Published in final edited form as:

Behav Ther. 2017 July ; 48(4): 490–500. doi:10.1016/j.beth.2017.02.001.

Advancing Personalized Medicine: Application of a Novel Statistical Method to Identify Treatment Moderators in the Coordinated Anxiety Learning and Management Study

Andrea N. Niles, M.A.¹, Amanda G. Loerinc, MA¹, Jennifer L. Krull, Ph.D.¹, Peter Roy-Byrne, M.D.², Greer Sullivan, M.D.³, Cathy Sherbourne, Ph.D.⁴, Alexander Bystritsky, M.D.⁵, and Michelle G. Craske, Ph.D.^{1,5}

¹University of California, Los Angeles, Department of Psychology

²University of Washington at Harborview Medical Center, Center for Healthcare

³University of Arkansas for Medical Sciences, Department of Psychiatry

⁴RAND Corporation, Santa Monica, CA

⁵University of California, Los Angeles, Department of Psychiatry and Biobehavioral Sciences

Abstract

Objective—There has been increasing recognition of the value of personalized medicine where the most effective treatment is selected based on individual characteristics. This study used a new method to identify a composite moderator of response to evidence-based anxiety treatment (CALM) compared to Usual Care.

Method—Eight hundred seventy-six patients diagnosed with one or multiple anxiety disorders were assigned to CALM or Usual Care. Using the method proposed by Kraemer (2013), thirty-five possible moderators were examined for individual effect sizes then entered into a forward-stepwise regression model predicting differential treatment response. *K*-fold cross validation was used to identify the number of variables to include in the final moderator.

Results—Ten variables were selected for a final composite moderator. The composite moderator effect size ($r = .20$) was twice as large as the strongest individual moderator effect size ($r = .10$). Although on average patients benefitted more from CALM, 19% of patients had equal or greater treatment response in Usual Care. The effect size for the CALM intervention increased from $d = .34$ to $d = .54$ when accounting for the moderator.

Conclusions—Findings support the utility of composite moderators. Results were used to develop a program that allows mental health professionals to prescribe treatment for anxiety based on baseline characteristics (<http://anxiety.psych.ucla.edu/treatmatch.html>).

Corresponding Author: Andrea N. Niles M.A., Department of Psychology, UCLA, 405 Hilgard Avenue, Los Angeles, CA 90095-1563, Tel.: 1 415 407 0988; aniles@ucla.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

anxiety disorders; CBT; primary care; moderators; model selection; personalized medicine; precision medicine

Although cognitive behavioral therapies are effective for the treatment of anxiety disorders (Hofmann & Smits, 2008; Norton & Price, 2007), response rates range from 38% to 77% depending on disorder, indicating that many patients do not respond (Hofmann, Asnaani, Vonk, Sawyer, & Fang, 2012). The need for personalized medicine, where the most effective treatment is selected based on individual characteristics (i.e. treatment moderators), has received increasing attention (Simon & Perlis, 2010) as evidenced by the \$215 million Precision Medicine Initiative (National Institute of Health, 2016). However, the research on moderators for anxiety disorder treatment is severely lacking due to insufficient sample sizes and underpowered statistical methods (Schneider, Arch, & Wolitzky-Taylor, 2015). Prior moderation analyses typically have compared response to two evidence-based treatments. This is a valuable question given that multiple therapeutic approaches may be effective. At the same time, many patients do not have access to evidence-based treatment (Stein et al., 2011; Weisberg, Dyck, Culpepper, & Keller, 2007). As reviewed by Kazdin (2015), a more basic question, and one that may inform the current treatment climate is whether there is additional benefit for a patient in receiving specialized anxiety treatment compared to what is already in use – i.e. who benefits from additional treatment beyond usual care.

Kraemer (2008, 2013) discusses problems with existing research on treatment moderation including disagreement over the definition of moderators and a failure to define a moderator effect size. Consequently, it is not possible to determine which moderators have the largest effect on the likelihood of benefiting from a treatment. The addition of effect sizes to moderator studies allows comparison of moderator effect size for the same treatment choice and outcome (Kraemer, 2013). Another limitation with existing approaches is that researchers typically examine one moderator per statistical model (Schneider et al., 2015). Because moderators tend to have modest effect sizes (Kraemer, 2013), combining multiple moderators into one model should produce larger effects and ultimately greater precision when assigning patients to treatments. Kraemer's approach (2013) uses effect size rather than statistical significance to directly compare moderator importance. In addition, her approach combines individual moderators into a composite, which results in greater power to predict which treatment will be most effective.

In this study, we evaluated moderators of response to an intervention comprised of computer-assisted CBT or psychotropic medications, jointly termed the Coordinated Anxiety Learning and Management program (CALM), or treatment as usual in primary care (Craske et al., 2011; Roy-Byrne et al., 2010). In this study, a moderator is defined as a baseline variable that helps identify on whom or under what conditions treatment has a causal effect on outcome (2013). CALM was more effective overall than Usual Care (UC) in reducing symptoms of anxiety and depression and improving functioning (Craske et al., 2011; Roy-Byrne et al., 2010). Although patients assigned to CALM showed superior outcomes than patients in UC, CALM was not effective for all patients. Furthermore, some

patients assigned to UC improved. These patterns of response highlight the need to identify moderators of response to CALM versus UC. Because CBT can be expensive and time-intensive (Barlow, Levitt, & Bufka, 1999) and medications can have side effects, identifying characteristics that help primary care providers determine whether to refer patients to an anxiety specialist or to continue with treatment as usual is critical.

In a recent review, Schneider and colleagues (2015) identified 24 studies that examined moderators of treatment for anxiety. Of those studies, 4 had a large sample size, 15 used high quality statistics, and only 1 had both. The studies compared a variety of treatments including CBT, mindfulness-based treatments, psychodynamic psychotherapy, pharmacotherapy, and capnometry-assisted respiratory training. The results were variable across studies, and the authors state that few conclusions can be drawn given the paucity of research, the limited sample sizes, and the high variability in terms of treatments compared and moderators assessed. The authors conclude that if our goal is to inform clinicians about how to match patients to treatments, the methodological quality and consistency across studies needs improvement.

To our knowledge, no prior studies have examined moderators of response to evidence-based treatment versus usual care. Thus we briefly review existing literature on predictors of response to CBT and medication for anxiety disorders. Prior studies have shown that race, ethnicity, gender, and socio economic status (Piacentini, Bergman, Jacobs, McCracken, & Kretchman, 2002; Schuurmans et al., 2009; Watanabe et al., 2010; Wolitzky-Taylor, Arch, Rosenfield, & Craske, 2012) are unrelated to treatment outcome with one exception for gender (Craske et al., 2014) where we found that women with social phobia had better outcomes than men from acceptance and commitment therapy and CBT (regardless of treatment condition). In terms of clinical variables, findings for the effect of baseline disorder severity on outcome are mixed with some studies showing that higher baseline scores predict poorer outcome (Kampman, Keijsers, Hoogduin, & Hendriks, 2008; Wergeland et al., 2016), and others showing no predictive effect (Watanabe et al., 2010; Wolitzky-Taylor et al., 2012). Comorbid depression has predicted worse treatment outcome in some studies (Chambless, Beck, Gracely, & Grisham, 2000; Steketee, Chambless, & Tran, 2001; Watanabe et al., 2010), but not others (Kampman et al., 2008; Rief, Trenkamp, Auer, & Fichter, 2000; Schuurmans et al., 2009; van Balkom et al., 2008), and comorbidity of other non-mood disorders generally has not significantly predicted treatment outcome (Kampman et al., 2008; Ollendick, Ost, Reuterskiöld, & Costa, 2010; Wolitzky-Taylor et al., 2012). Few studies have examined predictors of response to medications for anxiety, but in studies that have, earlier age of disorder onset, baseline disorder severity, and comorbidity of psychiatric conditions predicted worse outcomes in obsessive-compulsive disorder and post-traumatic stress disorder (Ackerman, Greenland, Bystritsky, Morgenstern, & Katz, 1994; Davidson et al., 1993), and no significant effects were found for age, gender, physiologic variables (e.g. heart rate), and symptom severity in social phobia (Stein, Stein, Pitts, Kumar, & Hunter, 2002). Taken together, these findings suggest that demographic features are generally unrelated to outcome whereas indicators of disorder severity and comorbidity may predict worse outcomes.

Two previous studies have employed Kraemer's method for identifying moderator effect size and created a combined moderator with promising results (Niles, Wolitzky-Taylor, Arch, & Craske, 2017; Wallace, Frank, & Kraemer, 2013). Wallace and colleagues compared psychotherapy to medication in treating depression. The authors examined 32 possible moderators of treatment response, and then used a data reduction method (principal component analysis) to identify the number of factors to include in the combined moderator. The final moderator included eight variables with an effect size of $r = .31$, which was larger than the largest individual moderator effect size of $r = .12$. Niles and colleagues (2017) compared acceptance and commitment therapy and CBT. They examined 26 possible moderators of treatment dropout, and then used a statistical learning method (stepwise regression with cross validation) to identify the number of factors to use in the combined moderator. Their final moderator included four variables with an effect size of $r = .28$, which was larger than the largest individual moderator effect size of $r = .17$. In these prior studies, the composite moderator approach results in larger effect sizes compared to individual moderators and thereby improves our ability to prescribe treatments and enhance personalized medicine. Notably, the composite moderator approach is a purely predictive exercise. The objective is to find the best composite moderator to predict differential treatment response rather than to develop a model based on theory or to test a substantive hypothesis about particular moderators. The advantage lies in the possibility that theoretical models may miss unexpected factors that moderate treatment response.

In the current study, we followed our prior approach (Niles et al., 2017) of identifying all variables representing unique constructs that could relate to treatment outcome and then reducing and combining those variables using forward stepwise regression and cross validation to create a composite moderator. After reviewing the principal component analysis and machine learning moderator selection approaches used in the prior two studies, we determined that machine learning includes fewer subjective and arbitrary decisions by the researcher and would thus result in more replicable results. In addition, with principal component analysis, the researcher identifies variables to include in the combined moderator based on zero-order correlations. When multiple variables are combined in a regression, each variable's relationship with outcome may differ substantially from zero-order correlations. The machine learning approach identifies new variables to include after controlling for effects of variables already in the model. Thus, we followed the forward stepwise regression with cross validation method (machine learning) in the current study. The goal of the current project was to identify a combined moderator of symptom reduction in CALM compared to UC and to make our results accessible to health-care providers for use in treatment prescription.

Method

Participants

Participants were 1004 primary care patients meeting criteria for panic (with or without agoraphobia), generalized anxiety, social anxiety, or posttraumatic stress disorders. Further details are presented elsewhere (Sullivan et al., 2007). The sample in the current report

included 876 participants with non-missing values on the dependent measure at baseline and 6-month follow-up, and 430 patients (49%) were allocated to UC and 446 (51%) to CALM.

Inclusion criteria—All eligible participants were patients in participating clinics; between ages 18 and 75; met DSM-IV criteria for the disorders mentioned above (all diagnoses made by the Mini International Neuropsychiatric Interview; MINI) (Lecrubier et al., 1997) administered by an Anxiety Clinical Specialist; and received a score of at least 8 (indicating moderate and clinically significant anxiety symptoms on a scale from 0–20) on the Overall Anxiety Severity and Impairment Scale (Norman et al., 2011).

Exclusion criteria—Patients were ineligible if they had unstable medical conditions, cognitive impairment, suicidality, psychosis or bipolar I disorder. Individuals with alcohol or marijuana abuse were included, however, those with substance dependence and other drug abuse were ineligible. Patients receiving CBT or who did not speak English or Spanish were excluded.

Materials

Moderators—We identified all variables available in the dataset that represented non-overlapping constructs and that could sensibly be related to treatment outcome, which resulted in 38 variables. We examined collinearity among these variables, and removed three with variance inflation factors over 2.5, resulting in 35 possible moderators. Similar to moderators and predictors examined in prior research, these moderators fell loosely into three categories: clinical characteristics, attitudes toward treatment, and demographics. Clinical characteristics included anxiety and depression symptom severity, psychological and medical comorbidity, physical functioning, medication use, alcohol use, and past psychological treatment. Attitudes toward treatment included beliefs about the effectiveness of psychological treatment, stigma of psychological treatment, preferences for psychological treatment, and satisfaction with current care. Demographics included age, gender, ethnicity, socio-economic status, education level, employment, family size, marital status, and insurance status. Additional variables examined included experience of discrimination and social support. For a full list of variables, see Table 1.

Anxiety and depression comorbidity and symptom severity were assessed with the MINI (Lecrubier et al., 1997), the Anxiety Sensitivity Index (Reiss, Peterson, Gursky, & McNally, 1986), and the Patient Health Questionnaire (Kroenke, Spitzer, & Williams, 2001) respectively. Physical functioning was assessed using the 12-item Short Form Health Questionnaire (Ware, Kosinski, & Keller, 1996). Social support was assessed using the RAND Social Support Survey Instrument (Sherbourne & Stewart, 1991). Alcohol use was assessed using the first three items of the Alcohol Use Disorders Identification Test (Saunders, Aasland, Babor, de la Fuente, & Grant, 1993). Additional variables were assessed with single items or with measures from the National Comorbidity Survey Replication (Kessler, Chiu, Demler, Merikangas, & Walters, 2005).

Outcome Measure—The dependent measure was the Brief Symptom Inventory General Severity Index (Derogatis, 2001). Outcome scores were residuals obtained from a regression

equation predicting the participants' BSI scores at 6 months from baseline BSI scores allowing us to control for baseline anxiety severity and consequently assess moderator effects on change in anxiety symptoms. Because the moderator methodology has not yet been adapted to statistical methods used to analyze multiple time points simultaneously (i.e. multi-level modeling), and because participants showed the greatest symptom improvement from baseline to 6 months, we chose to focus the analysis on only the post-treatment data at 6 months.

Randomization

After a baseline assessment, participants were randomized to the CALM intervention or UC. For more details about patient flow from referral through eligibility screening, consent, and randomization, see Roy-Byrne et al (2010).

Intervention (CALM)—Individuals in CALM had the choice to receive medication only (9%), CBT only (34%), or both (57%) over a 10- to 12-week period. The CBT program was called *CALM Tools for Living* (Craske et al., 2009), and covered the four primary anxiety disorders. It was targeted to each patient's most distressing and impairing anxiety disorder. The CBT program included generic modules, including self monitoring, psychoeducation, fear hierarchy, breathing retraining, and relapse prevention, and modules that were modified for each anxiety disorder (cognitive restructuring and exposure). CBT was computer-assisted to guide the novice therapist and patient as they sat side-by-side viewing the program on the computer. After the initial 10–12 weeks, participants who remained symptomatic had the choice to receive more of the same modality (CBT or medication) or the alternative modality for up to three more steps (another 10–12 weeks). Participants who chose medication had their medication adherence monitored.

Usual Care (UC)—Participants randomized to UC received whatever treatment the primary care provider administered normally. No restrictions were placed on providers in the UC condition. Patients reported upon treatment received. Patients assigned to UC reported the following changes in their treatment from baseline to 6 months following study enrollment: pharmacotherapy (17%), CBT (12%), or any type of counseling (4.3%) (Roy-Byrne et al., 2010). Because therapists in the CALM intervention were supervised by experts in CBT for anxiety, and because prior research has shown that only 10% of anxious patients in general receive therapy that includes the major components of CBT (Stein et al., 2004), CBT received by patients in the UC group was likely of lesser quality than that received by patients in CALM. After baseline assessment, participants in UC did not have contact with study personnel with the exception of phone assessments.

Procedure

After informed consent, measures were administered at baseline, 6, 12, and 18 months via telephone by the Research and Development (RAND) Survey Research Group. Only baseline and 6 month time points were used in the current study. Measures given at baseline were examined as moderators and the BSI, administered at baseline and 6-months, was used as the measure of treatment outcome.

Statistical Analyses

The statistical analyses combined methods described by Kraemer (2013) and James, Witten, Hastie and Tibshirani (2013). We used the method proposed by Kraemer to examine effect sizes for all possible moderators and to create the combined moderator. We used model selection combined with k -fold cross-validation with 5 folds as described by James and colleagues (2013) to identify the model used to calculate the combined moderator.

Combined Moderator Model Selection Using k-Fold Cross-Validation—We first identified 35 possible variables to examine as moderators. We then randomly divided the 876 observations into five folds of roughly equal size. Four folds were used as the training dataset (80% of the data) and the remaining one fold was used as a validation dataset (20%).

Following the method of Kraemer (2013), in the training datasets, we paired every patient assigned to CALM with every patient assigned to UC. As indicated by Kraemer, in the paired dataset, $r(O, AM)$, or the correlation between the difference in outcome and the average moderator for each pair, is the moderator effect size. This produces an effect size equivalent to a correlation coefficient that falls between -1 and $+1$, with null value 0 and greater absolute magnitudes indicating stronger moderation. Therefore, within the paired dataset, for all 35 moderators assessed, we calculated the average value of the moderator and the difference in the outcome for each pair in the dataset. We then calculated correlations between the difference in the outcome and each of the averaged moderators to determine the *individual* effect size for each of our moderators.

To build the combined moderator, in the training dataset, we ran forward stepwise regression predicting the difference in outcome from the average across pairs of each of the 35 moderators. Rather than using a p value cut off for inclusion in the model, we allowed the regression to continue adding predictors until all 35 variables had been included, resulting in 35 models with increasing numbers of variables from 1 to 35. This stepwise approach allowed us to subsequently examine the fit of all 35 models. Using the *unpaired* validation datasets (i.e. the original dataset with one variable representing treatment group and the 35 standardized moderators), we used the unstandardized beta coefficients from each of the 35 models to create 35 composite moderators M_1 through M_{35} .

In the *unpaired* validation datasets:

$$\begin{aligned} M_1 &= B_1 X_1 \\ M_2 &= B_1 X_1 + B_2 X_2 \\ &\dots \\ M_{35} &= B_1 X_1 + B_2 X_2 \dots B_{35} X_{35} \end{aligned}$$

where i is the model number, M_i is the combined moderator, B_1 through B_i are the beta coefficients obtained from the regression model, and X_1 through X_i are the individual moderators included in the regression model.

After creating 35 composite moderators M_1 through M_{35} , we ran a regression model in which the dependent variable was treatment outcome (BSI), and the predictors were Group

(CALM = 1, UC = 0), M_j , and Group $\times M_j$. We calculated the mean squared error (MSE) for each of the 35 regression models using the equation:

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2$$

We repeated the entire above procedure five times, training the models on 80% of the data and testing the models on the remaining 20%, each time using a different 20%. We then averaged the mean squared error across each of the five validation datasets for each of the models and plotted the MSE for moderators M_1 through M_{35} resulting in a curve representing the fit of models 1 through 35 in the validation datasets (results displayed in Figure 1). We then identified the minimum on the curve to chose a model that would minimize MSE (James et al., 2013).

Once we identified the number of variables to include in the final model, we used forward-stepwise regression on the full data, including only the number of variables identified by the k -fold cross-validation results. Following the method described by Kraemer (2013), we then used the beta coefficients obtained from the model to create our combined moderator (M^*). Finally, to estimate the effect size of the combined moderator, we conducted a regression analysis predicting anxiety outcome from the combined moderator M^* , treatment group, and the interaction between M^* and treatment group, graphed the results, and characterized those that would benefit more from one treatment over the other.

Results

The learning curve resulting from the k -fold cross-validation algorithm is displayed in Figure 1. This graph depicts the model fit (with lower numbers representing better fit) of the 35 models with increasing numbers of variables from 1 to 35. Because the curve minimum occurred at model 10, we included 10 variables in our final model. Compared to a model using only intervention group as the predictor, in the validation datasets, the inclusion of the Group by M^* interaction, where M^* was a composite of 10 variables, increased Cohen's d from .29 to .38. This is a conservative estimate of the predicted effect size of our moderator for new data.

Independent effect sizes for each of our moderators are displayed in Table 1. Effect sizes ranged from $-.002$ (alcohol use) to $.096$ (anxiety sensitivity) with negative numbers indicating better outcome in UC than CALM for higher values of the moderator and positive numbers indicating better outcome in CALM than UC for higher values of the moderator. We then ran forward stepwise regression on the full dataset (as described in the methods) to identify the first ten variables that would be included in the final model. The variables selected in order from first to last were anxiety sensitivity, gender, greater than high school education (compared to high school education), number of non-psychotropic medications, perceived social standing, less than high school education (compared to high school education), satisfaction with medical care, depression, number of psychotropic medications, and African American race (compared to any other ethnicity). We visually inspected a

histogram of the composite moderator and it was normally distributed. Weights used to calculate the combined moderator are shown in Table 1. The combined moderator effect size (for direct comparison to individual moderator effects) was .20.

To estimate the extent to which the combined moderator improved our ability to predict outcome, we conducted a regression analysis (using the original unpaired data) predicting anxiety outcome from M^* , Group, and the interaction between M^* and Group. Group explained 2.8% of the variance in the outcome (Cohen's $d = .34$), and inclusion of the main effect of M^* and the interaction between Group and M^* more than doubled the variance explained to 6.9%. (Cohen's $d = .54$).

We next compared the combined moderator to the largest individual moderator (anxiety sensitivity). Using “seemingly unrelated regression” (Zellner, 1962) to combine non-nested regression models, we compared the model testing the interaction of M^* with Group to the model testing the interaction of anxiety sensitivity with Group (all main effects were also included). Using a Wald test, we compared the coefficients associated with the interaction term in the model. The beta associated with the $M^* \times$ Group term ($b = -.39$) was significantly larger than the beta associated with the Anxiety Sensitivity $\times$ Group term ($b = -.19$) $\chi^2(1, N = 876) = 6.91, p = .009$.

The predicted lines for the two treatment groups crossed at $M^* = -.34$. Results are displayed in Figure 2. Above the cut point, anxiety reduction was greater in CALM than UC, and below the cut point, anxiety reduction was greater in UC than CALM. Approximately 19% of the sample fell below the cut point, and 81% fell above the cut point.

To characterize participants above and below the cut point for M^* , in Table 2, we report descriptive statistics separately for the ten moderators included in the calculation of M^* .

Discussion

The goal of the current study was to implement a new statistical approach to identifying patient characteristics that predict better response to one treatment over another. We compared computer-assisted CBT and/or medications to usual care for the treatment of anxiety disorders. We used the method described by Kraemer (2013) to examine 35 purported moderators then developed a combined moderator (M^*) using 10 of the 35 variables.

Consistent with findings from Wallace et al. (2013) and Niles et al. (2017), who identified combined moderators of response to psychotherapy compared to medication for depression and dropout from CBT and acceptance and commitment therapy for anxiety respectively, our combined moderator effect size ($r = .20$) was twice as large as that of any individual moderator (largest $r = .10$). The relatively small moderator effect sizes for individual variables render prescriptions based on a single variable problematic. In addition, testing individual moderators only allows for the use of one characteristic in treatment prescription at a time, whereas when moderators are combined, multiple characteristics can be used simultaneously and inter-correlations between them are accounted for. Our findings show that the composite moderator approach advances personalized medicine significantly

because it allows us to harness more predictive value from moderators by combining them into a composite.

In addition to increasing moderator power, this statistical approach allowed us to identify a cut point for M^* above which anxiety reduction was greater in CALM than in UC, and below which, anxiety reduction was greater in UC than CALM. M^* is analogous to the Personalized Advantage Index (DeRubeis et al., 2014) because values of M^* can be used to determine which of two treatments should be prescribed and the extent to which the prescribed treatment should produce superior outcomes over the alternative. The cut point for M^* was $-.34$, which fell in the observed range for M^* with 19% of the sample below and 81% above the cut point. Although the majority of the sample benefited more from CALM than UC, which is consistent with main outcomes from the study (Craske et al., 2011; Roy-Byrne et al., 2010), for the 19% of participants below the cut point, CALM was *no better than* or *worse than* UC. In addition, inclusion of the moderator in predicting treatment outcome increased the effect size of the intervention from $d = .34$ to $d = .54$ and more than doubled the variance explained in treatment outcome. Although these findings require confirmation in an independent dataset designed to test the composite moderator with new data, the M^* identified in this study may be valuable for identifying who will benefit most from an additional evidenced-based anxiety intervention and who will not.

Identification of a cut point also allowed us to describe individual characteristics of patients who fell above and below it, providing information about profiles of patients who may benefit more from CALM or UC. However, within the combined moderator, the contribution of each individual variable must be interpreted in conjunction with the other variables included in the moderator, meaning that each variable included in the composite was related to outcome over and above all other variables in the model. Patients who benefited more from CALM than UC tended to have more severe psychiatric symptoms (higher depression and anxiety scores), suggesting that an evidence-based treatment is indicated over usual care for those with more severe psychopathology. They also were less satisfied with current care, were less likely to be receiving psychotropic medications, and were less likely to have received education beyond highschool, indicating that referral to evidence-based treatment may be particularly beneficial when a person is not receiving adequate treatment from current providers or may not have the resources to seek out such treatment. Interestingly, in a prior paper, (Stein et al., 2011) we showed that receipt of quality psychotherapy was the only predictor of satisfaction with care, suggesting that patients who were more satisfied with care likely had previously received therapy with some components of CBT. Patients with lower perceived social standing, who were taking more medications for non-mental health concerns, and those with less than high school education benefitted less from CALM, perhaps suggesting that for those with concerns other than mental health such as physical illness or stress, referral to evidence-based anxiety treatment may be less beneficial. The finding that women benefitted more than men from CALM compared to UC and that white participants benefitted more than black participants from CALM compared to UC was unexpected given that neither gender nor ethnicity are typically predictors of treatment outcome (Piacentini et al., 2002; Schuurmans et al., 2009; Watanabe et al., 2010; Wolitzky-Taylor et al., 2012). It is important to note that the UC group was heterogenous in the

treatment received and that response could reflect placebo, spontaneous remission, or (less likely) response to eclectic treatment.

This was an exploratory study conducted to determine the feasibility of a newer methodological approach to identifying combined moderators in randomized controlled trials. When future trials validate this methodology, it is important to note that the M^* we developed is only one of many possible moderators as there are many combinations that may have been appropriate. Additionally, the M^* in the current study was based on baseline variables that were available to us. There may be other moderator variables that carry more weight than those we measured, such as personality, life stress, number of treatment sessions completed, or patient treatment preference (although presumably all patients who agreed to participate were amenable to receiving the CALM intervention). In addition, the inclusion of interactions between variables and non-linear effects may have increased the predictive power of our moderator. However, inclusion of these effects can be computationally challenging due to the number of possible combinations and could also increase the complexity of the model. It is necessary to balance model fit with model complexity (James et al., 2013), and k -fold cross-validation is one of multiple possible approaches. Finally, although the sample included patients with psychological and medical comorbidities, due to the exclusion criteria used for the study (e.g. bipolar disorder, psychosis, substance use) it is possible that different moderators would have emerged in the absence of these exclusion criteria.

In conclusion, findings from the current project provide strong evidence for the utility of this novel approach to identifying moderators. These findings can inform clinical decision-making, allowing mental health professionals to assign specific treatments to individuals based on baseline characteristics. Wallace and colleagues (2013), discussed the possibility of building a computer program designed for the input of baseline levels of relevant variables to assist in making treatment decisions. We have taken this step with the current findings and have designed a Treatment Matching Calculator, which is available online for use and can be completed by the patient with or without guidance from the treatment provider (<http://anxiety.psych.ucla.edu/treatmatch.html>). The current project takes a significant step in the direction of personalized medicine, and has produced a product that will help apply our findings directly to the improvement of patient care.

Acknowledgments

Grant Funding and Disclosures

Funding support for this project included grants MH57858 and MH065324 (Dr. Roy-Byrne), and MH58915-03 (Dr. Craske), U01 MH070022 (Dr. Sullivan), and U01 MH070018 (Dr. Sherbourne).

Dr. Sullivan reports grants and personal fees from NIMH during the conduct of the study; Dr. Stein reports grants from National Institute of Mental Health, from Veterans Affairs, and from Department of Defense during the conduct of the study; Dr. Sherbourne reports grants from NIMH during the conduct of the study. Dr. Roy-Byrne reports grants from NIMH during the conduct of the study; personal fees from Valiant Medical Solutions, personal fees from Editor and Chief Depression and Anxiety, personal fees from UpToDate Psychiatry & Journal Watch Psychiatry outside the submitted work; Ms. Niles has nothing to disclose; Dr. Craske reports grants and personal fees from NIMH during the conduct of the study; Ms. Loerinc has nothing to disclose; Dr. Krull has nothing to disclose; Dr. Bystritsky has nothing to disclose.

References

- Ackerman DL, Greenland S, Bystritsky A, Morgenstern H, Katz RJ. Predictors of treatment response in obsessive-compulsive disorder: Multivariate analyses from a multicenter trial of clomipramine. *Journal of Clinical Psychopharmacology*. 1994; 14(4):247–254. [PubMed: 7962680]
- Barlow DH, Levitt JT, Bufka LF. The dissemination of empirically supported treatments: A view to the future. *Behaviour Research and Therapy*. 1999; 37(Suppl 1):S147–162. [PubMed: 10402700]
- Chambless DL, Beck AT, Gracely EJ, Grisham JR. Relationship of cognitions to fear of somatic symptoms: A test of the cognitive theory of panic. *Depression and Anxiety*. 2000; 11(1):1–9. [PubMed: 10723629]
- Craske MG, Niles AN, Burklund LJ, Wolitzky-Taylor KB, Plumb C, Arch JJ, Lieberman MD. Randomized controlled trial of cognitive behavioral therapy and acceptance and commitment therapy for social phobia: Outcomes and moderators. *Journal of Consulting and Clinical Psychology*. 2014; 82(6):1034–1048. <https://doi.org/10.1037/a0037212>. [PubMed: 24999670]
- Craske MG, Rose RD, Lang A, Welch SS, Campbell-Sills L, Sullivan G, Roy-Byrne PP. Computer-assisted delivery of cognitive behavioral therapy for anxiety disorders in primary-care settings. *Depression and Anxiety*. 2009; 26(3):235–242. <https://doi.org/10.1002/da.20542>. [PubMed: 19212970]
- Craske MG, Stein MB, Sullivan G, Sherbourne C, Bystritsky A, Rose RD, Golinelli D. Disorder-specific impact of coordinated anxiety learning and management treatment for anxiety disorders in primary care. *Archives of General Psychiatry*. 2011; 68(4):378–388. [PubMed: 21464362]
- Davidson JR, Kudler HS, Saunders WB, Erickson L, Smith RD, Stein RM, Cavenar JO. Predicting response to amitriptyline in posttraumatic stress disorder. *The American Journal of Psychiatry*. 1993; 150(7):1024–1029. <https://doi.org/10.1176/ajp.150.7.1024>. [PubMed: 8317571]
- Derogatis, L. Brief Symptom Inventory 18: administration, scoring, and procedures manual. Minneapolis, MN: NCS Assessments; 2001.
- DeRubeis RJ, Cohen ZD, Forand NR, Fournier JC, Gelfand LA, Lorenzo-Luaces L. The personalized advantage index: Translating research on prediction into individualized treatment recommendations. A demonstration. *PLOS ONE*. 2014; 9(1):e83875. <https://doi.org/10.1371/journal.pone.0083875>. [PubMed: 24416178]
- Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*. 2012; 36(5):427–440. <https://doi.org/10.1007/s10608-012-9476-1>. [PubMed: 23459093]
- Hofmann SG, Smits JAJ. Cognitive-behavioral therapy for adult anxiety disorders: A meta-analysis of randomized placebo-controlled trials. *The Journal of Clinical Psychiatry*. 2008; 69(4):621–632. [PubMed: 18363421]
- James, G., Witten, D., Hastie, T., Tibshirani, R. An introduction to statistical learning with applications in R. New York, NY: Springer; 2013.
- Kampman M, Keijsers GPJ, Hoogduin CAL, Hendriks GJ. Outcome prediction of cognitive behaviour therapy for panic disorder: Initial symptom severity is predictive for treatment outcome, comorbid anxiety or depressive disorder, cluster C personality disorders and initial motivation are not. *Behavioural and Cognitive Psychotherapy*. 2008; 36(01):99–112. <https://doi.org/10.1017/S1352465807004018>.
- Kazdin, AE. Treatment as usual and routine care in research and clinical practice. *Clinical Psychology Review*. 2015. <https://doi.org/10.1016/j.cpr.2015.08.006>
- Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*. 2005; 62(6):617–627. <https://doi.org/10.1001/archpsyc.62.6.617>. [PubMed: 15939839]
- Kraemer HC. Toward non-parametric and clinically meaningful moderators and mediators. *Statistics in Medicine*. 2008; 27(10):1679–1692. <https://doi.org/10.1002/sim.3149>. [PubMed: 18008395]
- Kraemer HC. Discovering, comparing, and combining moderators of treatment on outcome after randomized clinical trials: A parametric approach. *Statistics in Medicine*. 2013; 32(11):1964–1973. <https://doi.org/10.1002/sim.5734>. [PubMed: 23303653]

- Kroenke K, Spitzer RL, Williams JB. The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*. 2001; 16(9):606–613. [PubMed: 11556941]
- Lecrubier Y, Sheehan D, Weiller E, Amorim P, Bonora I, Harnett Sheehan K, Dunbar G. The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: Reliability and validity according to the CIDI. *European Psychiatry*. 1997; 12(5):224–231. [https://doi.org/10.1016/S0924-9338\(97\)83296-8](https://doi.org/10.1016/S0924-9338(97)83296-8).
- National Institute of Health. All of us: The precision medicine initiative. 2016. Retrieved from <https://www.nih.gov/research-training/allofus-research-program>
- Niles AN, Wolitzky-Taylor KB, Arch JJ, Craske MG. Applying a novel statistical method to advance the personalized treatment of anxiety disorders: A composite moderator of comparative drop-out from CBT and ACT. *Behaviour Research and Therapy*. 2017; 91:13–23. <https://doi.org/10.1016/j.brat.2017.01.001>. [PubMed: 28110111]
- Norman SB, Campbell-Sills L, Hitchcock CA, Sullivan S, Rochlin A, Wilkins KC, Stein MB. Psychometrics of a brief measure of anxiety to detect severity and impairment: The Overall Anxiety Severity and Impairment Scale (OASIS). *Journal of Psychiatric Research*. 2011; 45(2): 262–268. <https://doi.org/10.1016/j.jpsychires.2010.06.011>. [PubMed: 20609450]
- Norton PJ, Price EC. A meta-analytic review of adult cognitive-behavioral treatment outcome across the anxiety disorders. *The Journal of Nervous and Mental Disease*. 2007; 195:521–531. <https://doi.org/10.1097/01.nmd.0000253843.70149.9a>. [PubMed: 17568301]
- Ollendick TH, Ost LG, Reuterskiöld L, Costa N. Comorbidity in youth with specific phobias: Impact of comorbidity on treatment outcome and the impact of treatment on comorbid disorders. *Behaviour Research and Therapy*. 2010; 48(9):827–831. <https://doi.org/10.1016/j.brat.2010.05.024>. [PubMed: 20573338]
- Piacentini J, Bergman RL, Jacobs C, McCracken JT, Kretchman J. Open trial of cognitive behavior therapy for childhood obsessive–compulsive disorder. *Journal of Anxiety Disorders*. 2002; 16(2): 207–219. [https://doi.org/10.1016/S0887-6185\(02\)00096-8](https://doi.org/10.1016/S0887-6185(02)00096-8). [PubMed: 12194545]
- Reiss S, Peterson RA, Gursky DM, McNally RJ. Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour Research and Therapy*. 1986; 24(1):1–8. [https://doi.org/10.1016/0005-7967\(86\)90143-9](https://doi.org/10.1016/0005-7967(86)90143-9). [PubMed: 3947307]
- Rief W, Trenkamp S, Auer C, Fichter MM. Cognitive behavior therapy in panic disorder and comorbid major depression. A naturalistic study. *Psychotherapy and Psychosomatics*. 2000; 69(2):70–78. <https://doi.org/10.1159/000012369>. [PubMed: 10671827]
- Roy-Byrne PP, Craske MG, Sullivan G, Rose RD, Edlund MJ, Lang AJ, Stein MB. Delivery of evidence-based treatment for multiple anxiety disorders in primary care. *The Journal of the American Medical Association*. 2010; 303(19):1921–1928. <https://doi.org/10.1001/jama.2010.608>. [PubMed: 20483968]
- Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption–II. *Addiction (Abingdon, England)*. 1993; 88(6):791–804.
- Schneider RL, Arch JJ, Wolitzky-Taylor KB. The state of personalized treatment for anxiety disorders: A systematic review of treatment moderators. *Clinical Psychology Review*. 2015; 38:39–54. <https://doi.org/10.1016/j.cpr.2015.02.004>. [PubMed: 25795293]
- Schuermans J, Comijs H, Emmelkamp PMG, Weijnen IJC, van den Hout M, van Dyck R. Long-term effectiveness and prediction of treatment outcome in cognitive behavioral therapy and sertraline for late-life anxiety disorders. *International Psychogeriatrics/IPA*. 2009; 21(6):1148–1159. <https://doi.org/10.1017/S1041610209990536>.
- Sherbourne CD, Stewart AL. The MOS social support survey. *Social Science & Medicine*. 1991; 32(6):705–714. [https://doi.org/10.1016/0277-9536\(91\)90150-B](https://doi.org/10.1016/0277-9536(91)90150-B). [PubMed: 2035047]
- Simon GE, Perlis RH. Personalized medicine for depression: Can we match patients with treatments? *The American Journal of Psychiatry*. 2010; 167(12):1445–1455. <https://doi.org/10.1176/appi.ajp.2010.09111680>. [PubMed: 20843873]
- Stein DJ, Stein MB, Pitts CD, Kumar R, Hunter B. Predictors of response to pharmacotherapy in social anxiety disorder: An analysis of 3 placebo-controlled paroxetine trials. *The Journal of Clinical Psychiatry*. 2002; 63(2):152–155. [PubMed: 11874217]

- Stein MB, Roy-Byrne PP, Craske MG, Campbell-Sills L, Lang AJ, Golinelli D, Sherbourne CD. Quality of and patient satisfaction with primary health care for anxiety disorders. *The Journal of Clinical Psychiatry*. 2011; 72(7):970–976. <https://doi.org/10.4088/JCP.09m05626blu>. [PubMed: 21367351]
- Stein MB, Sherbourne CD, Craske MG, Means-Christensen A, Bystritsky A, Katon W, Roy-Byrne PP. Quality of care for primary care patients with anxiety disorders. *The American Journal of Psychiatry*. 2004; 161(12):2230–2237. <https://doi.org/10.1176/appi.ajp.161.12.2230>. [PubMed: 15569894]
- Steketee G, Chambless DL, Tran GQ. Effects of axis I and II comorbidity on behavior therapy outcome for obsessive-compulsive disorder and agoraphobia. *Comprehensive Psychiatry*. 2001; 42(1):76–86. <https://doi.org/10.1053/comp.2001.19746>. [PubMed: 11154720]
- Sullivan G, Craske MG, Sherbourne C, Edlund MJ, Rose RD, Golinelli D, Roy-Byrne PP. Design of the Coordinated Anxiety Learning and Management (CALM) study: Innovations in collaborative care for anxiety disorders. *General Hospital Psychiatry*. 2007; 29(5):379–387. [PubMed: 17888803]
- van Balkom AJLM, van Boeijen CA, Boeke AJP, van Oppen P, Kempe PT, van Dyck R. Comorbid depression, but not comorbid anxiety disorders, predicts poor outcome in anxiety disorders. *Depression and Anxiety*. 2008; 25(5):408–415. <https://doi.org/10.1002/da.20386>. [PubMed: 17960642]
- Wallace ML, Frank E, Kraemer HC. A novel approach for developing and interpreting treatment moderator profiles in randomized clinical trials. *JAMA Psychiatry*. 2013; 70(11):1241–1247. <https://doi.org/10.1001/jamapsychiatry.2013.1960>. [PubMed: 24048258]
- Ware JE, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care*. 1996; 34(3):220. [PubMed: 8628042]
- Watanabe N, Furukawa TA, Chen J, Kinoshita Y, Nakano Y, Ogawa S, Noda Y. Change in quality of life and their predictors in the long-term follow-up after group cognitive behavioral therapy for social anxiety disorder: A prospective cohort study. *BMC Psychiatry*. 2010; 10(1):81. <https://doi.org/10.1186/1471-244X-10-81>. [PubMed: 20942980]
- Weisberg RB, Dyck I, Culpepper L, Keller MB. Psychiatric treatment in primary care patients with anxiety disorders: A comparison of care received from primary care providers and psychiatrists. *The American Journal of Psychiatry*. 2007; 164(2):276–282. <https://doi.org/10.1176/ajp.2007.164.2.276>. [PubMed: 17267791]
- Wergeland GJH, Fjermestad KW, Marin CE, Bjelland I, Haugland BSM, Silverman WK, Heiervang ER. Predictors of treatment outcome in an effectiveness trial of cognitive behavioral therapy for children with anxiety disorders. *Behaviour Research and Therapy*. 2016; 76:1–12. <https://doi.org/10.1016/j.brat.2015.11.001>. [PubMed: 26583954]
- Wolitzky-Taylor KB, Arch JJ, Rosenfield D, Craske MG. Moderators and non-specific predictors of treatment outcome for anxiety disorders: A comparison of cognitive behavioral therapy to acceptance and commitment therapy. *Journal of Consulting and Clinical Psychology*. 2012; 80(5) <https://doi.org/10.1037/a0029418>.
- Zellner A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*. 1962; 57(298):348–368. <https://doi.org/10.1080/01621459.1962.10480664>.

Highlights

- Identified composite moderator of response to EST for anxiety vs Usual Care
- Composite moderator effect size was twice that of the strongest individual moderator
- Composite moderator increased effect size from $d = .34$ to $.54$
- Findings support the utility of composite moderators

Figure 1. Learning curve depicting mean squared error in validation datasets from k -fold cross validation analysis with standard errors bars showing variability across the five validation datasets

Figure 2.

Predicted Standardized BSI Scores for Participants in the Usual Care and CALM Groups Across Observed Values of the Combined Moderator M^*

Note. The cross point of $M^* = -.34$ may indicate that people with values of M^* greater than $-.34$ may benefit more from the CALM intervention than usual care whereas those with values of M^* less than $-.34$ may benefit more from usual care than from CALM.

Table 1

Effect sizes and final model weights for moderators

Variable	Individual Effect Size	Weight in Final Model
Included in Final Model (in order of inclusion in the stepwise regression)		
Anxiety Sensitivity	.096	.155
Female	.072	.139
> HS Education	-.079	-.228
Number Non-Psychotropic Meds	-.065	-.118
Perceived Social Standing	.030	.116
< HS Education	-.054	-.185
Satisfaction With Medical Care	-.058	-.096
Patient Health Questionnaire	.071	.106
Number of Psychotropic Medications	-.051	-.113
Black Race	-.034	-.081
Not included in Final Model		
Income	-.069	
US Born	-.055	
Comorbid Multiple Anxiety and Depression	.047	
Satisfaction With Psychiatric Care	-.039	
Hispanic Race	.038	
Visits to Psychiatrist	.037	
Social Support	-.037	
Physical Functioning SF-12	-.036	
Comorbid Multiple Anxiety	-.031	
Previous CBT	.028	
Number Chronic Medical Conditions	-.026	
Use of Alternative Medicine	.020	
Belief in Medication	-.020	
Discrimination	.019	
Age	-.013	
European Quality of Life	-.012	
Comorbid One Anxiety and Depression	-.012	
Employed	-.008	
Comfort Getting Psych Help	.007	
Visits to Therapist	.007	
Number of Kids	.006	
Married	.005	
White Race	-.004	
Perceived Socioeconomic Standing	.004	
Has Insurance	-.003	
Belief in Therapy	-.003	
Alcohol Use	-.002	

Note. Positive effect sizes indicate that higher values of the moderator suggest preference for CALM over Usual Care and negative effect sizes indicate that higher values of the moderator suggest preference for Usual Care over CALM.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2Characteristics of participants above and below $M^* = -.34$

	Usual Care Preferable to CALM ($M^* < -.34$; n = 169)	CALM Preferable to Usual Care ($M^* > -.34$; n = 707)
Anxiety Sensitivity (ASI)	20.94 (11.54)	31.38 (13.43)
> High School Education	88.8%	76.9%
Female	37.9%	79.5%
Number of Non-Psychotropic Medications		
0	20.1%	45.4%
1	20.1%	22.1%
2	11.8%	15.1%
>3	47.9%	17.4%
Satisfaction with Care	4.12 (.80)	3.67 (1.06)
< High School Education	10.0%	3.4%
Social Standing	5.13 (2.01)	5.47 (2.07)
Depression (PHQ)	9.80 (5.56)	13.50 (6.22)
Number of Psychotropic Medications		
0	33.1%	37.3%
1	29.8%	30.3%
2	20.4%	18.5%
>3	16.2%	14.3%
Black Race	7.0%	12.5%