studies show promising results, they also leave caveats. The STAR*D study predicted overall outcome rather than outcomes of specific treatments. The strongest predictor was race, raising questions about how the findings generalize to populations with different ethnic composition. The study of cognitive-behavioural therapy and antidepressants established differential prediction, but due to a limited sample size, it had to derive a small number of predictors based on results obtained in the same sample and relied on a less stringent leave-one-out cross-validation.

Therefore, in the present study we evaluate to what extend can demographic and clinical variables predict outcomes with specific treatments at the level of individual. We have applied statistical learning to a study comparing treatment with two different antidepressants in an ethnically homogeneous sample large enough to allow robust 10-fold-split-sample cross-validation and permutations (Kohavi, 1995; Perez-Guaita et al., 2015).

2. Materials and methods

2.1. Study design and sample

The Genome-based Therapeutic Drugs for Depression (GENDEP) is a 12-week comparative study that aims to personalize treatment choice in major depressive disorder using clinical and genetic predictors of response to a serotonin-reuptake-inhibiting antidepressant escitalopram and a norepinephrine-reuptake-inhibiting antidepressant nortriptyline (Uher et al., 2009a, 2010). GENDEP included 868 treatment-seeking adults of White-European ethnicity from nine centers, diagnosed with ICD-10/DSM-IV major depressive disorder and a current depressive episode of at least moderate severity established with the Schedules for Clinical Assessment in Neuropsychiatry (SCAN) interview (Wing et al., 1990). Exclusion criteria were personal or family history of bipolar disorder or schizophrenia and active substance dependence. Eligible patients with no contraindications were randomly allocated to receive treatment with one of the two antidepressants for 12 weeks. Escitalopram is a selective serotonin reuptake inhibitor (SSRI) and has no effect on norepinephrine reuptake. Nortriptyline is a second-generation tricyclic antidepressant (TCA) with a much higher affinity for the norepinephrine transporter than for the serotonin transporter. A protocol guided treatment with escitalopram 10–30 mg daily and nortriptyline 50–150 mg daily, adjusted according to therapeutic effect and tolerability (Uher et al., 2009a). Participants with contraindications or history of intolerance of one of the drugs were offered treatment with the other drug non-randomly (Uher et al., 2009a). Seventy-six percent of GENDEP participants remained on the allocated antidepressant for 8 weeks or longer. In the present study, we include 793 participants (328 on nortriptyline and 465 on escitalopram), who had four or more depression severity measurements, a minimum needed to establish at least an initial trend in clinical response. Since participants non-randomly allocated to escitalopram and nortriptyline differed on some clinical characteristics (Uher et al., 2009a), we also repeated analyses restricting the sample to randomly allocated participants ($n = 450$) to provide drug-specific estimates in comparable samples. The ethics boards of all centers approved the protocol and all participants signed an informed consent.

2.2. Outcomes

The clinician-rated Montgomery–Åsberg Depression Rating Scale (MADRS) (Montgomery and Asberg, 1979), the 17-item Hamilton Rating Scale for Depression (HRSD) (Hamilton, 1967) and the self-report Beck Depression Inventory (BDI) (Beck et al., 1961) were administered at baseline and then weekly for 12 weeks with high inter-rater reliability (Uher et al., 2008, 2012c). Following a consensus reached in a meta-analysis (Investigators et al., 2013), we considered one primary continuous outcome and one primary categorical outcome. The primary continuous outcome was the percentage of improvement in MADRS score (the primary GENDEP outcome measure) over the twelve weeks, based on week twelve measurement if available and on the mixed effects model best unbiased linear estimate from earlier measurements if the week twelve measurement was missing, adjusted for center of recruitment, age and sex (Uher et al., 2010). On average, GENDEP participants improved by 56.2%, from a mean initial MADRS score of 29.0 to a mean end-of-treatment MADRS score of 12.7 (Uher et al., 2009a). The primary categorical outcome was remission, defined as a HRSD score of 7 or less on the last available measurement without imputation (we have selected the HRSD since this is the most established definition of remission; there is less agreement about which cut-off on the MADRS should be used as a threshold for defining remission). Secondary continuous and categorical outcomes included completion of an adequate treatment trial (six weeks or more on allocated antidepressant) and treatment resistant depression (TRD; lack of response to two adequate antidepressant treatment trials, including the GENDEP treatment and previous treatment trials). Of the analyzed sample, 326 (41.1%) participants achieved remission on HRSD-17, 710 (89.5%) completed an adequate treatment trial and 105 (13.3%) had TRD.