ABSTRACT

Title of Thesis: SUBTYPE OF DEPRESSION AS A MODERATOR OF RESPONSE IN THE TREATMENT OF MAJOR DEPRESSION WITH ST JOHN'S WORT

Christine Marie Ulbricht, Master of Public Health in Epidemiology, 2011

Thesis directed By: Professor Sunmin Lee
Department of Epidemiology and Biostatistics

Objective: To evaluate the effect of the melancholic and anxious subtypes of depression on treatment response in the Hypericum Depression Trial. Methods: 340 adults with depression were classified according to depression subtype at baseline. Linear and logistic regression models examined the effects of depression subtype and treatment assignment on treatment response. Results: 64.41% of participants had melancholic depression and 71.76% had anxious depression. The linear regression model demonstrated that melancholic depression status at baseline and the interaction of treatment assignment and baseline melancholic status had significant effects on depression severity. The linear regression model showed significant effects only for treatment and anxious depression status at baseline on depression severity. Conclusion: While depression subtype appears to be related to certain clinical characteristics, this study was

inconclusive and did not find melancholic or anxious depression subtypes to significantly moderate response to treatment with St John’s wort, placebo, or sertraline.

SUBTYPE OF DEPRESSION AS A MODERATOR OF RESPONSE IN THE TREATMENT OF MAJOR DEPRESSION WITH ST JOHN'S WORT

By

Christine Marie Ulbricht

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Master of Public Health
2011

Advisory Committee:
Dr. Sunmin Lee, Chair
Ms. Joanne Severe
Dr. Guangyu Zhang

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Depression is one of the most common mental disorders in the United States.[1] Despite many years of research on treatments for depression, it is still unclear as to how to successfully treat people so that they achieve remission of depression symptoms. A variety of treatments for major depressive disorder (MDD) are currently used, including standard pharmaceutical antidepressant medication and complementary and alternative medicine (CAM). Response to treatments for MDD varies widely in individuals and placebo response rates of approximately 30% are common.[2] Such placebo response rates can make it difficult to interpret study results and might increase the chance of dismissing potentially efficacious treatments.[3]

CAM use in the United States has been increasing over the last several decades and is particularly prevalent among people with MDD.[4] The botanical extract St John's wort (*Hypericum perforatum*) is popular for the treatment of depression in the U.S. and Europe despite inconsistent evidence of its efficacy, especially in treating severe MDD.[5] In the trial by the Hypericum Depression Study Group of St John's wort for MDD, neither the standard antidepressant sertraline nor St John's wort was more effective than placebo.[6] These results have raised concerns about placebo-response in studies of CAM treatment for MDD.[7] Such results, however, might be influenced by participant characteristics that could moderate treatment response. In particular, people with certain subtypes of depression seem to respond differently to treatment.[8] Given the possible consequences of not properly treating depression, it is important to elucidate what might moderate response to a popular treatment such as St John's wort.

Research Aims

Since the diagnostic criteria for depression consist of heterogeneous symptoms, there might be clinical utility in distinguishing different types of depression such as anxious depression or melancholic depression. Melancholic depression is major depression with anhedonia, significant weight loss, psychomotor retardation or agitation, early morning insomnia, and guilt.[9] Anxious depression is major depression with high levels of anxiety.[10] The main objective of this study was to examine the effect of the melancholic and anxious subtypes of major depression on treatment response in the Hypericum Depression Trial. The specific aims of this research were as follows.

Aim 1. To classify the Hypericum Depression Trial population in terms of depression subtype.

Participants were classified as having melancholic versus nonmelancholic depression and by having anxious depression versus nonanxious depression. The demographic and clinical correlates of each group were then examined. Based on previous studies of standard antidepressants in large community samples,[10] one hypothesis was that participants with melancholic depression would be more likely than those with nonmelancholic depression to be male, have greater severity of depression, and a shorter duration of the current depressive episode. It was also predicted that participants with anxious depression would be more likely than those with nonanxious depression to be female, have greater severity of depression, and have more functional impairment.

Aim 2. To compare treatment response in participants with melancholic depression versus nonmelancholic depression.

The hypothesis related to this aim was that people with melancholic depression would be more likely than those with nonmelancholic depression to respond to treatment with St John's wort or sertraline. Also, people with melancholic depression were predicted to be less likely to respond to placebo than those with nonmelancholic depression.

Aim 3. To compare treatment response in participants with anxious depression versus nonanxious depression.

The hypothesis corresponding to this aim was that people with anxious depression would be less likely than people with nonanxious depression to respond to treatment with St John's wort or sertraline. Additionally, people with anxious depression would be more likely than those with nonanxious depression to respond to placebo.

Chapter 2: Background

Overview of Major Depression

Several kinds of depression are recognized in the widely-used American Psychiatric Association's classification system of diagnostic criteria, including dysthymic disorder and major depression.[9] The exact cause of MDD is currently unknown but genetic, chemical, environmental, psychological, and social factors appear to intersect in the etiology of the disease. The American Psychiatric Association, in the most recent edition of their *Diagnostic and Statistical Manual* (*DSM-IV*), defines major depression as the presence of at least five of the following symptoms nearly every day during the same 2-week period.[9]

- depressed mood
- markedly diminished interest or pleasure in all, or almost all, activities
- significant weight loss or weight gain or decrease or increase in appetite
- insomnia or hypersomnia
- psychomotor agitation or retardation
- fatigue or loss of energy
- feelings of worthlessness or excessive or inappropriate guilt
- diminished ability to think or concentrate, or indecisiveness
- recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide

Combinations of these symptoms are considered to occur in episodes. An episode could occur only once but episodes usually recur throughout a person's life.

Depression is a major public health problem due to its prevalence and its association with significant disability, morbidity, and mortality. Depression is a leading cause of disease burden throughout the world.[11] The symptoms of depression have been shown to impair ability to function in work, household, relationship, and social roles in more than 50% of people with major depression.[12] Major depression is the leading cause

of the burden of disease, as measured by disability-adjusted life years, in middle- and high-income countries.[13] Depression is the fourth largest cause of disability worldwide and the largest source of disability for Americans between the ages of 15 and 44 years.[14] Lifetime prevalence for major depression has been estimated to be 15 percent-17 precent.[15] The annual 12-month prevalence of depression in the U.S. is estimated at 6.7 percent.[1]

In 2000, the cost of depression in the United States was estimated to be $83 billion. This included $26 billion in treatment costs and $57 billion in losses produced by absenteeism, reduced productivity, and lifetime earnings lost due to suicide-related deaths.[16] People with depression have reported an average of 35 days in the past year when they were unable to work or carry out their normal activities due to their depression.[17] In comparison, people with asthma, heart disease, or diabetes reported an average of 10.6 days, 8.8 days, and 6.4 days, respectively, when they were unable to perform their regular duties. Depression is also associated with increased mortality compared to the general population. Suicide-related deaths account for part of this increased mortality but it appears that premature mortality might also be related to chronic comorbid medical illnesses and social factors.[18-20]

Given the burden of major depression for the individual and for society in general, effective treatments are important. Antidepressant medication and psychotherapy are the most common methods of treatment. Despite the importance and availability of treatment, only 51.7% of people with depression receive treatment.[21] Of those who receive treatment, only 38.0% receive minimally adequate treatment, as defined by at least two months of appropriate medication plus at least four visits to a physician or eight

sessions of psychotherapy lasting an average of 30 minutes.[21] Of those people who do receive treatment for major depression, pharmacotherapy is effective for only approximately 70%.[22]

Even when people with major depression do receive treatment, response to treatment can vary widely. Poor response to one treatment does not necessarily mean that a different treatment will not be effective.[23] Unfortunately, many people do not try alternate treatments after the first one fails. Only about half of people make follow-up visits after starting antidepressant treatment.[24] Because there is usually such a narrow window of opportunity to effectively treat major depression, being able to provide personalized medicine by knowing the best first treatment for a particular person would be highly beneficial.[23] Characteristics of individuals with depression that reliably predict response to treatment must first be identified in order to develop personalized medicine for depression. Numerous characteristics that might moderate treatment response to standard antidepressant medication have been studied, including sociodemographic, clinical, and biological characteristics. Most studies of possible predictors of response have focused on the class of standard antidepressants known as selective serotonin reuptake inhibitors (SSRI).[25] One possible moderator of treatment response that has been considered is subtype of depression. While subtypes of depression have been studied as moderators of treatment with standard antidepressant medication.[8, 26, 27] there has been less research on how depression subtypes might affect response to CAM treatment.

Treatment for Depression

Standard Antidepressants

Major depression is usually treated with standard antidepressant medications but the effect sizes of these antidepressants in bringing about remission of depression symptoms have been modest in clinical trials.[28, 29] Despite the modest effectiveness, standard antidepressants were the fourth-largest pharmaceutical class in terms of sales in 2009, with sales of $9.9 billion.[30] SSRIs or the older tricyclic antidepressants are generally the first antidepressants to be tried when treating depression. Standard antidepressants are thought to work by affecting the metabolism and receptors of neurotransmitters in the brain.[31] Antidepressants are generally classified by the receptors believed to be involved in the antidepressant's mechanism of action. Some of the first pharmaceutical antidepressants used to treat major depression were monoamine oxidase inhibitors (MAOI) and tricyclic antidepressants (TCA).[32] MAOIs and TCAs eventually dropped out of favor because of their overwhelming side effects and possibly fatal contraindications/reactions. These older classes of antidepressants have more recently been replaced by selective serotonin reuptake inhibitors (SSRI) such as fluoxetine, paroxetine, and sertraline. SSRIs appear to be only partially effective at completely ameliorating major depression. The SSRIs now have competition from the serotonin norepinephrine reuptake inhibitors (SNRI) such as venlafaxine and duloxetine. Both newer and the older classes of standard antidepressants fall short in terms of inadequate efficacy in achieving complete symptom remission and undesirable side effects.[33] The most commonly prescribed antidepressants that received FDA-approval between 1987 and 1999 were the SSRIs fluoxetine, paroxetine, and citalopram, the SNRI venlafaxine,

and the serotonin norepinephrine dopamine reuptake inhibitor nefazodone.[34] It seems that SSRIs and SNRIs are both only moderately effective for people with the most severe depression.[28] In a pooled analysis of trials treating major depression with the SNRI venlafaxine, SSRIs or placebo, only 35% of people taking SSRIs experienced complete symptom remission.[35] Remission rates for placebo were 25%.

St John's Wort

The National Center for Complementary and Alternative Medicine defines complementary and alternative medicine as "a group of diverse medical and health care systems, practices, and products that are not presently considered part of conventional medicine."[36] Among adults who participated in the National Comorbidity Survey Replication and the National Survey on American Life: Coping with Stress in the 21st Century, 40% of those with a mood disorder reported using CAM in the previous 12 months.[37] Of those with an anxiety disorder, 35% used at least one form of CAM. Forty one percent of those with comorbid mood and anxiety disorders used CAM. Of study participants who met criteria for any *DSM-IV* disorder in the last 12 months and used CAM, 26% reported using herbal therapy.

One popular form of herbal therapy is St John's wort (*Hypericum perforatum*) supplements.[38] In Germany and other countries in Europe, St John's wort has been used for centuries for treating mild to moderate depression.[39] The regulation of St John's wort products varies by country and thus the formulations available on the market can differ considerably.[29] The composition of active ingredients in St John's wort supplements can vary considerably although hyperforin and total hypercerin are thought to be the most main components.[40] Despite its popularity, the exact nature of the efficacious ingredients

and the pharmacological mechanism of the antidepressant effects are relatively unknown.[39] Differential responses to St John's wort have been reported for different subtypes of depression.[41] Side effects of St John's wort are thought to be mild although there is concern over the possibility of herbal-drug interactions in people taking medication and supplements in addition to St John's wort.[42]

There is much debate over the efficacy and effectiveness of St John's wort in people with major depression. There is more support for the use of St John's wort in mild to moderate major depression but less support for using it for severe major depression. This lack of consensus might be due to heterogeneity in clinical trials methodology and inconsistency in which outcome measures are used.[29] One multi-site trial of St John's wort versus placebo for moderate to severe major depression found that St John's wort was not more effective than placebo.[43] In a trial of treatment with St John's wort, fluoxetine, or placebo for people with mild to moderate depression, St John's wort was significantly more effective than fluoxetine but not more effective than placebo.[39] In one review of studies of St John's wort compared with placebo or standard antidepressants for the treatment of depression, it was found that St John's wort was superior to placebo and similarly effective as standard antidepressants.[29] This same review also concluded that fewer side effects were associated with St John's wort than with standard antidepressants and that results of studies in Germany were generally more favorable than results from studies in other countries.

Measuring Treatment Response

In studies of treatment for major depression, treatment response is usually defined as a certain amount of change in depression severity between baseline and the end of

treatment. Severity of depression has been most often measured with the Hamilton Depression Rating Scale (HAM-D).[44] According to Hamilton, this observer-rated scale is a way to assess the initial severity of depression and any changes in the severity. The scale should not be used to diagnose depression.[45] The HAM-D was originally published in 1960 with 17 core items and was intended to systematically quantify results of clinical interviews with inpatients that had already been diagnosed with a depressive disorder.[46, 47] The 17 items assess depressed mood, guilt, suicide, initial insomnia, middle insomnia, delayed insomnia, work and interests, psychomotor retardation, psychic anxiety, somatic anxiety, genital symptoms, hypochondriasis, loss of insight, and weight loss.[46] Hamilton thought that four additional items of diurnal variation, depersonalization, paranoid symptoms, and obsessive symptoms could be included on ratings forms but should not be included in the final total score because these items either did not measure the intensity of depression or occurred too rarely to warrant inclusion.[46] Although 21- and 24-item versions of the HAM-D now exist, the 17-item version of the scale that was used in the Hypericum trial is the one most commonly used.[44] $HAM\text{-}D_{17}$ scores can range from 0-54. Depression severity increases as the score increases. A score of 0-7 usually indicates normal or remission of depression symptoms. A score of 20 or greater indicates moderate to severe depression.

The HAM-D is to be completed by clinicians using information obtained through a clinical interview. The items on the HAM-D are anchor-point descriptions of symptoms that are thought to be common in depression. Hamilton did not include a standardized interview for clinicians to use with the scale so the reliability of the scale is dependent on the clinician using it. Several structured interview guides have been developed in order to

improve the reliability of the scale.[44] The Williams' Structured Interview Guide for the Hamilton Depression Rating Scale has been found to improve the item reliability and facilitate rater training when compared to HAM-D ratings that are produced from unstructured interviews.[48, 49]

There has been much criticism of the HAM-D in the past fifty years, including that the scale is measuring a definition of depression that is only partially related to the common *DSM-IV* criteria for depression.[50] The scale does not recognize symptoms of depression such as helplessness, hopelessness, and worthlessness that are now considered to be relevant to defining depression.[51] Another criticism of the HAM-D is that its sensitivity to detect changes in depression severity is reduced when used with people who have higher levels of depression severity.[51] It has also been suggested that the HAM-D total score has an inherent bias for tricyclic antidepressants since the scale includes three items about sleep and one item about weight gain.[52] These items address common side effects of tricyclic antidepressants such as changes in appetite and weight and drowsiness. It appears that the HAM-D might also be affected by the side effects of SSRIs such as gastrointestinal symptoms, sleep disturbances, nervousness and agitation. These SSRI side effects might be confused with the depression symptoms in the HAM-D and thus people taking SSRIs might appear to have higher HAM-D scores and thus more severe depression.[52]

Depression severity can also be measured with the Clinical Global Impressions (CGI) scales. The CGI are two scales used in central nervous system trials to assess efficacy of treatment, usually as secondary outcome measures. The two scales are the CGI-Severity (CGI-S) and the CGI-Improvement (CGI-I). The CGI scales yield measures

of severity of illness (CGI-S), global improvement (CGI-I), and can include efficacy index of the interaction of therapeutic effectiveness and adverse reactions.[53] These scales were originally designed to provide a way for a study clinician to globally assess a study participant's condition before and then after receiving the study treatment.[54] The CGI scales were first published by the National Institute of Mental Health as part of a manual of assessments to be used in studies of psychotropic drugs. With the CGI-S, a clinician evaluates a participant at baseline to answer the question "considering your total clinical experience with this particular population, how mentally ill is the patient at this time?" The clinician rates the severity of the participant's illness on a 7-point scale where 1 is "normal, not at all mentally ill" and 7 equals "among the most extremely ill patients." Following the study treatment, the clinician uses the CGI-S again or uses the CGI-I to measure the change from pre- and post-treatment. The CGI-I is often used in clinical trials of antidepressant drugs to characterize treatment response.[3] For the 7-point CGI-I, the clinician determines if the participant's disorder has improved, become worse, or stayed the same. On the CGI-I, 1 is "very much improved" and 7 is "very much worse."

The CGI scales have been adapted routinely used in clinical trials of psychotropic treatments for bipolar disorder, anxiety, and schizophrenia but not in depression.[55] Although the CGI is often used in depression trials, there is still debate over the scale's specificity.[56] The validity of the CGI might be less than ideal because it relies on the clinician to be able to compare the severity of one person's illness to another's based on subjective experience, there is no standard interviewer guide to use with the scale, and because the format of the responses is ambiguous.[55]

Depression Subtypes

Given the heterogeneity of the diagnostic criteria for major depression, depression is likely a clinical syndrome or group of disorders rather than one distinct disease. People with depression can vary widely in terms of symptoms, comorbidities, clinical course, pathophysiology, severity, and treatment responsiveness.[57, 58] In order to further refine the diagnostic criteria for major depression and predict response to antidepressant treatment, different subtypes of depression have been proposed throughout the history of psychiatry. The proposed subtypes are based on differences in symptoms, onset of depression, trajectory, and the severity of symptoms.[57, 59-62] Subtypes have been found to be associated with different risk factors and characteristics.[60] In order for a subtype to be clinically useful, the subtype should be able to predict treatment response and thus have implications for the selection of treatment.[57]

Numerous distinct subtypes of major depression have been proposed, including psychotic depression, atypical depression, melancholic/endogenous depression, and anxious depression. While the distribution of such subtypes in the general population is unknown,[63] it is believed that these subtypes are neither mutually exclusive nor exhaustive. Anxiety symptoms are very commonly comorbid with depression in general[64] and it has been observed that anxiety disorders tend to be comorbid with atypical depression.[65] Additionally, in one study, 14% of participants met criteria for both melancholic depression and atypical depression.[66] In a post-hoc analysis of the large Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study, almost one quarter of participants met criteria based on the *DSM-IV* for melancholic depression. Seventy one percent of these participants with melancholic depression also had features

of anxious depression.[67] In the same STAR*D post-hoc analysis, 47.7% of participants with nonmelancholic depression had anxious depression. The different subtypes of depression also might not be stable and thus might vary by episode of depression. It is possible that a person might experience symptoms of melancholic depression during one episode but have anxious depression in another episode.[67]

Melancholic Depression

While the idea of melancholic depression is centuries-old,[68] there is no widely accepted clinically useful definition. Melancholic depression is also sometimes referred to as "endogenous" depression, based on the idea that it is a type of depression that "grows from within" an individual.[27] The Newcastle Scale, the Research Diagnostic Criteria,[69] and several editions of the *DSM* have all defined a version of melancholic depression. Although the concept of melancholic depression has been debated for many years, it was first formally included in the third edition of the *DSM* in 1980.[67] "Depression with melancholic features" is now one of two subtypes of major depression currently included in the *DSM-IV*.[9] The *DSM-IV* criteria for melancholic features emphasize having either the loss of pleasure in activities or lack of reactivity to usually pleasurable stimuli. It also includes the presence of at least three of the following: distinct depressed mood, worse symptoms in the morning, waking up at least two hours earlier than usual, notable psychomotor agitation or retardation, significant weight loss, and excessive guilt. The *DSM-IV* criteria have been criticized as being imprecise, inadequate for use in making decisions about treatment, and for contributing to diagnostically heterogeneous samples in clinical trials.[68] There is considerable controversy over how melancholic depression will appear in the upcoming *DSM-5*. This controversy centers on

whether melancholic depression should continue to be thought of as a category of depression features or if it should be recognized as a distinct syndrome.[70]

Despite the lack of agreement over how to define melancholic depression, unreactive mood, guilt, changes in weight, lack of a precipitating event, and psychomotor disturbance have continually been the main diagnostic components.[71, 72] A review of the available literature in 2005 determined that there is evidence that melancholic depression is distinct from nonmelancholic depression in terms of biological functioning, personality, treatment response and suicidality.[73] The presence of psychomotor disturbance has been proposed as a marker of an underlying neuropathological process specific to melancholic depression.[74] The distinct biology of melancholic depression appears to include having the long allele of the serotonin promoter polymorphism, loss of hippocampal volume, and intracellular signal transduction abnormalities.[75-77] The melancholic subtype also seems to be at least partially determined by genetics since the subtype independently aggregates in families.[59] Adding to the body of evidence pointing to melancholic depression having a distinct biological component, sleep patterns have been observed to differ for people with melancholic depression as compared to people with nonmelancholic depression.[78]

Numerous demographic and clinical features are also associated with melancholic depression. Studies applying applied different versions of diagnostic criteria have found that people with melancholic depression, as compared to those with nonmelancholic depression, were more likely: to be older, to have more severe depression, to have better response to somatic treatment, and to have poorer response to psychotherapy.[27] Additionally, people with depression who are hospital inpatients are more likely to have

melancholic depression than those who are outpatients.[79] In STAR*D, melancholic depression was associated with greater severity of depression as measured by the HAM-D and the Quick Inventory of Depressive Symptomatology than was nonmelancholic depression.[27] Having melancholic depression during the STAR*D trial was also related to a lower likelihood of symptom remission with standard antidepressants than was having nonmelancholic depression. Study investigators have hypothesized that this treatment resistance might be related to the presence of comorbid anxious depression.[67]

It is unclear if the presence of the melancholic subtype of depression is useful for the selection of medication for treating outpatients although it does appear to be useful with inpatients.[67, 79] Evidence on how people with melancholic depression respond to different kinds of antidepressants is varied. It appears that the older classes of TCAs and MAOIs are more effective than placebo for treating melancholic depression.[80-83] People with melancholic depression did not respond well to the SSRI citalopram in the first phase of STAR*D In the later study phases of STAR*D, however, the presence of melancholic depression did not predict a differential response to sertraline as compared to the sustained-release bupropion or sustained-release venlafaxine.[84] Inpatient samples have demonstrated that SSRIs are not effective for people with melancholic depression.[85] It remains uncertain if this result indicates lower efficacy of all SSRIs for people with melancholic depression.

Anxious Depression

It has long been observed that depression and anxiety often co-occur.[86] There is no accepted definition of anxious depression but it is often defined with a dimensional approach as major depression with high levels of anxiety symptoms.[10, 87] It can also be

defined with a syndromal approach as major depression with a comorbid anxiety disorder.[10] The dimensional approach seems more appropriate for clinical use because many people with major depression also have definite symptoms of anxiety but these symptoms might not be clearly discernible as distinct from the depression or might not fully qualify as an anxiety disorder under the *DSM-IV* or *ICD-10* diagnostic criteria.[88, 89] Anxious depression is not included as a diagnosis separate from depression in either the *DSM-IV* or *ICD-10* but there is evidence that it is a unique depression subtype.[90] When defined using the dimensional approach, anxious depression seems to be a fairly common subtype of depression.[91] Of people diagnosed with depression, 50-70% experience moderate levels of comorbid anxiety and 20-25% have severe levels of anxiety.[92]

A score of 7 or greater on the Guy and Cleary anxiety/somatization factor of the HAM-D is often used in studies to determine the presence of anxious depression as defined as major depression with high levels of anxiety symptoms.[10, 53, 93, 94] The anxiety/somatization factor includes items on psychic anxiety, somatic anxiety, gastrointestinal somatic symptoms, general somatic symptoms, hypochondriasis, and insight. According to the definition using the anxiety/somatization factor, 44-46% of participants in the STAR*D trial had anxious depression.[10, 90] Anxious depression, when compared to nonanxious depression, is associated with greater severity of symptoms, greater functional impairment, greater chronicity of depression, and longer episodes of depression.[10, 58, 90, 95, 96] It is also related to having more episodes of depression than nonanxious depression.[58] People with anxious depression also appear to have an increase risk for committing suicide.[22] People with anxious depression are more likely than those with nonanxious depression to report suicidal ideation and previous suicide attempts.[58]

In terms of treatment response, people with anxious depression tend to have poorer responses to antidepressant treatment than do people with nonanxious depression.[91, 94] Results of previous studies are mixed but anxious depression does seem to be related to differential responses to different classes of standard antidepressants. In one study, greater severity of psychic and somatic anxiety symptoms of major depression at baseline predicted an increased likelihood of non-response to fluoxetine, regardless of depression severity at baseline.[97] It appears that SSRIs might have a slight advantage over bupropion, a norepinephrine and dopamine reuptake inhibitor.[91] In a pooled analysis of data from randomized controlled trials of the SSRI escitalopram, there was no difference in response to escitalopram, older SSRIs, or SNRIs.[98] In contrast, the SNRI venlafaxine has appeared to be superior to the SSRI fluoxetine and to placebo in treating anxious depression.[99] In another study, mirtazapine, a noradrenergic and specific serotonergic antidepressant, was more effective in reducing the symptoms of major depression with anxiety than paroxetine was during the first few weeks of treatment.[100] This trial also indicated that reducing the anxiety symptoms quickly may increase treatment adherence.

Interpreting results of treatment trials for anxious depression can be difficult since people with anxious depression also tend to experience a delayed treatment response.[101] Because of the delayed treatment response associated with comorbid anxiety and depression, people with anxious depression may need to be evaluated for a longer period of time than those with nonanxious depression. It has been suggested that people with anxious depression be evaluated for 9-12 weeks before the efficacy of a specific treatment is established.[102] Standard antidepressants are generally thought to start producing measurable change in depression severity in 2-4 weeks following initiation of

treatment with some data suggesting onset of action might begin as soon as 3 days after treatment begins.[103]

Chapter 3: Research Design and Methods

This thesis research is a secondary data analysis of longitudinal data from a clinical trial of St John's wort for the treatment for major depression (ClinicalTrials.gov identifier NCT00005013). The publicly available limited access dataset used for this thesis was obtained from the National Institute of Mental Health (NIMH) through the institute's standard data use certification process.[104]

The Treatment of Major Depression with St John's Wort Trial

The Treatment of Major Depression with St John's Wort trial, also known as the Hypericum trial, was a study sponsored by NIMH and the National Center for Complementary and Alternative Medicine to determine the acute antidepressant efficacy of a standardized extract of St John's wort (LI-160) for the treatment of major depression.[105] It was a double-blind, randomized, placebo-controlled trial in which participants received St John's wort, the standard SSRI sertraline, or placebo. There was a one-week placebo run-in phase followed by randomization to St John's wort, sertraline, or placebo. This acute treatment phase lasted for 8 weeks. Assessments were completed weekly or biweekly until week 8. Participants who had a partial or full response by week 8 of the acute phase could enter a 4-month continuation phase. The main aim of the original Hypericum trial was to test if St John's wort was superior to placebo after treatment for 8 weeks. Sertraline was included as an active comparator to calibrate the validity of the trial. The main result was that neither St John's wort nor sertraline differed significantly from placebo on the change in HAM-D total score from baseline to week 8 and the incidence of full treatment response at week 8.[6]

Participants

The Hypericum trial consisted of 340 adult outpatient participants who were recruited from 12 academic and community clinics between December 1998 and June 2000.[6] These participants had moderate to severe major depression as diagnosed with the modified Structured Clinical Interview for Axis I *DSM-IV* disorders.[106] The inclusion criteria for the original Hypericum trial were: 1) at least 18 years of age, 2) moderate to severe depression as determined by a minimum total score of 20 on the 17-item HAM-D at screening and at baseline, 3) maximum score of 60 on the Global Assessment of Functioning at screening and at baseline, 4) HAM-D score cannot decrease by 25% or more between screening and baseline, 5) capacity to give informed consent and to follow study procedures, 6) identification of a close personal contact who would be notified of any clinical concerns, and 7) abstinence or effective contraception used throughout the duration of the study.

Details of enrollment and outcomes during the Hypericum trial are detailed in Figure 1. Four hundred twenty eight people entered the one-week placebo run-in period of the Hypericum trial. Three hundred forty people were randomized after the one-week run-in period. One hundred eleven participants were assigned to sertraline, 113 participants were assigned to St John's wort, and 116 participants were assigned to placebo. In the main outcome paper of the original trial, it was reported that two participants in the sertraline group had HAM-D total scores below the enrollment requirement of a HAM-D of at least 20. In the original Hypericum trial outcomes analysis, these two people have been included in the baseline analysis but excluded from the efficacy analysis.

This thesis research used limited access data from the participants. The original protocol and consent forms were approved by the institutional review boards at each clinical site, Duke University Medical Center, and Research Triangle Institute International.[107] Participants provided informed consent prior to entering the trial. During the consent process, participants were informed that they had 1 in 3 chances of receiving St John's wort, sertraline, or placebo. The Hypericum trial was monitored by the NIMH Data and Safety Monitoring Board. All personal identifiers in the dataset used for this thesis were removed and other data elements modified by NIMH prior to releasing the dataset. This was done to reduce to likelihood that any individual participant can be identified from the data. In order to access the data, the appropriate NIMH Data Use Certification and approval from the University of Maryland, College Park Institutional Review Board was obtained.

Efficacy Measures

Treatment Response

The main outcome variable was response to treatment for depression. The primary measure was the incidence of response at week 8 or at early study termination. Response was defined as a HAM-D total score of 12 or less, a 50% reduction in HAM-D score, and a CGI-I score of 1 or 2.

Severity of Depression

The CGI-S and CGI-I were ascertained as secondary outcome measures in the original Hypericum trial and were also reported separately from treatment response in this paper. For the HAM-D and the CGI-S, a higher score indicates more severe

depression. A higher score on the CGI-I indicates a lack of improvement or worsening of symptoms.

Functioning

Functioning was defined as the general ability to carry out daily activities. Functioning was ascertained using scores on the Global Assessment of Functioning Scale (GAF) and the Sheehan Disability Scale (SDS). The clinician-rated GAF is a numeric scale of 0-100 used by clinicians to assess social, occupational, and psychological functioning. On the GAF, 0 indicates severe impairment and a persistent danger of severely hurting oneself or others while 100 indicates superior functioning in a variety of activities. The brief, self-report SDS assesses impairment in work, social life, and family life. The SDS ranges from 0-30. On the SDS, a score of 0 signifies unimpaired while a score of 30 is highly impaired.

Independent Variables

Subtype of Depression

The original Hypericum trial did not classify participants according to depression subtype so classifying needed to be done post-hoc for the purposes of this study. Because a standard accepted way of identifying melancholic depression using scores on the HAM-D_{17} does not presently exist, a modified version of the Hamilton Endogenomorphy Subscale (HES) for identifying melancholic depression was used in this analysis.[108, 109] The HES was modified for use in this thesis research because not all the HES items are asked in the 17-item version of the HAM-D that was used in the Hypericum trial. The version of the HES used here does not include items on diurnal variation and hopelessness. These two items are only available on the 24-item version of the HAM-D.

Additionally, the definition used here includes an item on guilt. Although guilt was not included in the original HES, it is included here because it is one of the factors in the *DSM-IV* criteria for major depression with melancholic features.[9] The modified HES used in this study incorporated the items from the HAM-D_{17} that address depressed mood, guilt, late insomnia, retardation, agitation, and weight loss. Scores can range from 0-20. The presence of melancholic depression on the modified HES is indicated by a score of 9 or greater. Nonmelancholic depression is considered to be having a modified HES score of 8 or less.

Anxious depression was defined as major depression with high levels of anxiety symptoms as measured by a score of greater than or equal to 7 on the HAM-D anxiety/somatization factor derived by Cleary and Guy.[110] This measure has been used in other studies to classify anxious depression.[10] This factor includes the psychic anxiety, somatic anxiety, gastrointestinal somatic symptoms, general somatic symptoms, hypochondriasis, and insight factors on the HAM-D_{17}. Scores for the HAM-D anxiety/somatization can range from 0 to 18.

Treatment Assignment

Participants were assigned at the beginning of the trial to receive St John's wort, sertraline, or placebo. The St John's wort extract was provided by Lichtwer Pharma and was standardized to be between 0.12% and 0.28% hypericin. Hypericin is one of the active components of St John's wort. Sertraline was provided by Pfizer, Inc. Following a week-long run-in period of placebo tablets, participants were given either 900 mg/d of St John's wort, 50 mg/d of sertraline, or placebo. These treatments were given 3 times per

day. Daily doses of all treatments could be increased after week 3 or 4, depending on the severity of the participant's depression during the study.

Covariate

The HAM-D total score at baseline, before treatment, was the main measure of the severity of depression symptoms at the beginning of the study. This baseline score was used in the ANCOVA models as a covariate.

Statistical Analysis

The first task for analyzing the study data was to classify participants as having melancholic or nonmelancholic depression and having anxious or nonanxious depression. This was done according to the previously outlined melancholic and anxious HAM-D factor scores. In order to establish a cutoff point on the modified HES to define melancholic depression, the distribution of total scores on the modified HES was evaluated as was originally done by Thase et al.[109] The distributions of total scores on the HAM-D anxiety/somatization factor were also examined to confirm whether a cutoff score of 7 should be used to determine anxious depression. Comparisons were not made between participants with anxious depression and those with melancholic depression because of the possibility that these two subgroups are not mutually exclusive.

After participants were categorized having melancholic or nonmelancholic depression and as having anxious or nonanxious depression, baseline demographic and clinical features were compared for any significant differences among participants with melancholic versus nonmelancholic depression and with anxious versus nonanxious depression in order to determine if there were differences in the distributions of subtypes across the treatment groups at baseline. Demographic characteristics were age, sex, and

race. Clinical characteristics included length of current episode of depression, severity of depression at baseline, and level of functioning. Intent-to-treat analyses were performed. All 340 participants in the original Hypericum trial were included in these analyses, including analyses of baseline measures and analyses of results at week 8. This is in contrast to the original Hypericum trial's primary outcomes analyses which included all 340 participants in the baseline analyses but excluded 2 participants from the efficacy analyses because they did not meet the inclusion criteria of having a HAM-D total score of at least 20.[6]

Descriptive statistics were calculated in order to characterize the baseline demographic and clinical characteristics of participants by subtype of depression. These statistics include frequencies and percentages for categorical variables and mean and standard deviations for continuous variables. T tests were used to determine if there were differences within each depression subtype for the continuous variables age and length of current episode of depression. Similarly, chi-square and Fisher exact tests were used for comparing differences for the categorical variables sex and race in each subtype of depression. Wilcoxon-Mann-Whitney U tests were used for the ordinal variables of depression severity at baseline and level of functioning.

For each subtype of depression, an analysis of covariance (ANCOVA) model evaluated the severity of depression at the end of treatment as measured by HAM-D total score at week 8, with treatment and depression subtype level as main effects and baseline severity of depression as a covariate. Included in the model was the interaction term for treatment and subtype level in order to evaluate whether there is a differential response to treatment dependent on subtype level, e.g., in anxious depression versus nonanxious

depression. ANCOVA models were also used to obtain adjusted means of the other clinical characteristics.

The percentage of study participants remaining in each treatment group at week 8 was also compared within each subtype of depression using chi-square tests. Clinical characteristics and status of treatment response at week 8 by depression subtype were also examined for differences within subtypes of depression by treatment group.

The last observation available was substituted for missing observations at week 8. With the last observation carried forward (LOCF) method, missing values are replaced with the most recently obtained value for the same participant. LOCF was used here in order to match the methods used in the primary outcome paper of the original Hypericum trial as closely as possible.[6] LOCF is also used because it is unknown why missing observations occurred and thus it cannot be said with certainty that the missing observations were missing at random, completely at random, or were ignorable. Unlike in the original Hypericum trial, the LOCF imputation performed in this thesis was used for observations only as far back as week 2. This was done because it did not seem like a measurable response to treatment could have been detected unless the medication had been taken for at least two weeks. Because of this, some missing values of study measurements still remained in the dataset.

Because of concern that the number of data substitutions would be high and thus the LOCF data might not accurately represent the true treatment outcome at week 8, mixed effects linear regression models were also fit for each depression subtype in order to utilize all available data points and thus possibly more accurately reflect the actual response. Linear, rather than quadratic, models were used because the means of the

outcomes variables appeared to decrease linearly with time. The models explored the effects of treatment, depression subtype level, and time on HAM-D total score. The time points included in these models were baseline and weeks 1-8. Comparisons on goodness of fit statistics AIC and BIC were made between compound symmetry, unstructured, compound heterogeneous symmetry, autoregressive, and heterogeneous autoregressive variance-covariance structures.

The probability of achieving at least a moderate level of treatment response, versus no response, was also modeled for each depression subtype. Treatment response was examined using multivariate logistic regression. In each depression subtype, the logistic regression model examined the effect of predictor variables that included treatment assignment and depression subtype status at baseline. Of the treatment assignments, St John's wort was compared to placebo and sertraline was compared to placebo. St John's wort and sertraline were not directly compared to each other because the intent of the original Hypericum trial was to compare the efficacy of St John's wort to placebo. Sertraline was originally used as an active comparator.

The significance level for all analyses was set at a p-value of 0.05. All statistical analyses were performed using SAS 9.2 software (SAS Institute Inc., Cary, NC).

Chapter 4: Results

In order to classify people as having melancholic or nonmelancholic depression, a cutoff point on the modified HES needed to be determined. Thase et al. validated the original HES in a group of 147 women outpatients with depression who were not receiving treatment and used HES ≥ 8 as the determination of having high levels of endogenous/melancholic depression.[109] This differs from the work of Kovacs et al. that used HES ≥ 9 to indicate that people had high levels of endogenous/melancholic depression.[108] Because a clinically validated and accepted cutoff score for the modified HES does not exist, the distribution of scores on the modified HES was examined. Figure 2 displays the cumulative percentage of participants for each HES score. The frequency of participants for each HES score is displayed in Figure 3. The cumulative percent for HES=9 was 56.76 and the frequency was n=193. Thus, for the purposes of this thesis, melancholic depression was defined as HES ≥ 9. Nonmelancholic depression was then defined as HES ≤8.

Participants were categorized as having anxious depression or nonanxious depression according to scores on the HAM-D anxiety/somatization factor. As described in Chapter 3, a cutoff score of 7 or greater to indicate the presence of anxious depression was originally chosen based on recommendations in the literature. An examination of the distribution of anxiety/somatization factor scores in this Hypericum sample revealed that the cumulative percent for a score of 7 was 52.65 (Figure 4). The cumulative frequency at this score was 179 (Figure 5). This distribution confirmed that a score of 7 or greater should be used to delineate anxious depression from nonanxious depression.

The frequency of the melancholic and anxious depression subtypes in the Hypericum trial sample at baseline is displayed in Table 1. Slightly more than 64% of participants had at least melancholic depression. A little more than 71% of participants had at least anxious depression. Almost 43% of the sample had both melancholic and anxious depression. Approximately 29% of participants had both nonmelancholic and anxious depression. About 22% of participants had both melancholic and nonanxious depression. Less than 7% of participants had both nonmelancholic and nonanxious depression.

Baseline demographic and clinical characteristics of participants with melancholic and nonmelancholic depression are detailed in Table 2. Almost two-thirds of participants were classified as having melancholic depression (n=219) while 121 participants had nonmelancholic depression. Significant differences between the melancholic and nonmelancholic groups were observed only for age and HAM-D, CGI-S, GAF, and SDS total scores. People with melancholic depression were older on average than those with nonmelancholic depression. The mean age was 44.03 years for those with melancholic depression and 40.54 years for those with nonmelancholic depression ($t(338) = -2.35$, $p = 0.02$). The range of HAM-D total scores for people with melancholic depression was 18 to 33. The range of HAM-D total scores for people with nonmelancholic depression was 20-27. The average HAM-D total score was 23.51 for participants with melancholic depression and 21.17 for participants with nonmelancholic depression. People with melancholic depression had significantly higher ratings of depression severity, as measured by HAM-D total score, than those with nonmelancholic depression

($|z|_{1-\alpha/2}$= 7.11, $p < 0.0001$). People with melancholic depression also had higher levels of depression severity according to the CGI-S than did people with nonmelancholic depression ($|z|_{1-\alpha/2}$= 4.39, $p < 0.0001$). In terms of levels of functioning, the melancholic depression group had significantly higher levels of functional impairment according to scores on the SDS but had lower impairment according to scores on the GAF than did the nonmelancholic depression group ($|z|_{1-\alpha/2}$= 2.40, $p = 0.02$ and $|z|_{1-\alpha/2}$= 2.40, $p = 0.02$, respectively).

Baseline characteristics of participants with anxious or nonanxious depression are detailed in Table 3. Almost 72 percent of participants were classified as having anxious depression (n=244) versus nonanxious depression (n=96) at baseline. For the anxious subtype, there were no significant differences among the demographic characteristics. Of the clinical characteristics, differences between the anxious and nonanxious groups were significant for HAM-D, CGI-S, and SDS total scores but not for GAF score. People with anxious depression had significantly higher levels of depression severity than did those with nonanxious depression according to HAM-D score ($|z|_{1-\alpha/2}$= 7.39, p <0.0001) and CGI-S ($|z|_{1-\alpha/2}$ = 2.08, $p = 0.04$). People with anxious depression also had significantly more functional impairment as measured by SDS total score than did people with nonanxious depression ($|z|_{1-\alpha/2}$= 2.06, $p = 0.04$).

Differences in baseline clinical and demographic characteristics of the melancholic subtype were also compared by treatment assignment within each level of the depression subtype (Table 4). Differences by treatment group for sex, race, and length of current depression episode were not significant. For age, the only significant difference

was for those assigned to receive sertraline. In the sertraline group, people with melancholic depression were older than those with nonmelancholic depression ($t(109) = -2.04$, $p = 0.04$). People with melancholic depression who were assigned to sertraline had the highest mean age of all treatment groups. In all treatment assignments, people with melancholic depression had higher levels of depression severity according to HAM-D and CGI-S scores than did the people with nonmelancholic depression. The differences in HAM-D scores were significant for the people assigned to St John's wort ($|z|_{1-\alpha/2} = 5.64$, $p < 0.0001$), to placebo ($|z|_{1-\alpha/2} = 3.61$, $p = 0.0003$), and to sertraline ($|z|_{1-\alpha/2} = 2.94$, $p = 0.003$). The differences in CGI-S scores between melancholic and nonmelancholic depression groups were significant for the people receiving St John's wort ($|z|_{1-\alpha/2} = 2.65$, $p = 0.008$), for those receiving placebo ($|z|_{1-\alpha/2} = 2.65$, p = 0.008), and those receiving sertraline ($|z|_{1-\alpha/2} = 2.25$, $p = 0.02$). People with melancholic depression assigned to receive St John's wort had the highest mean severity of depression. Of those assigned to receive St John's wort, people with melancholic depression had higher levels of functioning as measured by GAF scores compared to people with nonmelancholic depression ($|z|_{1-\alpha/2}= 2.36$, $p = 0.02$). Differences in GAF scores were not significant for people assigned to receive placebo or sertraline. When functioning was measured by SDS total score, people with melancholic depression assigned to the placebo group were significantly more impaired than those in the placebo group with nonmelancholic depression ($|z|_{1-\alpha/2} = 2.33$, $p = 0.02$). Differences in SDS total scores were not significant for people receiving St John's wort or sertraline.

The baseline clinical and demographic characteristics of participants by treatment group for the two levels of the anxious depression subtype are available in Table 5. There

were no significant differences for the variables of sex, race, and length of current depression episode by treatment group. In the St John's wort group, people with anxious depression were significantly older than those with nonanxious depression ($t(111) = -2.06$, $p = 0.04$). Conversely, in the sertraline group, people with anxious depression were significantly younger than those with nonanxious depression ($t(109) = 2.92$, $p = 0.004$). There was not a significant difference in age for the placebo group. People with anxious depression had higher levels of depression severity than did people with nonanxious depression in all the treatment groups, according to HAM-D score: St John's wort: ($|z|_{1-\alpha/2} = 3.43$, $p = 0.0006$), placebo: ($|z|_{1-\alpha/2} = 4.54$, $p < 0.0001$), and sertraline: ($|z|_{1-\alpha/2} = 4.69$, $p < 0.0001$). In the placebo group, people with anxious depression had significantly higher severity according to CGI-S score than did people with nonanxious depression ($|z|_{1-\alpha/2} = 2.17$, $p = 0.03$). Differences in CGI-S were not significant for the St John's wort or sertraline groups. Differences in GAF and SDS total scores were also not significant by treatment group.

When study continuation was defined as having a HAM-D score at week 8 of treatment, 245 participants, or 72%, remained in the study at week 8. Ninety five people discontinued before week 8. As seen in Table 6, study continuation rates at week 8 were similar for people with melancholic depression and nonmelancholic depression. There were no significant differences in study continuation rates between melancholic and nonmelancholic depression groups by treatment assignment. As seen in Table 7, study continuation rates were also similar for people with anxious depression and nonanxious depression. There was no significant difference in the continuation rates for people with anxious versus nonanxious depression by treatment assignment.

Rates of treatment response (response versus no response) at week 8 by treatment group for those with melancholic and nonmelancholic depression are shown in Table 8. In each treatment by melancholic depression subtype group, more than half of participants did not respond to treatment. Response rates, however, did not differ significantly between the melancholic and nonmelancholic depression groups. There were also no significant differences in treatment response rates by treatment group. For the anxious depression subtype, rates of treatment response at week 8 are shown in Table 9. In each treatment by anxious depression subtype group, more than half of participants did not respond to treatment. Rates of treatment response did not differ significantly between the anxious and nonanxious depression groups. There was also no significant difference within each treatment group by subtype.

As shown in Table 10, the primary ANCOVA using LOCF revealed no significant main effects for treatment ($F_{2,300} = 1.77$, $p = 0.17$) or melancholic depression status at baseline ($F_{1,300} = 1.1.7$, $p = 0.28$) on depression severity at week 8. The covariate of depression severity at baseline was significant ($F_{1,300} = 24.58$, $p < 0.0001$). There was no significant interaction between treatment and baseline melancholic depression subtype level ($F_{2,300} = 0.93$, $p = 0.40$). Melancholic depression status at baseline did have a significant effect on CGI-S score at week 8 ($F_{1,301} = 4.59$, $p = 0.03$) after controlling for baseline severity of depression. The ANCOVA models demonstrated that of the people assigned to St John's wort, the melancholic depression group had higher mean HAM-D, CGI-S, and CGI-I total scores at week 8 (Table 11). These means were adjusted for HAM-D score at baseline. In the placebo group, the mean adjusted HAM-D, CGI-S, CGI-I scores were higher for the melancholic depression group than for the

nonmelancholic depression group. Conversely, in the sertraline group, the mean adjusted HAM-D, CGI-S, CGI-I scores were higher for the people in the nonmelancholic group than in the melancholic group. The nonmelancholic depression group assigned to placebo had the highest adjusted mean HAM-D score at week 8. The primary ANCOVA revealed no significant main effects for treatment ($F_{2,300} = 1.77$, $p = 0.17$) or melancholic depression status at baseline ($F_{1,300} = 1.1.7$, $p = 0.28$) on depression severity at week 8. The covariate of depression severity at baseline was significant ($F_{1,300} = 24.58$, $p < 0.0001$). There was no significant interaction between treatment and baseline melancholic depression subtype level ($F_{2,300} = 0.93$, $p = 0.40$). Melancholic depression status at baseline did have a significant effect on CGI-S score at week 8 ($F_{1,301} = 4.59$, $p = 0.03$) after controlling for baseline severity of depression.

As seen in Table 12, the primary ANCOVA model revealed no main effects for treatment ($F_{2,300} =2.74$, $p = 0.07$) or anxious depression status ($F_{1,300} = 0.04$, $p = 0.84$) on depression severity at week 8. There was no significant interaction between treatment and subtype level ($F_{2,300} =0.72$, $p = 0.49$). The covariate of depression severity at baseline was significant ($F_{1,300}=22.40$, $p < 0.0001$). Treatment had a significant effect on CGI-S score ($F_{1,301} = 3.23$, $p = 0.04$) and on CGI-I ($F_{1,301} = 3.23$, $p = 0.03$) after controlling for baseline severity of depression. The ANCOVA models for the anxious depression subtype showed that, of those who were assigned St John's wort, the anxious depression group had lower mean HAM-D, CGI-S, and CGI-I scores when adjusted for baseline depression severity than did the nonanxious depression group (Table 13). Of the people assigned to placebo and sertraline, the nonanxious depression groups had lower mean HAM-D, CGI-S, and CGI-I scores than did the anxious depression groups, when means

were adjusted for baseline severity of depression. The nonanxious depression St John's wort group had the highest adjusted HAM-D total score of all the treatment by anxious depression subtype groups.

For the mixed effects regression model of longitudinal HAM-D scores in the melancholic depression subtype, an unstructured covariance structure was used after comparing the goodness of fit statistics AIC and BIC. Because of the randomization that was originally done following the run-in phase of the trial, the model was not adjusted for additional covariates. Observations included in this model were those available at baseline and weekly from weeks 1-8. The type 3 fixed effects for the model and the solution for fixed effects indicated that HAM-D scores decreasing over time ($F_{2,1617} = 569.87$, $p < 0.0001$), with an estimated decrease of 1.18 points per unit of time ($p < 0.0001$). The type 3 tests of fixed effects also showed that there were significant effects of melancholic depression status at baseline ($F_{1,1617} = 26.10$, $p < 0.0001$). However, the interaction of treatment assignment by melancholic depression status at baseline is also significant ($F_{2,1617} = 5.81$, $p = 0.003$) and therefore supersedes the main effect of melancholic depression status. In the interaction of treatment by melancholic depression status at baseline, people with nonmelancholic depression assigned to St John's wort have a lower estimated HAM-D total score when compared to those with melancholic depression assigned to St John's wort ($p = 0.0009$).

An unstructured covariance structure was also used for the mixed effects regression model of longitudinal HAM-D scores in the anxious depression subtype due to this structure having the smallest AIC and BIC of the structures evaluated. The model was not adjusted for covariates because of the randomization that was done in the trial

after the run-in phase. Observations included in this model were those available at baseline and weekly from weeks 1-8. As shown in Table 15, the type 3 tests of fixed effects for the model showed significant effects for treatment ($F_{2,1951} = 6.94$, $p = 0.001$), anxious depression status at baseline ($F_{1,1951} = 84.83$, $p < 0.0001$), and time ($F_{1,319} = 304.17$, $p < 0.0001$). The solution for fixed effects showed that HAM-D total scores were predicted to be lower for people with nonanxious depression than for people with anxious depression ($p < 0.0001$), holding all other variables constant. HAM-D scores also decreased with time ($p < 0.0001$), holding all other variables constant.

Multivariate logistic regression models were used to examine the effects of depression subtype and treatment assignment in the probability of achieving treatment response. However, as shown in Table 16, none of these odds ratios were statistically significant. Results were similar for the logistic regression for the anxious depression subtype. As shown in Table 17, none of the odds ratios was statistically significant.

The clinical characteristics of participants at week 8 were examined by subtype and treatment response. In the melancholic depression group, those who did not respond to treatment had more severe depression and were more functionally impaired when compared to people with who did respond to treatment across all measures ($p <0.0001$ for HAM-D, CGI-S, GAF, and SDS scores) (Table 18). As can thus be expected, those who did not respond to treatment showed less improvement in depression severity as measured by the CGI-I ($p <0.0001$). Results were similar in the nonmelancholic depression group in that the differences in clinical measures were significantly different between those who responded to treatment and those who did not ($p <0.0001$ for HAM-

D, CGI-S, GAF, and SDS scores). These differences also held true for the anxious versus nonanxious depression groups (Table 19).

When the clinical scores at week 8 for participants in with melancholic or nonmelancholic depression were examined by treatment assignment, there were significant differences in all scores between the group with melancholic depression who responded and those who did not (Table 20). The group with melancholic depression assigned to St John's wort who did not respond had the highest mean HAM-D and CGI-S scores of all the treatment groups. Differences between those who responded and did not in the melancholic group were significant for HAM-D, CGI-S, CGI-I, and GAF scores for all treatment assignments (p <0.001). In the nonmelancholic group, the difference in SDS score was significant for people assigned to St John's wort (p = 0.03) and for people assigned to placebo (p = 0.001). Differences were also significant between those who did and did not respond by treatment assignment in anxious depression group (p <0.05) (Table 21).

Chapter 5: Discussion

The Hypericum trial appears to have enrolled high numbers of people with melancholic depression and with anxious depression, with higher rates of melancholic depression and anxious depression than have been seen in trials like STAR*D. Study results did not support the hypothesis that people with melancholic depression would be more likely than people with nonmelancholic depression to be male and have a shorter duration of current depressive episode. Participants with melancholic depression were, however, found to be significantly older, to have more severe depression, and to have more functional impairment than participants with nonmelancholic depression. The hypothesis that participants with anxious depression were more likely than participants with nonanxious depression to have more severe depression and more functional impairment was confirmed. Participants with anxious depression, however, were not found to be more likely than participants with nonanxious depression to be female.

Contrary to the original hypothesis, neither depression subtype nor treatment assignment appeared to have an effect on study continuation by week 8. Response rates also did not appear to be influenced by depression subtype or treatment assignment. People with melancholic depression were not significantly more likely than those with nonmelancholic depression to respond to treatment with St John's wort or with sertraline. Also, people with melancholic depression were not less likely than those with nonmelancholic depression to respond to placebo. People with anxious depression were not significantly less likely than people with nonanxious depression to respond to treatment with St John's wort or sertraline. People with anxious depression were not significantly more likely than those with nonanxious depression to respond to placebo.

There were several differences between the ANCOVA models with LOCF and the mixed effects linear regression models for HAM-D total scores in both the melancholic and anxious subtypes. The ANCOVA for the melancholic subtype did not find a main effect of treatment assignment, melancholic depression status at baseline, or the interaction between treatment and subtype status for depression severity at the end of treatment. A significant effect for treatment was also not found in the mixed effects model. The mixed effects model, however, demonstrated a significant influence of melancholic depression status at baseline and for the interaction between treatment assignment and baseline melancholic status. In the anxious depression subtype, the ANCOVA also did not find that anxious depression status at baseline, treatment assignment, and the interaction between anxious depression status and treatment assignment were significant main effects. The mixed effects regression model showed significant effects for treatment and anxious depression status at baseline but not for the interaction of the two. Given the methodological flaws associated with assuming one week's outcome will hold true for the next week, as is the practice in LOCF, the mixed effects regression model might be a better indicator of true effects.

Depression subtype does appear to be at least somewhat related to depression severity throughout treatment but it is difficult to use subtype status as a predictor of treatment response. The discrepant findings between the ANCOVA, mixed effects regression, and logistic regression analyses seem to indicate that additional research needs to be conducted. St John's wort appeared to be more effective for people with nonmelancholic depression rather than melancholic depression. Sertraline seemed more effective for people with nonanxious versus anxious depression and for people with

melancholic versus nonmelancholic depression. Placebo seemed more effective for people with nonanxious versus anxious depression and for people with nonmelancholic versus melancholic depression. These effects should be interpreted with more than a modicum of caution though, since effects were not consistently significant.

Study Strengths and Limitations

This study has several strengths. Since this was a secondary data analysis from a trial conducted by experienced academic researchers with funding and oversight by the federal government, the conduct of the original trial is presumed to be free of conflict of interest that might have biased the results. The results of this study can be used to generate hypotheses for future studies that could have practical clinical implications for developing personalized medicine for depression. An additional strength is that the operational definitions of several variables such as depression severity at baseline, the criteria for anxious depression, and treatment response have been used in other studies. This should allow for comparisons to be made across similar studies.

While this study has several strengths, there are also numerous limitations. Since this is a secondary data analysis, the variables are limited to what was originally collected. The influence of typical demographic variables such as education, income, and marital status could not be analyzed because they were not included in the original trial assessments. Additionally, the atypical depression subtype could not be included because the original trial did not ask about some of the distinguishing atypical features such as hypersomnia and rejection sensitivity. Another limitation inherent in using the Hypericum trial data instead of designing a new study is that people might not have been followed long enough to adequately capture the treatment effect. Anxious depression, in

particular, has been associated with delayed treatment response and it is possible that the original 8-week trial was not long enough to capture response. Furthermore, analysis of outcome was limited to interpretations of HAM-D, CGI-S, and CGI-I. These measures might not be the best indicators of treatment response. Additionally, the quality, including the inter-rater reliability, of the ratings of these measures is unknown. The quality of the clinical interviews during which these ratings has been previously shown to impact the ability separate the effects of active drug from placebo.[111]

Because the original randomization did not consider depression subtype, people with anxious depression and those with melancholic depression were not equally distributed throughout the treatment groups. Additionally, using scores on the modified HES to identify melancholic depression might not be a clinically valid method. The criteria for melancholic depression might have been overly broad or strict, thus possibly misclassifying people on the melancholic subtype. The HAM-D anxiety/somatization factor accounts for a limited amount of anxiety symptoms and thus might not have adequately identified all the participants who had anxious depression.

In general, because the Hypericum trial had numerous inclusion and exclusion criteria, it is possible that the sample of participants was overly homogenous and thus these results do not adequately represent effects that might be found in real-world clinical settings. An analysis of participants in the STAR*D trial, a trial which had broad inclusion criteria, demonstrated that regular phase III depression intervention trials do not recruit typical depression patients.[112] This analysis estimated that less than one in four people with depression would be eligible for a typical trial. The Hypericum trial exclusion criteria might have particularly reduced variability in the predictor variables.

Excluding those with a positive drug urine screen or those at risk for suicide might have disproportionately excluded people with anxious depression since people with anxious depression have an increased risk of suicide and drug abuse.[10, 22, 90, 97] Furthermore, the generalizability of the Hypericum trial results could be further compromised by the characteristics of people interested in receiving CAM treatment. Given that people who use CAM therapies for mental disorders tend to differ on sociodemographic and clinical variables compared to people who do not use CAM, participants in trials of CAM therapies might not adequately represent the general population.[37]

Efficacy also might not have been demonstrated in the Hypericum trial because of issues with treatment adherence. Because St John's wort is readily available as an over-the-counter supplement, participants randomized to placebo or sertraline might have been able to independently take St John's wort in addition to their assigned treatment. Another post-hoc analysis of the Hypericum trial did indeed find that, based on blood tests, 17% of participants assigned to placebo had taken St John's wort.[107] Treatment adherence could also have been compromised if people did not take all doses of their treatment assignment throughout the trial. This might have been more likely in the active treatments of St John's wort and sertraline due to possible side effects. Adverse events were more frequent in the St John's wort and sertraline groups in the original trial although it is unclear if these events impacted treatment adherence.[6]

Using the LOCF imputation technique for missing observations at week 8 likely produced inaccurate estimates of response. LOCF could have led to bias in underestimating or overestimating the effect of treatment. Previous studies have shown that LOCF can favor the experimental treatment or placebo.[113] To avoid the effects of

missing data on analysis, future studies should be designed to reduce as much participant attrition as possible by minimizing participant burden and treatment side effects.[114] Since it is likely that there will always be some missing observations in trials, studies should also explore analyzing data using mixed-effects regression models, multiple imputation, or nonignorable models, as appropriate.[115]

Conclusion

This study explored the effect of the melancholic and anxious depression subtypes on response to treatment with St John's wort, placebo, or sertraline. This work contributes to the body of literature on personalized medicine for major depression by exploring how to predict what treatments will work well for a particular person with depression. Current treatments fail to produce complete symptom remission in many people with depression. If it were possible to accurately predict who would respond to what treatment, treatment efficiency could be increased and the public health burden of depression could be decreased. Being able to differentiate treatment responses between people with depression could contribute to the determination of a biological mechanism of depression. This study also adds to the understanding of how well CAM treatments work for major depression. Clinicians need a better understanding of the evidence about the usefulness of CAM treatments for depression.[116]

While this thesis research provides a unique contribution to the field, additional research is necessary to provide conclusive evidence of how depression subtypes do or do not moderate treatment response. In addition to determining the etiology of major depression, a large longitudinal epidemiologic study examining the prevalence of the depression subtypes in the general population would be very valuable in informing

clinician decision-making.[63] Clinical trials of treatments that randomize participants according to subtype might also provide more conclusive evidence regarding whether or not depression subtypes moderate antidepressant treatment response. New studies might also want to consider the instability of depression subtypes over time and examine how people's depression subtypes might change throughout treatment.

Tables

Table 1: Frequency of melancholic and anxious depression subtypes in the Hypericum trial

No. (% of total population)	Melancholic	Nonmelancholic	Total
Anxious	146 (42.94)	98 (28.82)	244 (71.76)
Nonanxious	73 (21.47)	23 (6.76)	96 (28.24)
Total	219 (64.41)	121 (35.59)	340 (100)

Table 2: Baseline demographic and clinical characteristics of participants with melancholic versus nonmelancholic depression

Characteristic*	Melancholic Depression (n= 219)	Nonmelancholic Depression (n=121)	*p*
Age, mean (SD), yrs	44.03 (13.23)	40.54 (12.90)	0.02
Sex, No. (%)			0.87
Male	74 (33.79)	42 (34.71)	
Female	145 (66.21)	79 (65.29)	
Race, No. (%)			0.25
Asian/ Pacific Islander	6 (2.74)	8 (6.61)	
Black	25 (11.42)	10 (8.26)	
White	163 (74.43)	94 (77.69)	
Hispanic	21 (9.59)	6 (4.96)	
Native American	1 (0.46)	1 (0.83)	
Other	3 (1.37)	2 (1.65)	
Duration of current depression episode, No. (%)			0.42
<6 mos	60 (27.40)	32 (26.45)	
6-24 mos	99 (45.21)	38 (31.40)	
>24 mos	60 (27.40)	51 (42.15)	
HAM-D total score, mean (SD)	23.51 (2.80)	21.17 (1.74)	< 0.0001
CGI-S score, mean (SD)	4.31 (0.52)	4.09 (0.37)	< 0.0001
GAF total score, mean (SD)	53.23 (4.75)	54.36 (4.37)	0.02
SDS total score, mean (SD)	16.27 (8.18)	14.11 (7.71)	0.02

Abbreviations: HAM-D = Hamilton Depression Rating Scale, CGI-S = Clinical Global Impression Severity, GAF = Global Assessment of Functioning, SDS = Sheehan Disability Scale
* Lower clinical scores denote less severity or impairment, except for the GAF. For the GAF, lower scores denote more functional impairment.

Table 3: Baseline demographic and clinical characteristics of participants with anxious versus nonanxious depression

Characteristic*	**Anxious Depression (n=244)**	**Nonanxious Depression (n=96)**	***p***
Age, mean (SD), yrs	42.38 (12.68)	43.82 (14.45)	0.37
Sex, No. (%)			0.48
Male	86 (35.25)	30 (31.25)	
Female	158 (64.75)	66 (68.75)	
Race, No. (%)			0.43
Asian/ Pacific Islander	12 (4.92)	2 (2.08)	
Black	27 (11.07)	8 (8.33)	
White	181 (74.18)	76 (79.17)	
Hispanic	20 (8.20)	7 (7.29)	
Native American	2 (0.82)	0	
Other	2 (0.82)	3 (3.13)	
Duration of current depression episode, No. (%)			0.07
<6 mos	75 (30.74)	18 (18.56)	
6-24 mos	91 (37.30)	46 (47.42)	
>24 mos	78 (31.97)	33 (34.02)	
HAM-D total score, mean (SD)	23.41 (2.75)	21.19 (1.51)	<0.0001
CGI-S score, mean (SD)	4.26 (0.49)	4.17 (0.45)	0.04
GAF total score, mean (SD)	53.56 (4.62)	53.81 (4.70)	0.66
SDS total score, mean (SD)	16.00 (8.10)	14.19 (7.88)	0.04

Abbreviations: HAM-D = Hamilton Depression Rating Scale, CGI-S = Clinical Global Impression Severity, GAF = Global Assessment of Functioning, SDS = Sheehan Disability Scale

* Lower clinical scores denote less severity or impairment, except for the GAF. For the GAF, lower scores denote more functional impairment.

Table 4: Baseline demographic and clinical characteristics of participants with melancholic versus nonmelancholic depression, by treatment group

	SJW (n=113)			Placebo (n=116)			Setraline (n=111)		
Characteristic	**Mel (n=74)**	**Nonmel (n=39)**	***p***	**Mel (n=72)**	**Nonmel (n=44)**	***p***	**Mel (n=73)**	**Nonmel (n=38)**	***p***
Age, mean (SD), yrs	44.3 (13.7)	42.3 (13.0)	0.45	41.5 (12.6)	38.9 (11.2)	0.26	46.2 (13.0)	40.7 (14.7)	0.04
Sex, No. (%)			0.08			0.052			0.57
Male	22 (29.73)	18 (46.15)		29 (40.28)	10 (22.73)		23 (31.51)	14 (36.84)	
Female	52 (70.27)	21 (53.85)		43 (59.72)	34 (77.27)		50 (68.49)	24 (63.16)	
Race, No. (%)			0.58			0.26			0.70
Asian/ Pacific Islander	2 (2.70)	2 (5.13)		2 (2.78)	3 (6.82)		2 (2.74)	3 (7.89)	
Black	8 (10.81)	2 (5.13)		7 (9.72)	5 (11.36)		10 (13.70)	3 (7.89)	
White	55 (74.32)	32 (82.05)		54 (75.00)	34 (77.27)		54 (73.97)	28 (76.68)	
Hispanic	8 (10.81)	2 (5.13)		8 (11.11)	1 (2.27)		5 (6.85)	3 (7.89)	
Native American	0	0		0	1 (2.27)		1 (1.37)	0	
Other	1 (1.35)	1 (2.56)		1 (1.39)	0		1 (1.37)	1 (2.63)	

Abbreviation: SJW = St John's wort, Mel = melancholic depression, Nonmel = nonmelancholic depression

Table 4, continued: Baseline demographic and clinical characteristics of participants with melancholic versus nonmelancholic depression, by treatment group

	SJW (n=113)			Placebo (n=116)			Setraline (n=111)		
Characteristic	**Mel (n=74)**	**Nonmel (n=39)**	***p***	**Mel (n=72)**	**Nonmel (n=44)**	***p***	**Mel (n=73)**	**Nonmel (n=38)**	***P***
Duration of current depression episode, No.(%)			0.07			0.37			0.40
<6 mos	15 (20.27)	9 (23.08)		21 (29.17)	12 (27.27)		24 (32.88)	11 (28.95)	
6-24 mos	38 (51.35)	12 (30.77)		36 (50.00)	13 (29.55)		25 (34.25)	13 (34.21)	
>24 mos	21 (28.38)	18 (46.2)		15 (20.83)	19 (43.2)		24 (32.88)	14 (36.84)	
HAM-D total score, mean (SD)	24.05 (2.69)	21.28 (1.68)	<0.0001	23.47 (3.00)	21.52 (1.64)	0.0003	23.00 (2.64)	21.61 (1.92)	0.0003
CGI-S score, mean (SD)	4.34 (0.53)	4.10 (0.38)	0.008	4.31 (0.52)	4.07 (0.33)	0.008	4.29 (0.51)	4.11 (0.39)	0.02
GAF total score, mean (SD)	52.35 (5.24)	54.67 (3.95)	0.02	53.74 (4.32)	54.75 (4.21)	0.11	53.61 (4.56)	53.50 (4.90)	0.91
SDS total score, mean (SD)	16.42 (8.31)	14.56 (8.12)	0.25	16.71 (8.38)	12.95 (7.10)	0.02	15.66 (7.92)	14.92 (7.95)	0.62

Abbreviation: SJW = St John's wort, Mel = melancholic depression, Nonmel = nonmelancholic depression

Table 5: Baseline demographic and clinical characteristics of participants with anxious versus nonanxious depression, by treatment group

	SJW (n=113)			Placebo (n=116)			Sertraline (n=111)		
Characteristic	**Anx (n=83)**	**Nonanx (n=30)**	***p***	**Anx (n=85)**	**Nonanx (n=31)**	***p***	**Anx (n=76)**	**Nonanx (n=35)**	***p***
Age, mean (SD), yrs	45.16 (12.99)	39.34 (13.93)	0.04	40.17 (11.60)	41.41 (13.58)	0.63	41.83 (13.11)	49.79 (13.97)	0.004
Sex, No. (%)			0.78			0.48			0.11
Male	30 (36.14)	10 (33.33)		27 (31.76)	12 (38.71)		29 (38.16)	8 (22.86)	
Female	53 (63.86)	20 (66.67)		58 (68.24)	19 (61.29)		47 (61.84)	27 (77.14)	
Race, No. (%)			0.51			0.23			0.93
Asian/ Pacific Islander	4 (4.82)	0		4 (4.71)	1 (3.23)		4 (5.26)	1 (2.86)	
Black	8 (9.64)	2 (6.67)		11 (12.94)	1 (3.23)		8 (10.53)	5 (14.29)	
White	64 (77.11)	23 (76.67)		61 (71.76)	27 (87.10)		56 (73.68)	26 (74.29)	
Hispanic	6 (7.23)	4 (13.33)		8 (9.41)	1 (3.23)		6 (7.89)	2 (5.71)	
Native American	0	0		1 (1.18)	0		1 (1.32)	0	
Other	1 (1.20)	1 (3.33)		0	1 (3.23)		1 (1.32)	1 (2.86)	

Abbreviation: SJW = St John's wort, Anx = anxious depression, Nonanx = nonanxious depression

Table 5, continued: Baseline demographic and clinical characteristics of participants with anxious versus nonanxious depression, by treatment group

	SJW (n=113)			Placebo (n=116)			Sertraline (n=111)		
Characteristic	**Anx (n=83)**	**Nonanx (n=30)**	***p***	**Anx (n=85)**	**Nonanx (n=31)**	***p***	**Anx (n=76)**	**Nonanx (n=35)**	***p***
Duration of current depression episode, No. (%)			0.14			0.90			0.09
<6 mos	23 (27.71)	1 (3.33)		26 (30.59)	7 (22.58)		26 (34.21)	9 (25.71)	
6-24 mos	33 (39.76)	17 (56.67)		31 (36.47)	18 (58.06)		27 (35.53)	11 (31.4)	
>24 mos	27 (32.53)	12 (40.00)		28 (32.94)	6 (19.35)		23 (30.26)	15 (42.9)	
HAM-D total score, mean (SD)	23.61 (2.82)	21.67 (1.81)	0.0006	23.36 (2.85)	21.00 (1.24)	<0.0001	23.25 (2.57)	20.94 (1.39)	<0.0001
CGI-S score, mean (SD)	4.28 (0.50)	4.20 (0.48)	0.35	4.27 (0.52)	4.06 (0.25)	0.03	4.22 (0.45)	4.23 (0.55)	0.64
GAF total score, mean (SD)	53.09 (4.87)	53.40 (5.3)	0.66	53.98 (4.21)	54.52 (4.54)	0.45	53.59 (4.81)	53.54 (4.39)	0.66
SDS total score, mean (SD)	16.55 (8.01)	13.55 (8.69)	0.06	15.48 (8.4)	14.90 (7.25)	0.64	15.97 (7.90)	14.09 (7.90)	0.25

Abbreviation: SJW = St John's wort, Anx = anxious depression, Nonanx = nonanxious depression

Table 6: Comparison of study retention rates at week 8 of treatment for participants with melancholic versus nonmelancholic depression, by treatment group

Note: Study retention is defined as having a HAM-D score at week 8 of treatment

	Study participants remaining at week 8			
No (%)	**SJW**	**Placebo**	**Sertraline**	**Total**
Melancholic depression at baseline	57 (77.0)	55 (76.4)	50 (68.5)	162 (74)
Nonmelancholic depression at baseline	25 (64.1)	29 (65.9)	29 (76.3)	83 (69)
p	0.14	0.22	0.43	0.27
Total	82 (72.6)	84 (72.4)	79 (71.2)	245 (72.1)

*Percentages in cells are of number of participants in treatment assignment by depression subtype status. For example, 57 of the 74 participants (77.0%) with melancholic depression in the St John's wort group remained in the study at week 8.

Table 7: Comparison of study retention rates at week 8 of treatment for participants with anxious versus nonanxious depression, by treatment group

Note: Study retention is defined as having a HAM-D score at week 8 of treatment

	Study participants remaining at week 8			
No. (%)	**SJW**	**Placebo**	**Sertraline**	**Total**
Anxious depression at baseline	64 (77.1)	59 (69.4)	54 (71.1)	177 (72.5)
Nonanxious depression at baseline	18 (60.0)	25 (80.6)	25(71.4)	68 (69.4)
p	0.07	0.23	0.94	0.66
Total	82 (72.6)	84 (72.4)	79 (71.2)	245 (72.1)

*Percentages in cells are of number of participants in treatment assignment by depression subtype status. For example, 64 of the 83 participants (77.1%) with anxious depression in the St John's wort group remained in the study at week 8.

Table 8: Clinical response rates of participants with melancholic versus nonmelancholic depression at week 8, by treatment group

	SJW		Placebo		Sertraline		
Response*, No. (%)	**Mel**	**Nonmel**	**Mel**	**Nonmel**	**Mel**	**Nonmel**	**Total**
Any response	27 (36.49)	16 (41.03)	31 (43.06)	16 (36.36)	31 (42.47)	18 (47.37)	139 (40.89)
No response	47 (63.51)	23 (58.97)	41 (56.94)	28 (63.64)	42 (57.53)	20 (52.63)	201 (59.12)
p	0.64		0.48		0.63		0.90
Total	74 (21.76)	39 (11.47)	72 (21.18)	44 (12.94)	73 (21.47)	38 (11.18)	340 (100)

*Any response = HAM-D less than or equal to 12, 50% reduction in HAM-D score from baseline to end of treatment, and a CGI-I score of 1 or 2. LOCF is being used for HAM-D and CGI-I scores.

Abbreviations: SJW=St John's wort, Mel = melancholic depression, Nonmel = nonmelancholic depression

Table 9: Clinical response rates of participants with anxious versus nonanxious depression at week 8, by treatment group

	SJW		Placebo		Sertraline		
Response*, No. (%)	**Anx**	**Nonanx**	**Anx**	**Nonanx**	**Anx**	**Nonanx**	**Total**
Any response	33 (39.76)	10 (33.33)	34 (40.00)	13 (41.94)	33 (43.42)	16 (45.71)	139 (40.88)
No response	50 (60.24)	20 (66.67)	51 (60.00)	18 (58.06)	43 (56.58)	19 (54.29)	201 (59.12)
p	0.54		0.85		0.82		0.95
Total	83 (24.41)	30 (8.82)	85 (25.00)	31 (9.12)	76 (22.35)	35 (10.29)	340 (100)

*Any response = HAM-D less than or equal to 12, 50% reduction in HAM-D score from baseline to end of treatment, and a CGI-I score of 1 or 2. LOCF used for HAM-D and CGI-I scores.
Abbreviations: SJW=St John's wort, Anx = anxious depression, Nonanx = nonanxious depression

Table 10: Results of ANCOVA models at week 8 of treatment for the melancholic depression subtype

Source	**Baseline severity of depression (Covariate)**	**Treatment**	**Melancholic depression status at baseline**	**Treatment by melancholic depression status at baseline**
Outcome*	**F (*p*)**	**F (*p*)****	**F (*p***	**F (*p*)**
HAM-D total score	$F_{1,300}$ = 24.58 (<0.0001)	$F_{2,300}$ = 1.77 (0.17)	$F_{1,300}$ = 1.17 (0.28)	$F_{2,300}$ = 0.93 (0.40)
CGI-S score	$F_{1,301}$ = 15.43 (0.0001)	$F_{2,301}$ = 2.10 (0.12)	$F_{1,301}$ = 4.59 (0.03)	$F_{2,301}$ = 0.31 (0.73)
CGI-I score	$F_{1,301}$ = 6.91 (0.009)	$F_{2,301}$ = 1.93 (0.15)	$F_{1,301}$ = 3.02 (0.08)	$F_{2,301}$ = 1.19 (0.31)
GAF total score	$F_{1,301}$ = 4.57 (0.03)	$F_{2,301}$ = 0.04 (0.96)	$F_{1,301}$ = 3.05 (0.08)	$F_{2,301}$ = 0.46 (0.63)
SDS total score	$F_{1,234}$= 21.71 (<0.0001)	$F_{2,234}$ = 0.12 (0.89)	$F_{1,234}$= 0.01 (0.91)	$F_{2,234}$ = 0.61 (0.55)

* LOCF scores are used for the clinical characteristics, except for SDS total score. SDS total score is not using LOCF because SDS was only assessed as baseline and week 8
** *F* statistics correspond to the type III SS

Table 11: Adjusted means of clinical characteristics at week 8 of participants with melancholic versus nonmelancholic depression, by treatment group

Note: means are adjusted for severity of depression at baseline, as measured by HAM-D total score*

	Melancholic Depression			Nonmelancholic Depression		
Characteristic**	**SJW (n=68)**	**Placebo (n=69)**	**Sert (n=61)**	**SJW (n=34)**	**Placebo (n=42)**	**Sert (n=33)**
HAM-D total score, mean	13.83	12.78	11.21	13.22	14.81	12.73
CGI-S score, mean	3.10	2.95	2.72	3.28	3.39	3.04
CGI-I score, mean	2.54	2.36	2.11	2.51	2.86	2.45
GAF total score, mean	65.46	66.82	67.31	64.72	63.06	63.52
SDS total score, mean	11.48	10.38	11.68	11.75	11.90	10.24

* Dependent variable used week 8 HAM-D total if available (n=245) or last available observation after week 1 (n=62). 33 cases not included in analysis due to lack of observations beyond week 1.
** LOCF scores are used for the clinical characteristics, except for SDS total score. SDS total score is not using LOCF because SDS was only assessed as baseline and week 8

Table 12: Results of ANCOVA models at week 8 of treatment for the anxious depression subtype

Source	Baseline severity of depression (Covariate)	Treatment	Anxious depression status at baseline	Treatment by anxious depression status at baseline
Outcome*	**F** (*p*)**	**F (*p*)**	**F (*p*)**	**F (*p*)**
HAM-D total score	$F_{1,300}$= 22.40 (<0.0001)	$F_{2,300}$= 2.74 (0.07)	$F_{1,300}$= 0.04 (0.84)	$F_{2,300}$= 0.72(0.49)
CGI-S score	$F_{1,301}$=9.88 (0.002)	$F_{2,301}$= 3.23 (0.04)	$F_{1,301}$= 0.00 (0.99)	$F_{2,301}$ = 1.71(0.18)
CGI-I score	$F_{1,301}$=4.65 (0.03)	$F_{2,301}$= 3.48 (0.03)	$F_{1,301}$=0.21 (0.64)	$F_{2,301}$=0.73 (0.48)
GAF total score	$F_{1,301}$=1.83 (0.18)	$F_{2,301}$= 0.33 (0.72)	$F_{1,301}$=0.34(0.56)	$F_{2,301}$=0.37(0.69)
SDS total score	$F_{1,234}$=24.17 (<0.0001)	$F_{2,234}$= 0.05 (0.95)	$F_{1,234}$= 1.01 (0.31)	$F_{2,234}$= 0.17(0.84)

* LOCF scores are used, except for SDS total score. SDS total score is not using LOCF because SDS was only assessed as baseline and week 8

** *F* statistics correspond to the type III SS

Table 13: Adjusted means of clinical characteristics at week 8 of participants with anxious versus nonanxious depression, by treatment group

Note: means are adjusted for severity of depression at baseline, as measured by HAM-D total score*

	Anxious Depression			Nonanxious Depression		
Characteristic**	**SJW (n=83)**	**Placebo (n=85)**	**Sert (n=76)**	**SJW (n=30)**	**Placebo (n=31)**	**Sert (n=35)**
HAM-D total score, mean	13.2	13.6	12.0	14.9	13.4	11.1
CGI-S score, mean	3.07	3.18	2.87	3.44	2.94	2.73
CGI-I score, mean	2.52	2.67	2.24	2.86	2.59	2.21
GAF total score, mean	65.4	65.1	65.3	64.5	66.4	67.8
SDS total score, mean	11.4	10.2	10.9	12.1	12.4	11.8

* Dependent variable used week 8 HAM-D total if available (n=245) or last available observation after week 1 (n=62). 33 cases not included in analysis due to lack of observations beyond week 1.
** LOCF scores are used for the clinical characteristics, except for SDS total score. SDS total score is not using LOCF because SDS was only assessed as baseline and week 8

Table 14: Solutions for fixed effects and type 3 tests of fixed effects in the mixed effects regression model for severity of depression (HAM-D total score) in the melancholic depression subtype

Solutions for Fixed Effects			
Effect	**Estimate**	**Standard Error**	***p***
Intercept	21.25	0.44	<0.0001
Treatment assignment (sertraline vs placebo)	0.06	0.63	0.92
Treatment assignment (SJW vs placebo)	1.93	0.62	0.002
Melancholic depression (nonmelancholic vs melancholic)	-0.69	0.72	0.34
Time	-1.18	0.05	<0.0001
Sertraline by melancholic status (nonmelancholic vs melancholic)	-1.02	1.04	0.33
SJW by melancholic status (nonmelancholic vs melancholic)	-3.46	1.04	0.0009
Type 3 Tests of Fixed Effects			
Effect		**$F_{df\ num,\ df\ den}$**	***p***
Treatment assignment		$F_{2,1617}$=0.79	0.45
Melancholic depression status at baseline		$F_{1,1617}$=26.10	<0.0001
Time		$F_{1,319}$=569.87	<0.0001
Treatment assignment at baseline by melancholic depression status at baseline		$F_{2,1617}$=5.81	0.003

Table 15: Solutions for fixed effects and type 3 tests of fixed effects in the mixed effects regression model for severity of depression (HAM-D total score) in the anxious depression subtype

Solutions for Fixed Effects			
Effect	**Estimate**	**Standard Error**	***p***
Intercept	21.83	0.26	<0.0001
Treatment assignment (sertraline vs placebo)	-0.39	0.38	0.30
Treatment assignment (SJW vs placebo)	0.47	0.36	0.19
Anxious depression (nonanxious vs anxious)	-3.12	0.49	<0.0001
Time	-1.19	0.07	<0.0001
Sertraline by anxious status (nonanxious vs anxious)	0.36	0.69	0.60
SJW by anxious status (nonanxious vs anxious)	1.10	0.71	0.12

Type 3 Tests of Fixed Effects		
Effect	$\mathbf{F_{df\ num,\ df\ den}}$	***P***
Treatment assignment	$F_{2,1951}$= 6.94	0.001
Anxious depression status at baseline	$F_{1,1951}$=84.83	<0.0001
Time	$F_{1,319}$=304.17	<0.0001
Treatment assignment at baseline by anxious depression status at baseline	$F_{2,1951}$=1.26	0.29

Table 16: Odds ratios and 95% confidence intervals for the association between the probability of treatment response, treatment assignment, and melancholic depression subtype

	Odds ratio	Confidence Intervals	*P*
Treatment			
Setraline vs placebo	1.17	0.69-1.97	0.39
SJW vs placebo	0.90	0.53-1.54	0.45
Melancholic depression status			
Nonmelancholic vs. melancholic	1.03	0.66-1.62	0.90

Table 17: Odds ratios and 95% confidence intervals for the association between the probability of treatment response, treatment assignment, and anxious depression subtype

	Odds ratio	Confidence Intervals	*P*
Treatment			
Sertraline vs placebo	1.16	0.69-1.97	0.39
SJW vs placebo	0.90	0.53-1.53	0.45
Anxious depression status			
Nonanxious vs anxious	0.98	0.60-1.58	0.92

Table 18: Clinical characteristics at week 8 of participants with melancholic versus nonmelancholic depression, by treatment response

	Melancholic Depression			Nonmelancholic Depression		
Characteristic*	**Response (n=89)**	**No response (n=130)**	***p***	**Response (n=50)**	**No response (n=71)**	***P***
HAM-D total score, mean (SD)	6.40 (3.39)	18.85 (4.94)	<0.0001	6.68 (3.15)	17.61 (5.21)	<0.0001
CGI-S score, mean (SD)	1.97 (0.70)	3.86 (0.69)	<0.0001	2.10 (0.68)	3.95 (0.67)	<0.0001
CGI-I score, mean (SD)	1.40 (0.49)	3.31 (0.92)	<0.0001	1.48 (0.50)	3.53 (0.85)	<0.0001
GAF total score, mean (SD)	75.01 (8.93)	58.72 (7.94)	<0.0001	72.90 (13.50)	57.57 (8.52)	<0.0001
SDS total score, mean (SD)	6.83 (7.16)	17.00 (7.00)	<0.0001	6.52 (5.59)	14.63 (7.92)	<0.0001

*Scores use LOCF except for SDS total score

Table 19: Clinical characteristics at week 8 of participants with anxious versus nonanxious depression by treatment response

	Anxious Depression			Nonanxious Depression		
Characteristic*	**Response (n=100)**	**No response (n=144)**	***p***	**Response (n=39)**	**No response (n= 57)**	***P***
HAM-D total score, mean (SD)	6.83 (3.30)	18.94 (4.87)	<0.0001	5.67 (3.17)	17.02 (5.31)	<0.0001
CGI-S score, mean (SD)	2.08 (0.72)	3.93 (0.64)	<0.0001	1.85 (0.59)	3.78 (0.79)	<0.0001
CGI-I score, mean (SD)	1.42 (0.50)	3.43 (0.90)	<0.0001	1.46 (0.51)	3.28 (0.91)	<0.0001
GAF total score, mean (SD)	73.82 (11.51)	57.79 (7.30)	<0.0001	75.36 (8.74)	59.71 (10.01)	<0.0001
SDS total score, mean (SD)	7.20 (6.49)	16.76 (7.22)	<0.0001	5.33 (6.88)	15.17 (7.61)	<0.0001

*Scores use LOCF except for SDS total score

Table 20: Clinical characteristics at week 8 of participants with melancholic or nonmelancholic depression by treatment assignment and treatment response

Melancholic Depression									
	St John's Wort			Placebo			Sertraline		
Any response	**Yes**	**No**	***P***	**Yes**	**No**	***p***	**Yes**	**No**	***p***
Characteristic*									
HAM-D total score, mean (SD)	7.11 (3.40)	20.00 (4.50)	<0.0001	5.94 (3.12)	19.42 (5.52)	<0.0001	6.26 (3.65)	16.57 (4.04)	<0.0001
CGI-S score, mean (SD)	2.07 (0.68)	4.00 (0.63)	<0.0001	1.94 (0.63)	3.92 (0.71)	<0.0001	1.90 (0.79)	3.60 (0.67)	<0.0001
CGI-I score, mean (SD)	1.48 (0.51)	3.46 (0.87)	<0.0001	1.35 (0.49)	3.47 (0.89)	<0.0001	1.39 (0.50)	2.90 (0.92)	<0.0001
GAF total score, mean (SD)	74.81 (7.35)	57.98 (8.08)	<0.0001	76.52 (8.90)	58.11 (8.22)	<0.0001	73.68 (10.18)	60.53 (7.34)	<0.0001
SDS total score, mean (SD)	7.44 (6.73)	17.27 (7.72)	<0.0001	5.41 (6.44)	17.38 (7.24)	<0.0001	7.75 (8.18)	16.18 (5.84)	0.0001

*Scores use LOCF except for SDS total score

Table 20, continued: Clinical characteristics at week 8 of participants with melancholic or nonmelancholic depression by treatment assignment and treatment response

Nonmelancholic Depression									
	St John's Wort			Placebo			Setraline		
Any response	**Yes**	**No**	***p***	**Yes**	**No**	***P***	**Yes**	**No**	***p***
Characteristic*									
HAM-D total score, mean (SD)	6.44 (3.58)	16.83 (5.00)	<0.0001	6.44 (2.53)	18.35 (5.70)	<0.0001	7.11 (3.36)	16.27 (4.73)	<0.0001
CGI-S score, mean (SD)	2.00 (0.82)	4.11 (0.58)	<0.0001	2.13 (0.62)	3.96 (0.77)	<0.0001	2.17 (0.62)	3.75 (0.58)	<0.0001
CGI-I score, mean (SD)	1.44 (0.51)	3.50 (0.79)	<0.0001	1.50 (0.52)	3.69 (0.84)	<0.0001	1.50 (0.51)	3.31 (0.95)	<0.0001
GAF total score, mean (SD)	75.81 (10.85)	56.67 (10.28)	<0.0001	74.00 (6.76)	57.54 (8.79)	<0.0001	69.33 (18.94)	58.63 (5.89)	0.0002
SDS total score, mean (SD)	6.31 (6.58)	14.75 (9.27)	0.03	6.00 (5.83)	16.50 (6.79)	0.001	7.11 (4.85)	12.45 (7.65)	0.08

*Scores use LOCF except for SDS total score

Table 21: Clinical characteristics at week 8 of participants with anxious or nonanxious depression by treatment assignment and treatment response

Anxious Depression									
	St John's Wort			Placebo			Sertraline		
Any response	**Yes**	**No**	***p***	**Yes**	**No**	***p***	**Yes**	**No**	***p***
Characteristic*									
HAM-D total score, mean (SD)	7.03 (3.47)	19.33 (4.09)	<0.0001	6.38 (2.92)	19.54 (5.71)	<0.0001	7.09 (3.55)	17.53 (4.28)	<0.0001
CGI-S score, mean (SD)	2.06 (0.75)	4.00 (0.49)	<0.0001	2.06 (0.69)	4.06 (0.73)	<0.0001	212 (0.74)	3.67 (0.60)	<0.0001
CGI-I score, mean (SD)	1.48 (0.51)	3.43 (0.77)	<0.0001	1.38 (0.49)	3.65 (0.91)	<0.0001	1.40 (0.50)	3.12 (0.96)	<0.0001
GAF total score, mean (SD)	74.88 (9.31)	57.21 (6.26)	<0.0001	75.97 (7.78)	56.92 (8.25)	<0.0001	70.55 (15.62)	59.79 (6.89)	<0.0001
SDS total score, mean (SD)	7.74 (6.69)	17.03 (8.02)	<0.0001	6.09 (6.43)	17.48 (6.53)	<0.0001	7.78 (6.40)	15.55 (6.95)	0.0003

*Scores use LOCF except for SDS total score

Table 21, continued: Clinical characteristics at week 8 of participants with anxious or nonanxious depression by treatment assignment and treatment response

Nonanxious Depression									
	St John's Wort			Placebo			Sertraline		
Any response	**Yes**	**No**	***p***	**Yes**	**No**	***p***	**Yes**	**No**	***p***
Characteristic*									
HAM-D total score, mean (SD)	6.30 (3.47)	18.29 (6.43)	0.0001	5.38 (2.84)	17.31 (4.92)	<0.0001	5.50 (3.37)	15.00 (3.70)	<0.0001
CGI-S score, mean (SD)	2.00 (0.67)	4.12 (0.86)	<0.0001	1.85 (0.38)	3.56 (0.63)	<0.0001	1.75 (0.68)	3.62 (0.77)	<0.0001
CGI-I score, mean (SD)	1.40 (0.52)	3.59 (1.00)	<0.0001	1.46 (0.52)	3.31 (0.70)	<0.0001	1.50 (0.52)	2.85 (0.90)	0.0001
GAF total score, mean (SD)	76.20 (6.65)	58.47 (13.23)	0.001	74.85 (9.66)	60.75 (8.40)	0.001	75.25 (9.58)	60.08 (7.09)	0.0005
SDS total score, mean (SD)	4.00 (5.69)	15.18 (8.78)	0.02	4.33 (5.52)	16.38 (8.12)	0.002	6.86 (8.46)	13.73 (6.02)	0.01

*Scores use LOCF except for SDS total score

Figures

Figure 1: Enrollment and outcomes during the Hypericum Trial[6]

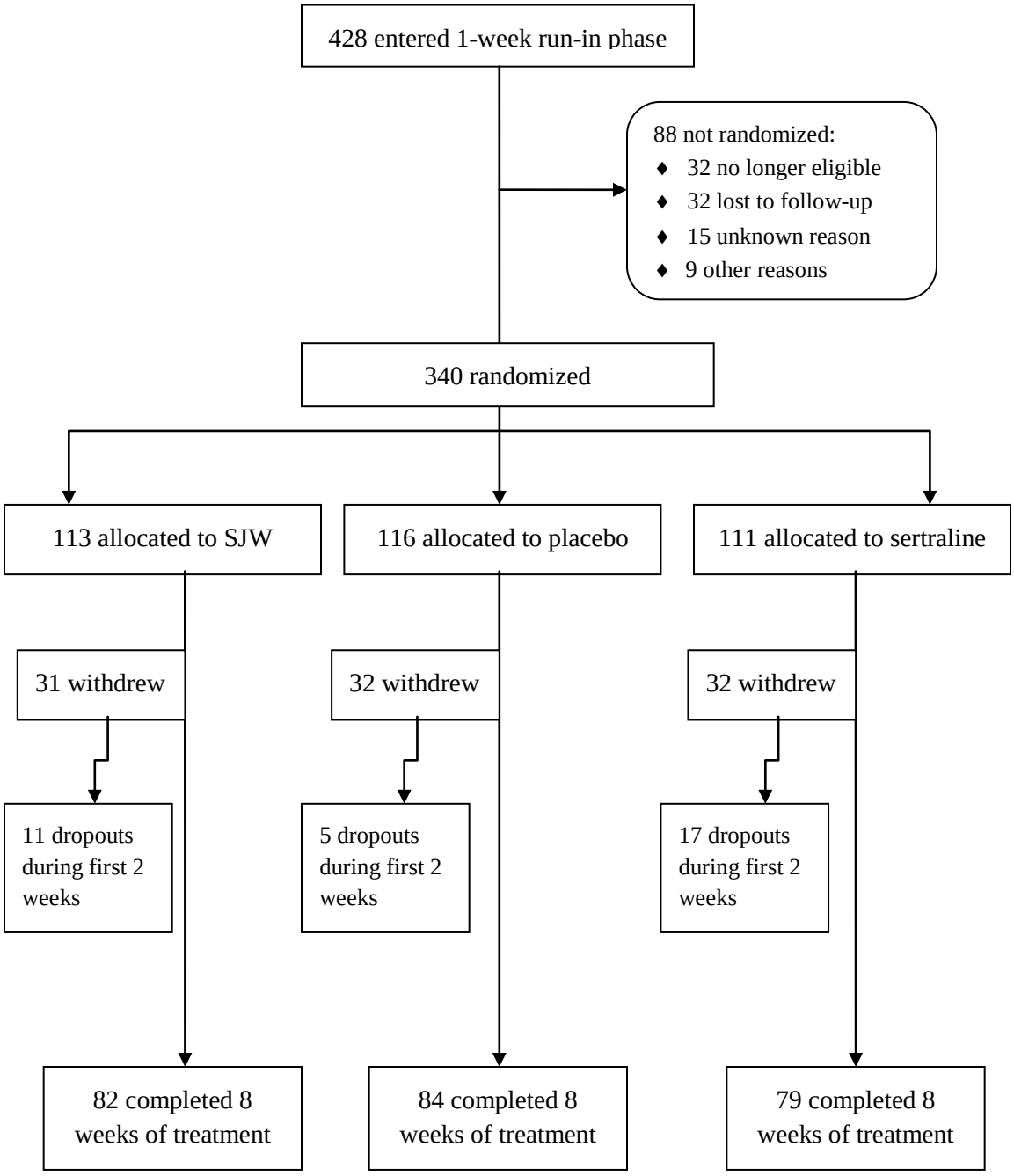


Figure 2: Cumulative frequency distribution of modified Hamilton Endogenomorphy Subscale scores at baseline in the Hypericum trial

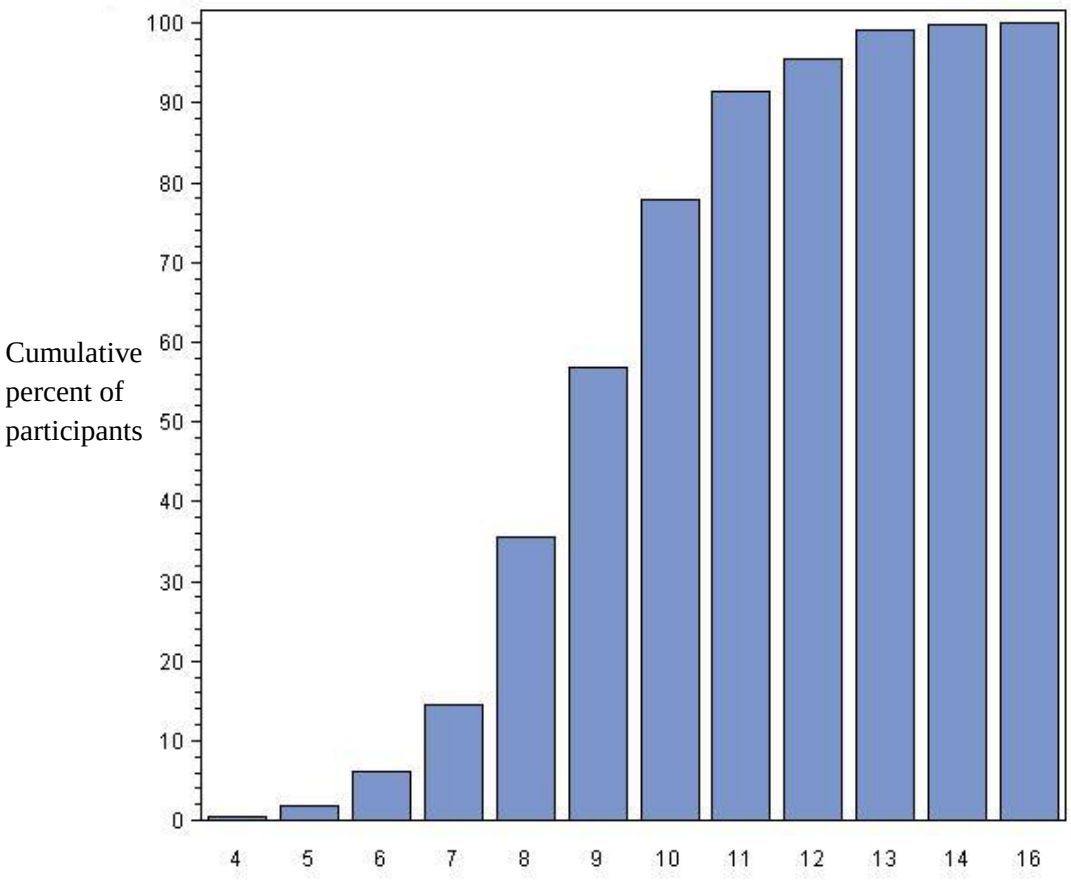


Figure 3: Frequency distribution of modified Hamilton Endogenomorphy Subscale scores at baseline in the Hypericum trial

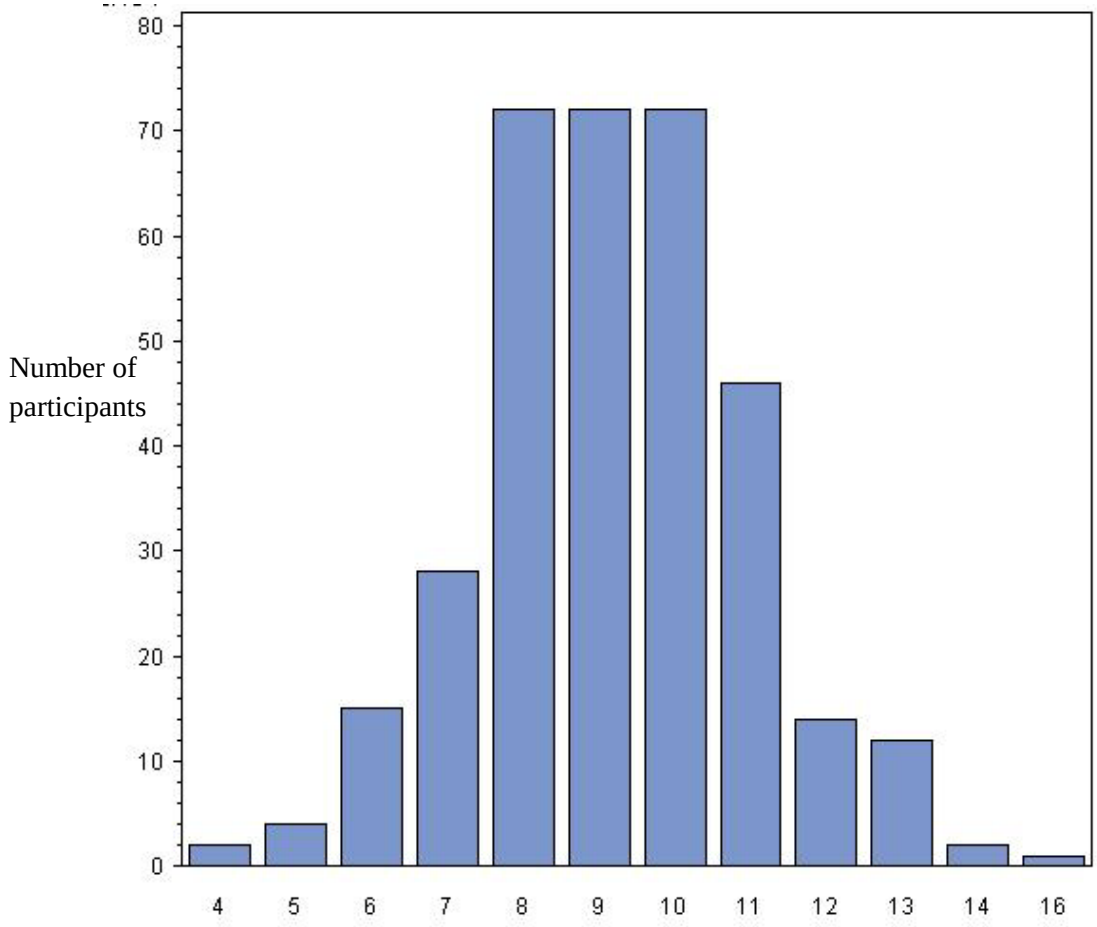


Figure 4: Cumulative frequency distribution of HAM-D anxiety/somatization factor scores at baseline in the Hypericum trial

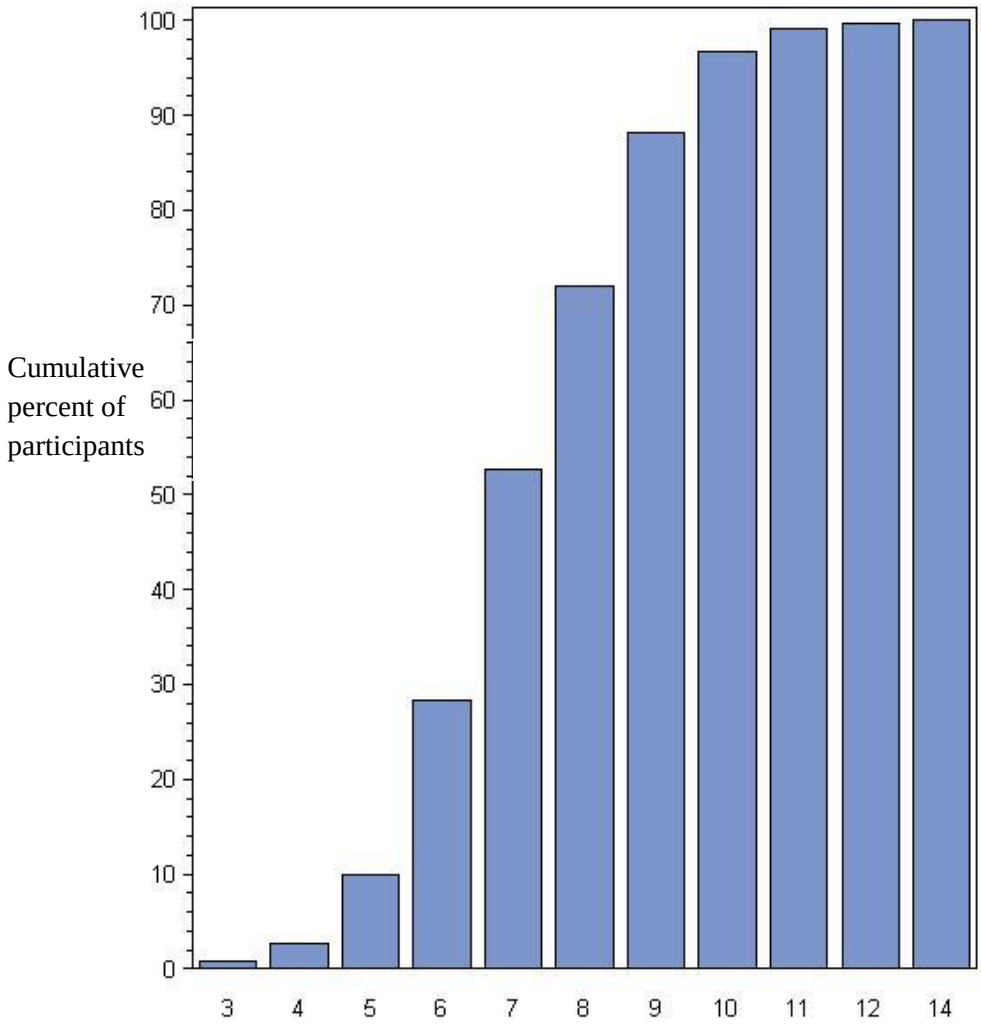
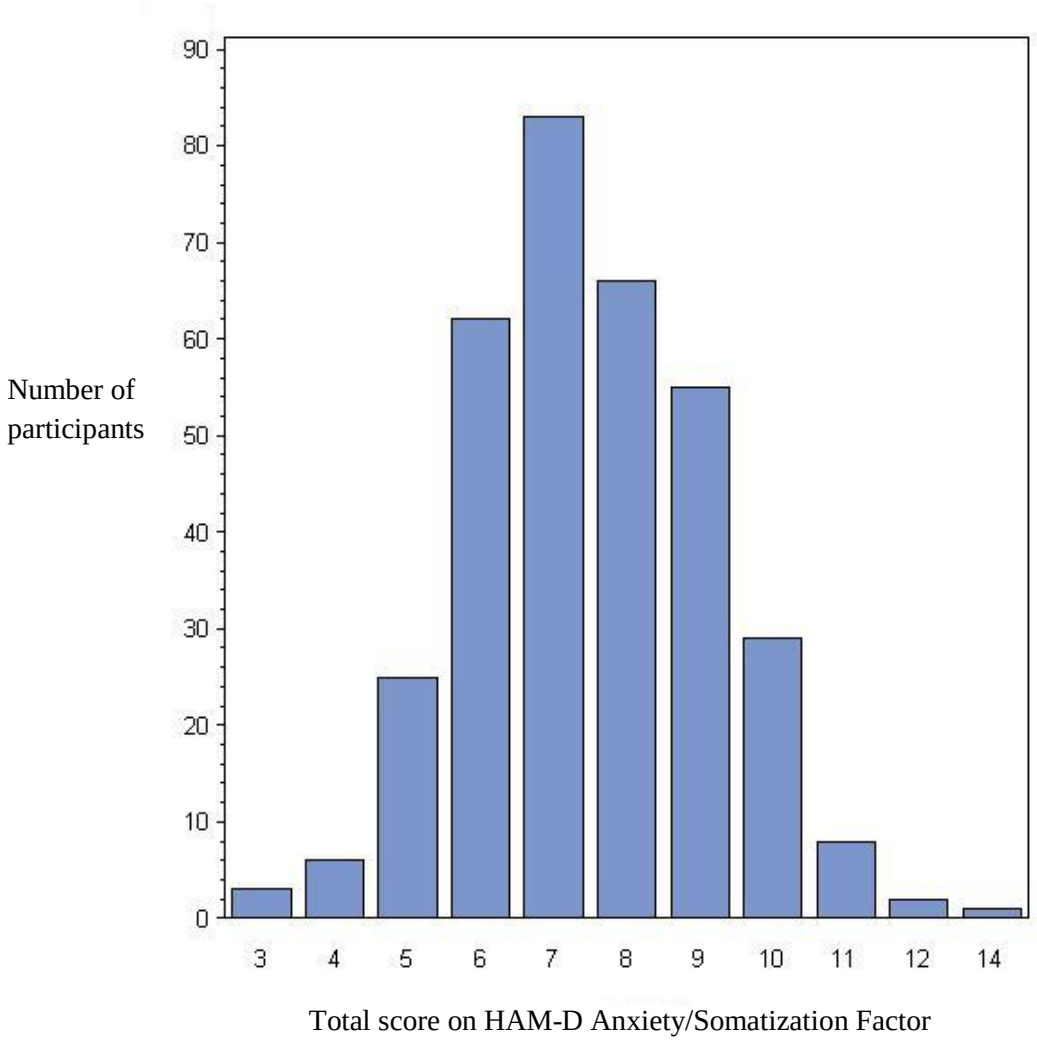


Figure 5: Frequency distribution of HAM-D anxiety/somatization factor scores at baseline in the Hypericum trial



References

1. Kessler R, Chiu W, Demler O, Merikangas K, Walters E. 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.
2. Insel TR, Wang PS. The STAR*D trial: revealing the need for better treatments. *Psychiatr Serv.* Nov 2009;60(11):1466-1467.
3. Walsh BT, Seidman SN, Sysko R, Gould M. Placebo Response in Studies of Major Depression. *JAMA: The Journal of the American Medical Association.* April 10, 2002 2002;287(14):1840-1847.
4. Freeman MP, Mischoulon D, Tedeschini E, et al. Complementary and alternative medicine for major depressive disorder: a meta-analysis of patient characteristics, placebo-response rates, and treatment outcomes relative to standard antidepressants. *J Clin Psychiatry.* Jun 2010;71(6):682-688.
5. Hammerness P, Basch E, Ulbricht C, et al. St. John's Wort: A Systematic Review of Adverse Effects and Drug Interactions for the Consultation Psychiatrist. *Psychosomatics.* August 1, 2003 2003;44(4):271-282.
6. Hypericum Depression Trial Study Group. Effect of Hypericum perforatum (St John's Wort) in Major Depressive Disorder: A Randomized Controlled Trial. *JAMA.* April 10, 2002 2002;287(14):1807-1814.
7. Freeman MP, Fava M, Lake J, Trivedi MH, Wisner KL, Mischoulon D. Complementary and alternative medicine in major depressive disorder: the American Psychiatric Association Task Force report. *J Clin Psychiatry.* Jun 2010;71(6):669-681.
8. Trivedi MH, Thase ME, Fava M, et al. Adjunctive aripiprazole in major depressive disorder: analysis of efficacy and safety in patients with anxious and atypical features. *J Clin Psychiatry.* Dec 2008;69(12):1928-1936.
9. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR).* Fourth ed. Alexandria, VA: American Psychiatric Association; 2000.
10. Fava M, Alpert JE, Carmin CN, et al. Clinical correlates and symptom patterns of anxious depression among patients with major depressive disorder in STAR*D. *Psychol Med.* Oct 2004;34(7):1299-1308.
11. Ustun TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJL. Global burden of depressive disorders in the year 2000. *The British Journal of Psychiatry.* May 1, 2004 2004;184(5):386-392.
12. Kessler RC, Berglund P, Demler O, et al. The Epidemiology of Major Depressive Disorder. *JAMA: The Journal of the American Medical Association.* June 18, 2003 2003;289(23):3095-3105.
13. World Health Organization. *The global burden of disease: 2004 update.* Geneva, Switzerland2008.
14. World Health Organization. *The world health report, 2002: Reducing risks, promoting healthy life.* Geneva, Switzerland2002.
15. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. Fourth ed. Washington, DC: American Psychiatric Association; 1994.

16. National Institute of Mental Health. Questions and Answers about the NIMH Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Study — Background. 2006; http://www.nimh.nih.gov/trials/practical/stard/backgroundstudy.shtml. Accessed 1/20/2011.
17. Wang PS, Beck A, Berglund P, et al. Chronic medical conditions and work performance in the health and work performance questionnaire calibration surveys. *J Occup Environ Med.* Dec 2003;45(12):1303-1311.
18. Ebmeier KP, Donaghey C, Steele JD. Recent developments and current controversies in depression. *Lancet.* Jan 14 2006;367(9505):153-167.
19. Osby U, Brandt L, Correia N, Ekbom A, Sparen P. Excess Mortality in Bipolar and Unipolar Disorder in Sweden. *Arch Gen Psychiatry.* September 1, 2001 2001;58(9):844-850.
20. Piatt EE, Munetz MR, Ritter C. An Examination of Premature Mortality Among Decedents With Serious Mental Illness and Those in the General Population. *Psychiatr Serv.* July 1, 2010 2010;61(7):663-668.
21. Wang PS, Lane M, Olfson M, Pincus HA, Wells KB, Kessler RC. Twelve-Month Use of Mental Health Services in the United States: Results From the National Comorbidity Survey Replication. *Arch Gen Psychiatry.* June 1, 2005 2005;62(6):629-640.
22. Tollefson GD, Holman SL, Sayler ME, Potvin JH. Fluoxetine, placebo, and tricyclic antidepressants in major depression with and without anxious features. *J Clin Psychiatry.* Feb 1994;55(2):50-59.
23. Simon GE, Perlis RH. Personalized medicine for depression: can we match patients with treatments? *Am J Psychiatry.* Dec 2010;167(12):1445-1455.
24. Simon GE, Von Korff M, Rutter CM, Peterson DA. Treatment Process and Outcomes for Managed Care Patients Receiving New Antidepressant Prescriptions From Psychiatrists and Primary Care Physicians. *Arch Gen Psychiatry.* April 1, 2001 2001;58(4):395-401.
25. Papakostas GI, Fava M. Predictors, moderators, and mediators (correlates) of treatment outcome in major depressive disorder. *Dialogues in Clinical Neuroscience.* 2008;10(4):439-451.
26. Clayton PJ. Depression subtyping: treatment implications. *J Clin Psychiatry.* 1998;59 Suppl 16:5-12; discussion 40-12.
27. Khan AY, Carrithers J, Preskorn SH, et al. Clinical and demographic factors associated with DSM-IV melancholic depression. *Ann Clin Psychiatry.* Apr-Jun 2006;18(2):91-98.
28. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial Severity and Antidepressant Benefits: A Meta-Analysis of Data Submitted to the Food and Drug Administration. *PLoS Med.* 2008;5(2):e45.
29. Linde K, Berner M, Kriston L. *St John's wort for major depression*: The Cochrane Collaboration;2008.
30. IMS Health. IMS Health Reports U.S. Prescription Sales Grew 5.1 Percent in 2009, to $300.3 Billion. . 2010; Press release. Available at: http://www.imshealth.com/portal/site/imshealth/menuitem.a46c6d4df3db4b3d88f611019418c22a/?vgnextoid=d690a27e9d5b7210VgnVCM100000ed152ca2RCR

D&vgnextchannel=41a67900b55a5110VgnVCM10000071812ca2RCRD&vgnextfmt=default. Accessed March 22, 2011, 2011.
31. Berman RM, Sporn J, Belanoff JK, Schatzberg A, Charney DS. Principles of the pharmacotherapy of depression. In: Charney DS, Nestler EJ, eds. *Neurobiology of Mental Illness*. Second ed. New York: Oxford University Press, Inc.; 2004:491-511.
32. Rosenzweig-Lipson S, Beyer CE, Hughes ZA, et al. Differentiating antidepressants of the future: Efficacy and safety. *Pharmacology & Therapeutics.* 2007;113(1):134-153.
33. Schechter LE, Ring RH, Beyer CE, et al. Innovative approaches for the development of antidepressant drugs: current and future strategies. *NeuroRx.* Oct 2005;2(4):590-611.
34. Kirsch I, Moore TJ, Scoboria A, Nicholls S. The emperor's new drugs: An analysis of antidepressant medication data submitted to the U.S. Food and Drug Administration *Prevention and Treatment.* 2002;5.
35. Thase ME, Entsuah AR, Rudolph RL. Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors. *The British Journal of Psychiatry.* March 1, 2001 2001;178(3):234-241.
36. National Center for Complementary and Alternative Medicine. What Is Complementary and Alternative Medicine? 2010; http://nccam.nih.gov/health/whatiscam/. Accessed March 19, 2011.
37. Woodward AT, Bullard KM, Taylor RJ, et al. Complementary and alternative medicine for mental disorders among African Americans, black Caribbeans, and whites. *Psychiatr Serv.* Oct 2009;60(10):1342-1349.
38. Wille SMR, Cooreman SG, Neels HM, Lambert WEE. Relevant Issues in the Monitoring and the Toxicology of Antidepressants. *Critical Reviews in Clinical Laboratory Sciences.* 2008;45(1):25-89.
39. Fava M, Alpert J, Nierenberg AA, et al. A Double-blind, randomized trial of St John's wort, fluoxetine, and placebo in major depressive disorder. *J Clin Psychopharmacol.* Oct 2005;25(5):441-447.
40. Nierenberg AA, Mischoulon D, DeCecco LM. St John's wort: a critique of antidepressant efficacy and possible mechanisms of action. In: Mischoulon D, Rosenbaum JF, eds. *Natural Medications for Psychiatric Disorders*. Philadelphia, PA: Lippincott Williams and Wilkins; 2002:3-12.
41. Mannel M, Kuhn U, Schmidt U, Ploch M, Murck H. St. John's wort extract LI160 for the treatment of depression with atypical features - a double-blind, randomized, and placebo-controlled trial. *J Psychiatr Res.* Sep 2010;44(12):760-767.
42. St. John's wort (Hypericum perforatum L.): Natural Standard Bottom Line Monograph. Natural Standard Research Collaboration; 2011. http://www.naturalstandard.org/databases/herbssupplements/patient-stjohnswort.asp. Accessed March 20, 2011.
43. Shelton RC, Keller MB, Gelenberg A, et al. Effectiveness of St John's Wort in Major Depression. *JAMA: The Journal of the American Medical Association.* April 18, 2001 2001;285(15):1978-1986.

44. Williams JB. Standardizing the Hamilton Depression Rating Scale: past, present, and future. *Eur Arch Psychiatry Clin Neurosci.* 2001;251 Suppl 2:II6-12.
45. Hamilton M. Development of a rating scale for primary depressive illness. *Brit J Soc Clin Psychol.* 1967;6:278-296.
46. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry.* Feb 1960;23:56-62.
47. Hedlund J, Viewer B. The Hamilton Rating Scale for Depression: A comprehensive review. *Journal of Operational Psychiatry.* 1979;10(2):149-165.
48. Moberg PJ, Lazarus LW, Mesholam RI, et al. Comparison of the standard and structured interview guide for the Hamilton Depression Rating Scale in depressed geriatric inpatients. *Am J Geriatr Psychiatry.* Winter 2001;9(1):35-40.
49. Williams JB. A structured interview guide for the Hamilton Depression Rating Scale. *Arch Gen Psychiatry.* Aug 1988;45(8):742-747.
50. Bagby RM, Ryder AG, Schuller DR, Marshall MB. The Hamilton Depression Rating Scale: Has the Gold Standard Become a Lead Weight? *Am J Psychiatry.* December 1, 2004 2004;161(12):2163-2177.
51. Evans KR, Sills T, DeBrota DJ, Gelwicks S, Engelhardt N, Santor D. An Item Response analysis of the Hamilton Depression Rating Scale using shared data from two pharmaceutical companies. *J Psychiatr Res.* May-Jun 2004;38(3):275-284.
52. Moller HJ. Methodological aspects in the assessment of severity of depression by the Hamilton Depression Scale. *Eur Arch Psychiatry Clin Neurosci.* 2001;251 Suppl 2:II13-20.
53. Guy W. *ECDEU Assessment Manual for Psychopharmacology, revised.* Rockville, MD: National Institute of Mental Health;1976.
54. Busner J, Targum SD. The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry (Edgmont).* Jul 2007;4(7):28-37.
55. Kadouri A, Corruble E, Falissard B. The improved Clinical Global Impression Scale (iCGI): development and validation in depression. *BMC Psychiatry.* 2007;7:7.
56. Prien RF, Levine J. Research and methodological issues for evaluating the therapeutic effectiveness of antidepressant drugs. *Psychopharmacol Bull.* Spring 1984;20(2):250-257.
57. Rush AJ. The varied clinical presentations of major depressive disorder. *J Clin Psychiatry.* 2007;68 Suppl 8:4-10.
58. Seo HJ, Jung YE, Kim TS, et al. Distinctive clinical characteristics and suicidal tendencies of patients with anxious depression. *J Nerv Ment Dis.* Jan 2011;199(1):42-48.
59. Kendler K, Eaves L, Walters E, Neale M, Heath A, Kessler R. The identification and validation of distinct depressive syndromes in a population-based sample of female twins. *Archives of general psychiatry.* 1996;53(5):391 - 399.
60. Lamers F, de Jonge P, Nolen WA, et al. Identifying depressive subtypes in a large cohort study: results from the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry.* Dec 2010;71(12):1582-1589.

61. Sullivan PF, Kessler RC, Kendler KS. Latent class analysis of lifetime depressive symptoms in the national comorbidity survey. *Am J Psychiatry.* Oct 1998;155(10):1398-1406.
62. Sullivan PF, Prescott CA, Kendler KS. The subtypes of major depression in a twin registry. *J Affect Disord.* Apr 2002;68(2-3):273-284.
63. Kessing LV. Epidemiology of subtypes of depression. *Acta psychiatrica Scandinavica.* 2007;115:85-89.
64. Rosenbaum JF. Depression and its subtypes: a treatment update. Conclusion. *J Clin Psychiatry.* 1998;59 Suppl 18:37-38.
65. Posternak MA, Zimmerman M. The prevalence of atypical features across mood, anxiety, and personality disorders. *Comprehensive Psychiatry.* 2002;43(4):253-262.
66. Nierenberg AA, Alpert JE, Pava J, Rosenbaum JF, Fava M. Course and treatment of atypical depression. *J Clin Psychiatry.* 1998;59 Suppl 18:5-9.
67. McGrath PJ, Khan AY, Trivedi MH, et al. Response to a selective serotonin reuptake inhibitor (citalopram) in major depressive disorder with melancholic features: a STAR*D report. *J Clin Psychiatry.* Dec 2008;69(12):1847-1855.
68. Fink M, Taylor MA. Resurrecting melancholia. *Acta psychiatrica Scandinavica.* 2007;115:14-20.
69. Spitzer RL, Endicott J, Robins E. Research Diagnostic Criteria: Rationale and Reliability. *Arch Gen Psychiatry.* June 1, 1978 1978;35(6):773-782.
70. Kaplan A. Whither Melancholia? *Psychiatric Times.* 2010;20(1). Accessed March 14, 2011.
71. Coryell W. The facets of melancholia. *Acta psychiatrica Scandinavica.* 2007;115:31-36.
72. Rush A, Weissenburger J. Melancholic symptom features and DSM-IV. *Am J Psychiatry.* April 1, 1994 1994;151(4):489-498.
73. Leventhal AM, Rehm LP. The empirical status of melancholia: implications for psychology. *Clin Psychol Rev.* Jan 2005;25(1):25-44.
74. Parker G. Classifying depression: should paradigms lost be regained? *American journal of psychiatry.* 2000;157(8):1195 - 1203.
75. Akin D, Manier DH, Sanders-Bush E, Shelton RC. Signal transduction abnormalities in melancholic depression. *Int J Neuropsychopharmacol.* Mar 2005;8(1):5-16.
76. Hickie I, Naismith S, Ward PB, et al. Reduced hippocampal volumes and memory loss in patients with early- and late-onset depression. *Br J Psychiatry.* Mar 2005;186:197-202.
77. Willeit M, Praschak-Rieder N, Neumeister A, et al. A polymorphism (5-HTTLPR) in the serotonin transporter promoter gene is associated with DSM-IV depression subtypes in seasonal affective disorder. *Mol Psychiatry.* Nov 2003;8(11):942-946.
78. Simons AD, Thase ME. Biological markers, treatment outcome, and 1-year follow-up in endogenous depression: electroencephalographic sleep studies and response to cognitive therapy. *J Consult Clin Psychol.* Jun 1992;60(3):392-401.

79. Stage KB, Bech P, Gram LF, Kragh-Sorensen P, Rosenberg C, Ohrberg S. Are in-patient depressives more often of the melancholic subtype? Danish University Antidepressant Group. *Acta Psychiatr Scand.* Dec 1998;98(6):432-436.
80. Angst J, Scheidegger P, Stabl M. Efficacy of moclobemide in different patient groups. Results of new subscales of the Hamilton Depression Rating Scale. *Clin Neuropharmacol.* 1993;16 Suppl 2:S55-62.
81. Coryell W, Turner R. Outcome with desipramine therapy in subtypes of nonpsychotic major depression. *J Affect Disord.* Sep 1985;9(2):149-154.
82. Paykel ES, Hollyman JA, Freeling P, Sedgwick P. Predictors of therapeutic benefit from amitriptyline in mild depression: a general practice placebo-controlled trial. *J Affect Disord.* Jan-Feb 1988;14(1):83-95.
83. Peselow E, Sanfilipo M, Difiglia C, Fieve R. Melancholic/endogenous depression and response to somatic treatment and placebo. *Am J Psychiatry.* October 1, 1992 1992;149(10):1324-1334.
84. Rush AJ, Wisniewski SR, Warden D, et al. Selecting among second-step antidepressant medication monotherapies: predictive value of clinical, demographic, or first-step treatment features. *Arch Gen Psychiatry.* Aug 2008;65(8):870-880.
85. Roose SP, Glassman AH, Attia E, Woodring S. Comparative efficacy of selective serotonin reuptake inhibitors and tricyclics in the treatment of melancholia. *Am J Psychiatry.* Dec 1994;151(12):1735-1739.
86. Mapother E. Discussion On Manic-Depressive Psychosis. *The British Medical Journal.* 1926;2(3436):872-879.
87. Fava M, Martinez JM, Greist J, et al. The efficacy and tolerability of duloxetine in the treatment of anxious versus non-anxious depression: a post-hoc analysis of an open-label outpatient study. *Ann Clin Psychiatry.* Jul-Sep 2007;19(3):187-195.
88. Sanderson W, Beck A, Beck J. Syndrome comorbidity in patients with major depression or dysthymia: prevalence and temporal relationships. *Am J Psychiatry.* August 1, 1990 1990;147(8):1025-1028.
89. Wiethoff K, Bauer M, Baghai TC, et al. Prevalence and treatment outcome in anxious versus nonanxious depression: results from the German Algorithm Project. *J Clin Psychiatry.* Aug 2010;71(8):1047-1054.
90. Fava M, Rush AJ, Alpert JE, et al. What clinical and symptom features and comorbid disorders characterize outpatients with anxious major depressive disorder: a replication and extension. *Can J Psychiatry.* Nov 2006;51(13):823-835.
91. Papakostas GI, Trivedi MH, Alpert JE, et al. Efficacy of bupropion and the selective serotonin reuptake inhibitors in the treatment of anxiety symptoms in major depressive disorder: a meta-analysis of individual patient data from 10 double-blind, randomized clinical trials. *J Psychiatr Res.* Jan 2008;42(2):134-140.
92. Fawcett J, Barkin RL. A meta-analysis of eight randomized, double-blind, controlled clinical trials of mirtazapine for the treatment of patients with major depression and symptoms of anxiety. *J Clin Psychiatry.* Mar 1998;59(3):123-127.
93. Fava M, Judge R, Hoog SL, Nilsson ME, Koke SC. Fluoxetine versus sertraline and paroxetine in major depressive disorder: changes in weight with long-term treatment. *J Clin Psychiatry.* Nov 2000;61(11):863-867.

94. Fava M, Rush AJ, Alpert JE, et al. Difference in treatment outcome in outpatients with anxious versus nonanxious depression: a STAR*D report. *Am J Psychiatry.* Mar 2008;165(3):342-351.
95. Joffe RT, Bagby RM, Levitt A. Anxious and nonanxious depression. *Am J Psychiatry.* Aug 1993;150(8):1257-1258.
96. Van Valkenburg C, Akiskal HS, Puzantian V, Rosenthal T. Anxious depressions : Clinical, family history, and naturalistic outcome -- comparisons with panic and major depressive disorders. *Journal of Affective Disorders.* 1984;6(1):67-82.
97. Papakostas GI, McGrath P, Stewart J, et al. Psychic and somatic anxiety symptoms as predictors of response to fluoxetine in major depressive disorder. *Psychiatry Research.* 2008;161(1):116-120.
98. Papakostas GI, Larsen K. Testing anxious depression as a predictor and moderator of symptom improvement in major depressive disorder during treatment with escitalopram. *Eur Arch Psychiatry Clin Neurosci.* Sep 22 2010.
99. Davidson JRT, Meoni P, Haudiquet V, Cantillon M, Hackett D. Achieving remission with venlafaxine and fluoxetine in major depression: its relationship to anxiety symptoms. *Depression & Anxiety (1091-4269).* 2002;16:4-13.
100. Kim JE, Yoon SJ, Kim J, et al. Efficacy and tolerability of mirtazapine in treating major depressive disorder with anxiety symptoms: an 8-week open-label randomised paroxetine-controlled trial. *International Journal of Clinical Practice.* 2011;65(3):323-329.
101. Clayton P, Grove W, Coryell W, Keller M, Hirschfeld R, Fawcett J. Follow-up and family study of anxious depression. *Am J Psychiatry.* November 1, 1991 1991;148(11):1512-1517.
102. Dunner DL. Management of anxiety disorders: The added challenge of comorbidity. *Depression and anxiety.* 2001;13(2):57-71.
103. Mitchell AJ. Two-week delay in onset of action of antidepressants: new evidence. *Br J Psychiatry.* Feb 2006;188:105-106.
104. National Institute of Mental Health. Policy for Distribution of Limited Access Datasets From NIMH Clinical Trials. 2010; http://www.nimh.nih.gov/trials/datasets/nimh-policy-for-distribution-of-data.shtml. Accessed March 13, 2011.
105. National Institute of Mental Health. Treatment of Major Depression With St. John's Wort (Hypericum). http://www.clinicaltrials.gov/show/NCT00005013. Accessed March 13, 2011, 2011.
106. First M, Spitzer R, Gibbon M, Williams J. *Structured Clinical Interview for DSM-IV Axis I Disorders, Research Version, Modified for NIMH Clinical Trial of Hypericum in Depressive Disorder (SCID-Hypericum).* New York, NY1997.
107. Vitiello B, Shader RI, Parker CB, et al. Hyperforin plasma level as a marker of treatment adherence in the National Institutes of Health Hypericum Depression Trial. *J Clin Psychopharmacol.* Jun 2005;25(3):243-249.
108. Kovacs M, Rush AJ, Beck AT, Hollon SD. Depressed outpatients treated with cognitive therapy or pharmacotherapy. A one-year follow-up. *Arch Gen Psychiatry.* Jan 1981;38(1):33-39.

109. Thase ME, Hersen M, Bellack AS, Himmelhoch JM, Kupfer DJ. Validation of a Hamilton subscale for endogenomorphic depression. *Journal of Affective Disorders.* 1983;5(3):267-278.
110. Cleary P, Guy W. Factor analysis of Hamilton Depression Scale. *Drugs Expt Clin Res.* 1977;1:115-120.
111. Kobak KA, Brown B, Sharp I, et al. Sources of unreliability in depression ratings. *J Clin Psychopharmacol.* Feb 2009;29(1):82-85.
112. Wisniewski SR, Rush AJ, Nierenberg AA, et al. Can Phase III Trial Results of Antidepressant Medications Be Generalized to Clinical Practice? A STAR*D Report. *Am J Psychiatry.* May 1, 2009 2009;166(5):599-607.
113. Molenberghs G, Thijs H, Jansen I, et al. Analyzing incomplete longitudinal clinical trial data. *Biostatistics.* July 1, 2004 2004;5(3):445-464.
114. Leon AC, Mallinckrodt CH, Chuang-Stein C, Archibald DG, Archer GE, Chartier K. Attrition in Randomized Controlled Clinical Trials: Methodological Issues in Psychopharmacology. *Biological Psychiatry.* 2006;59(11):1001-1005.
115. Siddique J, Brown CH, Hedeker D, et al. Missing Data in Longitudinal Trials -- Part B, Analytic Issues. *Psychiatric Annals.* 2008;38(12):793-801.
116. Freeman MP. Complementary and Alternative Medicine (CAM): considerations for the treatment of major depressive disorder. *J Clin Psychiatry.* 2009;70 Suppl 5:4-6.