2015-04-23

Latent Variable Approaches for Understanding Heterogeneity in Depression: A Dissertation

Christine M. Ulbricht
University of Massachusetts Medical School

LATENT VARIABLE APPROACHES FOR UNDERSTANDING HETEROGENEITY IN DEPRESSION

A Dissertation Presented

By

CHRISTINE MARIE ULBRICHT

Submitted to the Faculty of the
University of Massachusetts Graduate School of Biomedical Sciences, Worcester
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

APRIL 23, 2015

CLINICAL AND POPULATION HEALTH RESEARCH

LATENT VARIABLE APPROACHES FOR UNDERSTANDING HETEROGENEITY IN DEPRESSION

A Dissertation Presented
By

CHRISTINE MARIE ULBRICHT

The signatures of the Dissertation Defense Committee signify completion and approval as to style and content of the Dissertation

_____

Kate L. Lapane, Ph.D., Thesis Advisor

_____

Robert K. Heinssen, Ph.D., ABPP, Member of Committee

_____

Sharina D. Person, Ph.D., Member of Committee

_____

Anthony J. Rothschild, M.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

_____

Eric Mick, Sc.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the school.

_____

Anthony Carruthers, Ph.D.
Dean of the Graduate School of Biomedical Sciences

Clinical and Population Health Research

April 23, 2015

ACKNOWLEDGEMENTS

This work could not have been completed without the support and guidance of many people, to whom I would like to express my gratitude.

I would like to first profusely thank my mentor and program director Kate L. Lapane, PhD, for her unwavering support throughout the program and for her patience with me and these methods.

I am also grateful to my dissertation committee of Robert K. Heinssen, PhD, ABPP; Eric Mick, ScD; Sharina Person, PhD; and Anthony J. Rothschild, MD. Their time, advice, and support throughout this process were invaluable.

I never would have made it to this point without the mentoring and life coaching of my former colleagues Jean Baum and Joanne Severe. Their personal and professional accomplishments continually inspire me. Their sage wisdom is always appreciated, especially when delivered over a glass of wine.

I am indebted to the STAR*D participants and staff for their tremendous contribution to advancing the treatment of major depression and without whom this dissertation would not have been possible.

I have also benefitted tremendously from the support of the faculty and staff of the Clinical and Population Health Research Program/Department of Quantitative Health Sciences, especially: Kelley Baron; Terry Field, DSc; Stephenie Lemon, PhD; Sandy Stankus; and Mayra S. Tisminetzky, MD, PhD. Many thanks are also owed to all of my fellow CPHR students but especially to my cohort members for making sure I did not

quit at any point in the last four years: Carol Curtin, MSW (very soon to be PhD); Dan Frendl (soon to be MD/PhD); Camilla Pimentel, PhD; and Mollie Wood, PhD.

Last but definitely not least, I am beyond grateful to my friends and family, especially: Karl Anderson, Billie Giannone, Christopher Richards, Kate Ryan, and Kasim Te. Thank you so much for your unconditional support and for making me laugh for all of these years.

ABSTRACT

Background: Major depression is one of the most prevalent, disabling, and costly illnesses worldwide. Despite a 400% increase in antidepressant medication use since 1988, fewer than half of treated depression patients experience a clinically meaningful reduction in symptoms and uncertainty exists regarding how to successfully obtain symptom remission. Identifying homogenous subgroups based on clinically observable characteristics could improve the ability to efficiently predict who will benefit from which treatments.

Methods: Latent class analysis and latent transition analysis (LTA) were applied to data from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study to explore how to efficiently identify subgroups comprised of the multiple dimensions of depression and examine changes in subgroup membership during treatment. The specific aims of this dissertation were to: 1) evaluate latent depression subgroups for men and women prior to antidepressant treatment; 2) examine transitions in these subgroups over 12 weeks of citalopram treatment; and 3) examine differences in functional impairment between women's depression subgroups throughout treatment.

Results: Four subgroups of depression were identified for men and women throughout this work. Men's subgroups were distinguished by depression severity and psychomotor agitation and retardation. Severity, appetite changes, insomnia, and psychomotor disturbances characterized women's subgroups. Psychiatric comorbidities, especially anxiety disorders, were related to increased odds of membership in baseline moderate and severe depression subgroups for men and women. After 12 weeks of citalopram

treatment, depression severity and psychomotor agitation were related to men's chances of improving. Severity and appetite changes were related to women's likelihood of improving during treatment. When functional impairment was incorporated in LTA models for women, baseline functional impairment levels were related to both depression subgroups at baseline and chances of moving to a different depression subgroup after treatment.

Conclusion: Depression severity, psychomotor disturbances, appetite changes, and insomnia distinguished depression subgroups in STAR*D. Gender, functional impairment, comorbid psychiatric disorders, and likelihood of transitioning to subgroups characterized by symptom improvement differed between these subgroups. The results of this work highlight how relying solely on summary symptom rating scale scores during treatment obscures changes in depression that might be informative for improving treatment response.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

GAD – Generalized anxiety disorder

HRSD – Hamilton Rating Scale of Depression

LCA – Latent class analysis

LTA – Latent transition analysis

NIMH – National Institute of Mental Health

OCD – Obsessive compulsive disorder

PDSQ – Psychiatric Diagnostic Screening Questioannaire

PTSD – Post-traumatic stress disorder

QIDS-SR_{16} – self-report version of the Quick Inventory of Depressive Symptomatology

STAR*D – Sequenced Treatment Alternatives to Relieve Depression

WSAS – Work and Social Adjustment Scale

PREFACE

Chapter II of this dissertation is under preparation as:

Ulbricht CM, Rothschild AJ, Lapane KL. Subgroups of Depression in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Study: A Latent Class Analysis.

Chapter III of this dissertation is under preparation as:

Ulbricht CM, Rothschild AJ, Dumenci L, Lapane KL. Changes in Depression Subtypes for Men in STAR*D: A Latent Transition Analysis.

Ulbricht CM, Rothschild AJ, Dumenci L, Lapane KL. Changes in Depression Subtypes for Women in STAR*D: A Latent Transition Analysis.

Chapter IV of this dissertation is under preparation as:

Ulbricht CM, Rothschild AJ, Lapane KL. Functional Impairment and Changes in Depression Subtypes for Women In STAR*D: A Latent Transition Analysis.

CHAPTER I

INTRODUCTION

Major Depression

Major depression is one of the most prevalent and burdensome diseases worldwide,[1,2] conferring substantial disability, morbidity, and mortality. The 12-month prevalence of major depression in the United States has been estimated to be 7%[2] and lifetime prevalence of major depression has been estimated to range from 15-17%.[3] In 2010, more than 15 million adults in the U.S. had experienced major depression in the past year.[4] The incremental burden of individuals with major depression in the U.S. alone was $210.5 billion in 2010.[4] This includes treatment costs, losses from absenteeism and lack of productivity, and lifetime earnings lost due to suicide. Depression is a leading cause of disease burden throughout the world and is anticipated to be the leading cause of disability by 2030.[5] 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. Beyond suicide deaths, depression is also associated with increased mortality from related chronic comorbid medical conditions.

Given the burden of major depression, effective treatment is necessary but there is limited information on how to best treat people so that symptom remission and improved functioning are achieved.[6] Despite a 400% increase in the use of antidepressant medication between 1988-1994 and 2005-2008,[7] only 51.7% people with depression receive any treatment.[8] Of the people who do receive treatment, fewer than half experience a clinically meaningful reduction in symptoms.[6,9] Providing effective treatment in a timely manner after treatment initiation is critical since patients tend to not return for treatment if response to initial treatment is poor. Half of people who start

antidepressant treatment do not receive follow-up care.[10] This occurs despite the high likelihood that a good response may eventually be achieved by switching treatments.[11]

The lack of treatment response in depression might be partially explained by the non-specific symptomatology and variability in severity and trajectory of the disease. More than 1,400 combinations of DSM criteria symptoms are possible[12] and considerable differences in illness course, prominent symptoms, and treatment response have been observed.[13] Depression presents differently by a number of factors.[13–16] In particular, gender differences have been documented in depression rates, severity, course, risk factors, and symptoms, with women experiencing depression more often and more severely than men.[17–20] Women also seem to be more likely to experience somatic, atypical, and anxiety symptoms.[21–23] Additionally, more women than men experience a major depressive episode with severe functional impairment.[17] When men with depression are affected by functional impairment, the domains in which functioning is impaired also appear to differ by gender.[24]

Although the presence of heterogeneity is well-established in depression, how to best delineate subgroups of people who share similar features is still debatable.[25] Depression subtypes based on symptom patterns have been proposed but the clinical utility of these categorizations is unclear. When participants in the International Study to Predict Optimized Treatment in Depression (iSPOT-D) and Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trials were classified post-hoc into the classic melancholic, anxious, and atypical subtypes, a quarter of iSPOT-D and a third of STAR*D participants did not belong to any subtype.[26] Among those who could be

classified, 48% in iSPOT-D and 39% in STAR*D met criteria for more than one subtype. Furthermore, the extent to which these subtypes or even individual symptoms change over time is uncertain. Identifying distinct homogenous subgroups based on clinically observable characteristics could eventually improve the ability to predict who will benefit from which treatments.[11,27,28] It has also been seen that individual symptoms[29,30] and subtypes[31] can change throughout depressive episodes but few studies have been conducted and treatment has rarely been considered. Research about the longitudinal stability of subtypes, including transitions between subtypes during treatment, could ultimately inform efforts to address depression heterogeneity in personalizing treatment strategies, a goal of precision medicine and the National Institute of Mental Health's Research Domain Criteria initiative.[32]

Specific Aims

This dissertation explored using latent class analysis (LCA) and latent transition analysis (LTA) to efficiently identify subgroups comprised of the multiple dimensions of depression and examine changes in subgroup membership during the first level of the STAR*D study, during which participants received the antidepressant citalopram. The specific aims of this dissertation were as follows.

Aim 1. Evaluate baseline latent depression classes:

- Examine underlying depression classes based on clusters of depression symptoms.
- Evaluate gender differences in these classes.
- Identify baseline correlates of class membership.

Aim 2. Examine transitions in latent depression classes over 12 weeks of treatment with citalopram:

- Characterize patterns of depression symptoms during 12 weeks of treatment.
- Examine changes in the descriptive nature of each subgroup and how participants move between subgroups.
- Examine gender differences in these subgroups and transitions.

Aim 3. Examine the association between functional impairment and latent depression statuses in women:

- Characterize the association between functional impairment and major depression subgroups at baseline.
- Characterize changes in depression subgroups by level of functional impairment at the end of 12 weeks of citalopram treatment.

Data Source and Study Population

We used a publicly available de-identified dataset from STAR*D, the largest and longest community-based depression treatment trial conducted to date, for this dissertation. STAR*D was a pragmatic clinical trial originally designed to assess the effectiveness of a variety of treatments for moderate-to-severe nonpsychotic depression.[22] In order to best capture real-world clinical practice, a variety of clinical settings throughout the U.S. were included and broad eligibility criteria were employed. Adults with major depression who were seeking treatment for their depression were eligible to

participate and 4,041 adults were enrolled from 18 primary care and 23 outpatient psychiatric sites between July 2001 and April 2004.

In level 1 of STAR*D, all participants received citalopram for 12 to 14 weeks. Study visits were recommended at weeks 0, 2, 4, 6, 9, and 12, with an additional visit at week 14 for participants who experienced response or remission only at week 12. The main outcome measure was remission of depression symptoms and was defined as a score at study exit on the 17-item Hamilton Rating Scale of Depression (HRSD) of ≤ 7 or last observed 16-item Quick Inventory of Depressive Symptomatology (QIDS-SR_{16}) score of ≤ 5.[33] The secondary outcome was response, defined as at least 50% reduction in QIDS-SR_{16} baseline score. Those who achieved symptom remission could continue on citalopram for up to 12 months of follow-up while those with partial or no response could continue to subsequent randomized treatment levels. Approximately 28% of participants achieved remission as defined by the HRSD score and 47% of participants achieved response.[33]

In level 1, all participants received citalopram. In level 2, those who did not respond to citalopram in level 1 had the option of augmenting citalopram or to be randomly assigned to switch to sertraline, bupropion-SR, or venlafaxine-XR. Psychotherapy was also a treatment option. If remission was not achieved during this phase, participants could continue on to level 3 where they could choose to be randomly assigned to mirtazapine or nortriptyline. In level 4, participants who did not become symptom-free in level 3 could switch to be randomly assigned to receive tranylcypromine

or venlafaxine-XR with mirtazpine. Approximately half of the participants achieved remission after two levels of treatment.[9]

Data from level 1 was used for this dissertation. The original STAR*D investigators defined an evaluable sample of study participants as the 2,876 participants who had an HRSD score of at least 14 at baseline and had at least one post-baseline visit.[33] This evaluable sample served as the basis of the sample eligible for the analyses of this dissertation. For the LCA at baseline in Aim 1, men and women were included if they were not missing all QID-SR_{16} items at baseline and were not missing baseline covariates of age, race/ethnicity, and psychiatry comorbidity. This resulted in a sample of 2,772 participants. In Aim 2, only complete cases were used for the LTA model of baseline and week 12 visits. Participants must have completed study visits at baseline and week 12 and could not be missing all QIDS-SR_{16} items at these visits. The sample was thus comprised of 1,142 participants. The LTA models for Aim 3 included only women who had completed the QIDS-SR_{16} and the Work and Social Adjustment Scale at baseline and the QIDS-SR_{16} at week 12, resulting in a sample of 755 women.

Analytic Methods

Latent Class Analysis

Aim 1 was accomplished by conducting an LCA to examine subgroups of depression at baseline in STAR*D. LCA models are finite mixture models that assume that there are mutually exclusive and exhaustive classes of individuals within a population that can be distinguished by values of an unobserved categorical variable.[34] This unobserved latent variable and resulting subgroups are derived from indicator

variables of observable phenomenon such as depression symptoms. The observed indicator variables are considered to be a function of the latent variable and error. Unlike traditional variable-centered methods of subgroup analysis, LCA accommodates a large amount of information to organize people into homogenous subgroups and thus has the potential to efficiently identify subgroups comprised of the many dimensions of major depression.[35] In Aim 1, the latent depression variable was based on patterns of observed depression symptoms and known correlates of depression course were examined for each latent class in LCA models with covariates.

LCA estimates two sets of parameters: latent class membership prevalences (γ) and item-response probabilities (ρ). Individuals have a probability of membership in each of the latent classes but true class membership is unknown and inferred from response patterns of the observed depression symptoms.[34] The latent class membership prevalence is the probability of membership in that particular class. The item-response probabilities represent the relationship between each indicator variable and each latent class by identifying the response patterns of the observed characteristics that define each latent class.[34] Item-response probabilities approaching 0 or 1 indicate a strong relationship between the observed indicator variable and the latent variable, meaning that the particular response can be determined with a high degree of certainty conditional on latent class membership.

Figure 1.1 shows the overview of the LCA model that was used in Aim 1.[36] The latent depression variable was based on the DSM-IV symptoms of major depression as measured by the QIDS-SR_{16}. Gender was also considered as a grouping variable (model

not shown). Covariates of depression included age, race, and presence of comorbid psychiatric disorders such as generalized anxiety disorder and post-traumatic stress disorder. The overall model is:

$P(\mathbf{Y} = \mathbf{y}|\mathrm{X} = x) = \sum_{c=1}^{C} \gamma_c\,(x) \prod_{j=1}^{J} \prod_{r_j=1}^{R_j} \rho_{j,r_{j|c}}^{I(y_j=r_j)}$, where y_j is item j of the set of indicator variables y (e.g., early morning insomnia, sad mood, impaired concentration, etc.); r_j is the response to item j; X is a covariate (e.g., age); and c is the latent depression subgroup. $\gamma_c(x)$ is a logistic regression model of the probability of membership in latent depression class c =

$$P(L = c|\mathrm{X} = x) = \frac{e^{\beta_{0c} + \beta_{1c}Age + \beta_{2c}Race + \beta_{3c}PsychiatricComorbidity}}{1+\sum_{c=1}^{C-1} e^{\beta_{0c} + \beta_{1c}Age + \beta_{2c}Race + \beta_{3c}PsychiatricComorbidity}}.$$ [34,35]

Basic models without covariates or a grouping variable were first fit in order to gain an understanding of the overall latent structure and the appropriate number of classes. These models were fit with numbers of classes that varied from two to seven. The G^2 likelihood-ratio statistic, Akaike Information Criterion, Bayesian Information Criterion, solution stability, and interpretability of competing solutions were considered when selecting the model with the optimal number of latent classes.[34] After this basic model was determined, gender was considered as a grouping variable. A multiple-group LCA permits the exploration of measurement invariance to see if the latent construct is the same across the groups. This approach allows for gender-specific estimation of latent class membership probabilities, item-response probabilities, and predictors of latent class membership. By considering gender as a grouping variable, it was possible to evaluate

two areas: 1) if men and women experience the same depression subgroups; and 2) if so, if the prevalence of each subgroup is similar in men and women.

Measurement invariance between men and women was tested to see if the latent construct of depression was the same for men and women. If there is measurement invariance in the LCA, the number of classes and item-response probabilities will be the same across groups, indicating that the qualitative understanding of the depression symptoms characterizing the latent classes is the same between groups.[34] Measurement invariance was examined with the G^2 likelihood-ratio difference test to determine if the item-response probabilities differ in two nested models of the dichotomous QIDS-SR_{16} indicators: a multiple-group model with all the prevalence and item-response probability parameters free to vary by gender and a multiple-group model with the item-response probabilities constrained to be equal between men and women.[34] Beyond the results of the difference G^2 test, we considered the principle of parsimony and the interpretability of the latent classes, especially in terms of clinically meaningful differences in the item-response probabilities.

After measurement invariance of the item-response probabilities was tested, the equivalence of the latent class prevalences across men and women was examined with the difference G^2 test in a similar manner as measurement invariance. Two nested models were compared: a model where women's and men's item-response probabilities were constrained to be equal but latent class prevalences were allowed to vary and a model in which the item-response probabilities and latent class prevalences were constrained to be equal across the genders.[34] After the equivalence of the latent class prevalences was

tested, covariates were added to the multiple-groups model using multinomial logistic regression. When covariates are added to the LCA model, the resulting odds ratios are the difference in odds of membership in the reference class compared to the other classes for individuals endorsing a particular covariate. Covariates were initially added individually to the LCA model with gender as a grouping variable and then a model adjusted for all the covariates was fit. An LCA model with covariates was used instead of the "classify-analyze" approach of using posterior probabilities to assign individuals to their most likely latent class in order to traditionally analyze characteristics of the classes with descriptive statistics. This was done because the classify-analyze approach does not accommodate the uncertainty in classification in LCA and can thus introduce classification error.[35]

Latent Transition Analysis

Aims 2 and 3 employed LTA to examine changes in the latent depression classes over the 12 weeks of STAR*D level 1. LTA is the longitudinal extension of LCA. In LTA, "statuses" are analogous to "classes" in LCA. This distinction emphasizes that the statuses can be temporary states and people's membership in these statuses can change.[34] As in LCA, LTA models latent status prevalences and item-response probabilities. Unlike LCA, LTA also estimates transition probabilities. These transition probabilities convey how changes occur between the statuses over time, demonstrating the incidence of transitioning to a latent status conditional on membership in an earlier status.

In Aims 2 and 3, LTA models were fit to determine the probabilities for individual participants belonging to a particular depression subgroup at week 12 given

the individual's depression subgroup membership at baseline (Figure 1.2).[37] The basic LTA model for latent depression statuses at baseline and week 12 is:

$$P(\mathbf{Y} = \mathbf{y}) = \sum_{s_1=1}^{S} \sum_{s_2=1}^{S} \delta_{s_1} \tau_{s_2|s_1} \prod_{t=1}^{2} \prod_{j=1}^{J} \prod_{r_{j,t}=1}^{R_j} \rho_{j,r_{j,t}|s_t}^{I(y_{j,t}=r_{j,t})},$$

where δ_{s1} is the probability of membership in latent depression status s_1 at baseline; $\tau_{s2|s1}$ is the probability of transitioning to latent status s_2 at week 12 conditional on membership in latent status s_1 at baseline; and $\rho_{t,j,r_{jt|s_t}}$ is the probability of response $r_{j,t}$ to item j at time t, conditional on membership in latent class s_t at time t.[34] In LTA, measurement invariance can exist across groups and/or across times. As was done in the LCA for Aim 1, measurement invariance was tested in nested models with the G^2 difference likelihood ratio test. In Aim 2, measurement invariance between men and women and between time points was examined. In Aim 3, measurement invariance was tested between women's functional impairment groups and between time points.

Figure 1.1: Overview of the latent class analysis model to examine baseline subtypes of depression in STAR*D level 1

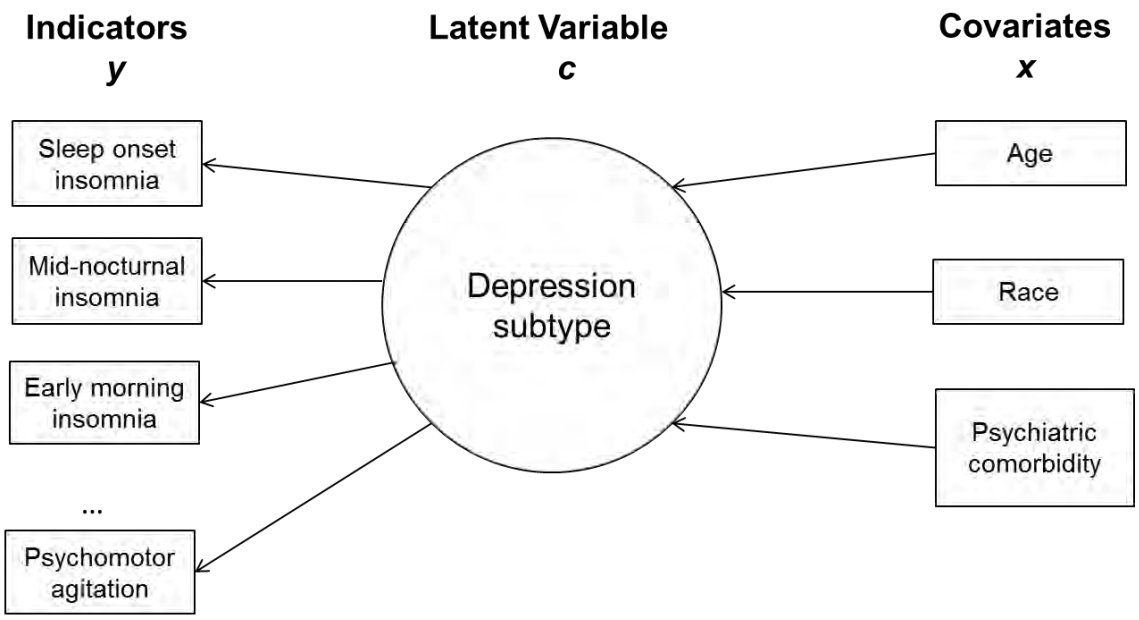
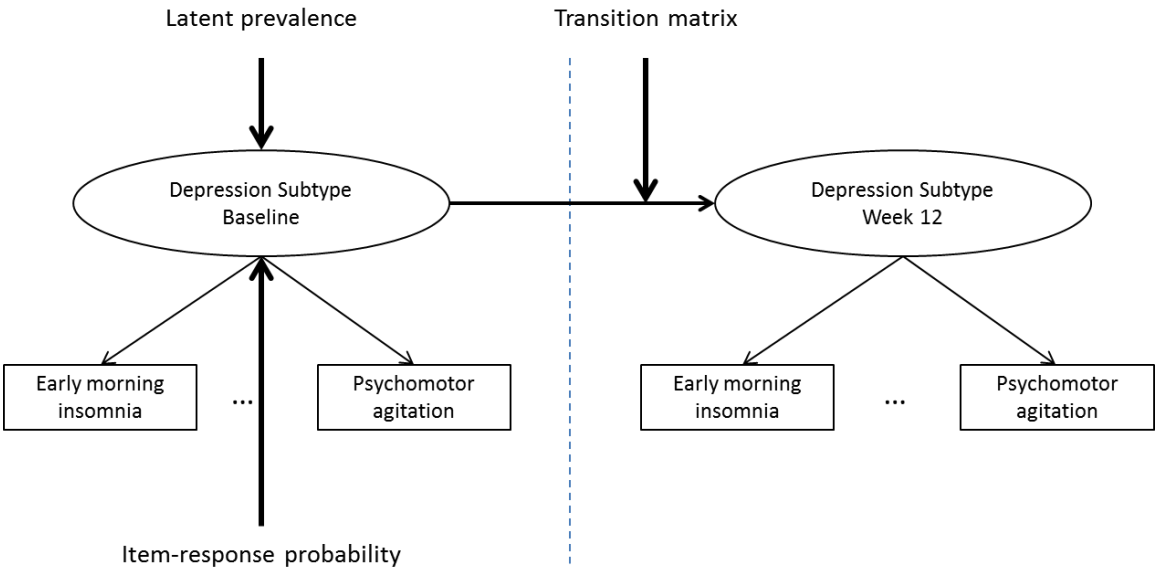


Figure 1.2: Overview of the latent transition analysis model to examine changes in depression subtype in STAR*D level 1



CHAPTER II

SUBGROUPS OF DEPRESSION IN THE SEQUENCED TREATMENT ALTERNATIVES TO RELIEVE DEPRESSION (STAR*D) STUDY - A LATENT CLASS ANALYSIS

Abstract

Objective: The objective of this study was to characterize latent classes of depression by symptoms, evaluate gender differences in these classes, and examine correlates of each class.

Method: Latent class analysis was applied to baseline data from 2,772 participants in level 1 of the Sequenced Treatment Alternatives to Relieve Depression trial. Items from the Quick Inventory of Depressive Symptomatology were used as indicators. Multinomial logistic models identified correlates of latent classes (adjusted odds ratios (aOR) and 95% confidence intervals (CI)).

Results: Four latent classes were identified: Mild (men: 37%, women: 27%), Moderate (men: 24%, women 21%), Severe with Increased Appetite (men: 13%, women: 22%), and Severe with Insomnia (men: 26%, women: 31%). Comorbid generalized anxiety disorder (aOR_{women}: 1.74; 95% CI: 1.06-2.85), bulimia (aOR_{women}: 5.21; 95% CI: 3.16-8.59; aOR_{men}: 12.29; 95% CI: 5.28-28.6), and social phobia (aOR_{women}: 3.68; 95% CI: 2.36-5.75; aOR_{men}: 3.22; 95% CI: 1.71-6.06) were correlated with Severe with Increased Appetite. Generalized anxiety disorder (aOR_{women}: 2.91; 95% CI: 1.91-4.41; aOR_{men}: 2.07; 95% CI: 1.24-3.46), post-traumatic stress disorder (aOR_{women}: 2.30; 95% CI: 1.51-3.51; aOR_{men}: 2.00; 95% CI: 1.24-3.24), and social phobia (aOR_{women}: 2.41; 95% CI: 1.60-3.63) were correlated with Severe with Insomnia.

Conclusions: Insomnia and increased appetite distinguished latent classes. Gender, anxiety disorders and other psychiatric comorbidities also differed between these classes.

These results suggest that sleep disturbances, appetite changes, and other mental disorders may play a role in the etiology of depression.

Introduction

Major depression is one of the most prevalent, disabling, and costly illnesses worldwide.[1,2] Despite a 400% increase in antidepressant medication use since 1988, the prevalence of depression remains around 7% for adults in the United States.[7,17,38] Fewer than half of treated depression patients experience a clinically meaningful reduction in symptoms and uncertainty exists regarding how to successfully obtain symptom remission.[6,9] Understanding this heterogeneity is necessary to identify predictors of response and ultimately improve the treatments and services available for depression.[32]

How to delineate subgroups of people who respond differently to antidepressants is debatable.[25] The heterogeneity is likely partially explained by the non-specific symptomatology and variability in severity and trajectory of depression. Depression presents differently by age, gender, race and ethnicity, and psychiatric comorbidities.[13,15–17] Numerous depression subtypes have been proposed but lack of clinical utility limits adoption.[39] Identifying homogenous subgroups based on clinically observable characteristics could improve the ability to efficiently predict who will benefit from which treatments.[11,27,28]

Latent class analysis (LCA) is a person-centered analytic approach which can be used to efficiently identify subgroups comprised of the multiple dimensions of depression.[35] LCA models assume mutually exclusive and exhaustive classes of individuals within a population differentiated by values of an unobserved categorical latent variable.[34] This latent variable and resulting classes are based on observed indicator

variables, such as depression symptoms. Individuals have a probability of membership in each of the latent classes, inferred from response patterns of indicator variables.[34]

We sought to evaluate the extent to which latent classes based on depression symptoms could be identified using LCA. Data were used from level 1 of the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial, the largest community-based study of major depression.[33] STAR*D provides a unique opportunity to study patients who would have been excluded from most other studies.[40,41] The objectives were to: 1) examine underlying depression classes based on patterns of depression symptoms; 2) evaluate gender differences in latent depression classes; and 3) identify correlates of the depression classes.

Methods

Study Participants

We used a de-identified dataset from STAR*D, a pragmatic clinical trial to assess the effectiveness of a variety of treatments for moderate-to-severe nonpsychotic depression.[42] From July 2001-April 2004, 4,041 treatment-seeking patients were enrolled from 18 primary care and 23 outpatient psychiatric sites.[43,44] The open-label, unblinded treatment protocol for level 1 allowed for flexible dosing of citalopram for 14 weeks.[33] Approximately 30% of participants achieved remission (primary endpoint: 17-item Hamilton Rating Scale of Depression (HRSD) of ≤ 7 or last observed 16-item Quick Inventory of Depressive Symptomatology ($QIDS\text{-}SR_{16}$) score of ≤ 5) and 47% achieved response (secondary outcome: $\geq 50\%$ reduction in $QIDS\text{-}SR_{16}$ baseline score) by the end of level 1.[33]

LCA employs full information maximum likelihood for missing data on individual indicator variables such as the QIDS-SR_{16} items but requires complete case analysis for missing covariates. Of the evaluable sample (participants with HRSD ≥ 14 at baseline and ≥ one post-baseline visit, n= 2,876),[45] 15 participants were excluded because they were missing all QIDS-SR_{16} indicator items at baseline. An additional 89 participants were excluded because they were missing the covariates of interest, resulting in a sample of 2,772 participants for this analysis. The University of Massachusetts Medical School Institutional Review Board determined that this secondary analysis was not human subject research.

Measures

Indicators of Latent Class Membership

The 16 baseline QIDS-SR_{16} items were the indicator variables from which the latent class construct was inferred: sad mood, impaired concentration, self-criticism, suicidal ideation, lack of general interest, fatigue, sleep disturbances, appetite and weight changes, and psychomotor agitation/retardation.[46] Items pertain to experiences in the past seven days, except for weight change (previous two weeks). Item scores ranged from 0-3, with scores ≥ 2 indicating the symptom met the DSM-IV depression threshold.[44] All items were dichotomized (≤ 1: absence; ≥2: presence of a symptom). Although the QIDS-SR_{16} instructions specify that only one item on decreased or increased appetite should be completed, these items were included as separate indicators to capture the direction of change. Weight changes were also treated this way.

Consideration of Gender in the LCA

Gender differences occur in depression rates, severity, course, risk factors, and symptoms, with women experiencing depression more often and more severely than men.[17–20] We did not *a priori* assume that men and women would have the same subtypes of depression. Instead, we evaluated: 1) if men and women experienced the same types of depression; and 2) gender differences in class prevalences.

Correlates of Depression Class Membership

We considered sociodemographic and clinical variables as correlates of depression class membership. Baseline age (<45 years versus ≥ 45 years), race, and psychiatric comorbidities were considered because these are previously identified demographic and clinical predictors of depression subtypes.[47] Comorbid DSM-IV conditions were assessed at baseline with the Psychiatric Diagnostic Screening Questionnaire (PDSQ),[48] which screens for 13 mood, anxiety, eating, substance use, and somatoform disorders.[49] A threshold of 90% specificity was used to determine the presence of a disorder.[50] We examined comorbid post-traumatic stress disorder (PTSD), social phobia, generalized anxiety disorder (GAD), and bulimia because the prevalence of each of these was ≥ 10%. We created one variable that defined other psychiatric comorbidity as conditions occurring with low frequency, including obsessive-compulsive disorder (OCD), panic disorder, psychosis, agoraphobia, alcohol abuse/dependence, drug abuse/dependence, somatization, and/or hypochondriasis.

Statistical Analysis

There were three analytic phases: 1) describing the sample and developing the LCA model; 2) evaluating gender differences in depression classes; and 3) evaluating class correlates. First, summary statistics of gender differences in baseline demographic and clinical characteristics were calculated using t-tests and chi-square tests as appropriate. We then constructed LCA models with numbers of classes varying from one to seven. Basic models without grouping variables and covariates were fit first to obtain a general understanding of the structure of the classes. These models were examined for fit, parsimony, interpretability, and sparseness of the indicator variables.[34] Fit was assessed with Akaike information criterion (AIC)[51] and Bayesian information criterion (BIC)[52] information criteria statistics.

Second, after the optimal number of latent classes was selected for the basic model, models with gender as a grouping variable were examined to see if the number of latent classes was identical between men and women. We explored measurement invariance using the difference G^2 likelihood ratio test to see if the item-response probabilities, and thus the qualitative nature of the latent construct, were the same by gender.[34] This test was used to determine if the item-response probabilities differed in two nested models of the QIDS-SR_{16} indicators: 1) a multiple-group model with all the prevalence and item-response probability parameters free to vary by gender; and 2) a multiple-group model with the item-response probabilities constrained to be equal between men and women.[34] We considered parsimony and the interpretability of the

latent classes, especially in terms of clinically meaningful differences in the item-response probabilities.

After measurement invariance in the item-response probabilities was confirmed, we used the difference G^2 test to examine the equivalence of latent class prevalences across men and women. Two nested models were compared: 1) a model where women's and men's item-response probabilities were constrained to be equal but latent class prevalences were allowed to vary; and 2) a model in which the item-response probabilities and latent class prevalences for men and women were constrained to be equal.[34]

The third phase evaluated correlates of the depression classes. Covariates were added to the four-class multiple-groups LCA in multinomial logistic models. Mild Depression served as the referent group to facilitate interpretation. Odds ratios of the difference in odds of membership in the Mild referent class compared to another class for individuals endorsing a particular covariate were derived from the multinomial logistic models. Unadjusted gender-specific odds ratios were estimated for each covariate and then the model was adjusted for all covariates. Sparseness was examined and ruled out.[34]

Analyses were conducted using PROC LCA[53,54] in SAS 9.3 and Mplus version 7.2.[55]

Results

Women and men differed on several demographic and clinical characteristics (Table 2.1). Women were younger than men. Men were more likely than women to be white. Women had a more severe depression at baseline but had experienced fewer

depressive episodes, on average, than men. Of the individual QIDS-SR items, women were more likely to report sad mood, appetite changes, weight gain, impaired concentration, negative self-view, fatigue, and psychomotor retardation. Men were more likely than women to report suicidal ideation and psychomotor agitation.

LCA Model

A four-class base model was selected as optimal after examining fit statistics, latent class prevalences, and interpretability (Supplementary Table 2.1). Although the traditional fit statistics of AIC and BIC suggested the five-class model, the four-class model was ultimately selected since the additional fifth class was not readily distinguished from the others and had a very low prevalence. Sparseness did not appear to be an issue since no latent class prevalence estimate approached zero. All four classes included participants likely to endorse sad mood but unlikely to endorse hypersomnia, decreased weight, suicidal ideation, or psychomotor disturbances (Supplementary Table 2.2). Two classes had high probabilities of endorsing most symptoms and thus seemed to represent more severe depression than the other classes. The Severe Depression with Insomnia class (31% of participants) had the highest probabilities of reporting sleep-onset insomnia, mid-nocturnal insomnia, and early morning insomnia. The Severe Depression with Increased Appetite (~16%) included those with a high probability of having increased appetite. The Moderate Depression class (23%) had many of the same symptoms as the severe classes but lower probabilities of endorsing these symptoms. The Mild Depression class (30%) was the least likely of all classes to have sad mood; members were unlikely to have early morning insomnia, hypersomnia, impaired

concentration, negative self-view, lack of general interest, fatigue, appetite or weight changes, and psychomotor disturbances.

Gender Differences in Class Membership

When gender was added as a grouping variable, a four-class solution was optimal for both genders (Supplementary Table 2.1). The difference G^2 tests indicated that measurement invariance did not hold and the model with all parameters freed to vary by gender fit best. However, because only the class prevalences and not the item-response probabilities varied by gender, the data were best described with the item-response probabilities constrained to be equal by gender (Table 2.2). According to tests of equivalence of latent class prevalences, the proportions of women and men in each class differed for the Severe with Increased Appetite ($p < 0.01$), Severe with Insomnia ($p = 0.02$) and Mild ($p < 0.01$) classes but not the Moderate class ($p = 0.20$). The Mild class was the most prevalent for men (37%) but the Severe with Insomnia class was most prevalent for women (31%). Almost twice as many women (22%) as men (13%) were in the Severe with Increased Appetite class.

Demographic and Clinical Correlates of Class Membership

Several demographic and clinical factors were correlates of class membership (Table 2.3). The Mild class served as the reference group in all models. Men in the Moderate class were less likely than those in the Mild class to be 45 years or older (aOR: 0.41; 95% CI: 0.24-0.72) and to have comorbid PTSD (aOR: 0.28; 95% CI: 0.09-0.88) but more likely to have comorbid social phobia (aOR: 2.05; 95% CI: 1.12-3.76). Those in the Severe with Increased Appetite class were more likely to have comorbid bulimia

(aOR: 12.29; 95% CI: 5.28-28.6) and social phobia (aOR: 3.22; 95% CI: 1.71-6.06) compared to men in the Mild class. Men in the Severe with Insomnia class were less likely to be 45 years or older (aOR: 0.65; 95% CI: 0.43-0.99) and more likely to have comorbid GAD (aOR: 2.07; 95% CI: 1.24-3.46), PTSD (aOR: 2.00; 95% CI: 1.24-3.24), and any other psychiatric comorbidity (aOR: 1.75; 95% CI: 1.14-2.68).

Women in the Moderate class were less likely to be 45 years or older (aOR: 0.23; 95% CI: 0.12-0.43) and to have any other psychiatric comorbidity (aOR: 0.40; 95% CI: 0.21-0.74). Women in the Moderate class had greater odds of being white (aOR: 2.35; 95% CI: 1.27-4.34) and of having comorbid social phobia (aOR: 3.12; 95% CI: 1.82-5.36). Women in the Severe with Increased Appetite class were more likely than those in the Mild class to have comorbid GAD (aOR: 1.74; 95% CI: 1.06-2.85), bulimia (aOR: 5.21; 95% CI: 3.16-8.59), and social phobia (aOR: 3.68; 95% CI: 2.36-5.75). Those in the Severe with Insomnia class had lower odds of being white compared to women in the Mild class (aOR: 0.64; 95% CI: 0.44-0.93). Those in the Severe with Insomnia class had greater odds of having comorbid GAD (aOR: 2.91; 95% CI: 1.91-4.41), PTSD (aOR: 2.30; 95% CI: 1.51-3.51), social phobia (aOR: 2.41; 95% CI: 1.60-3.63), and any other psychiatric comorbidity (aOR: 1.69; 95% CI: 1.19-2.41).

Discussion

This study sought to improve the understanding of depression subtypes based on clinically meaningful symptom patternsby utilizing the person-centered analytic approach of LCA. Four classes of depression were identified: Mild, Moderate, Severe with Increased Appetite, and Severe with Insomnia. Men and women experienced the same

patterns of depression symptoms but the proportions of men and women likely to be in the Mild, Severe with Increased Appetite, and Severe with Insomnia classes differed. Age, race, and psychiatric comorbidities were correlates of depression class membership.

While STAR*D participants were required to have at least "moderate" depression as determined by an HRSD score ≥ 14,[33,56] the Mild class had a relatively high probability of endorsing sad mood but low probabilities of endorsing all other DSM-IV depression symptom criteria. This class was the most common for men. The Moderate class was distinguished by probabilities of experiencing sad mood, impaired concentration, lack of general interest, and low energy that were higher than in the Mild class but lower than in the two Severe classes. Both Severe classes had very high probabilities of endorsing most of the depression criteria.

The Severe with Insomnia class was the most prevalent class for women but the second most prevalent for men. The salience of insomnia symptoms in distinguishing this class is in line with how sleep disturbances are commonly both precursors[57,58] and symptoms of depression.[59] The difference in the latent class prevalences for women and men is consistent with women having a higher risk of insomnia than men.[60] Consistent with previous reports linking sleep problems and anxiety,[61,62] both men and women with anxiety disorders were more likely to be in this class than in the Mild class. Men and women 45 years or older were less likely to belong to the Severe with Insomnia class than to the Mild class, although for women this finding was attenuated after adjustment of other factors. This finding is contrary to reports that older age is associated with decreased sleep quality[63] but differences between this sample and those of previous

studies, especially related to comorbid psychiatric conditions, could explain this contradiction. Addressing the insomnia symptoms for people in this class could be beneficial in reducing overall depression[64] because insomnia is associated with depression recurrence[57,65] and lack of antidepressant treatment response.[66]

The gender difference in the Severe with Increased Appetite class prevalence was striking, with almost twice as many women in this class. Participants in this class also had substantially increased odds of having comorbid bulimia. The presence of this class supports findings that depression is bi-directionally associated with obesity,[67,68] insulin resistance,[69,70]and diabetes.[71,72] The lack of concurrently increased weight in this class might be due to the limited timeframe addressed by the QIDS-SR_{16} or comorbid bulimia and related dieting and compensating for overeating. Numerous pathophysiological mechanisms for the relationship between depression and appetite disturbances have been proposed, including hypothalamic-pituitary-adrenal (HPA) axis dysregulation,[73] appetite-stimulating effects of psychotropic medication,[74] behaviors such as increased inflammation.[75] The complex involvement of HPA axis dysregulation, cortisol, and brain-derived neurotrophic factor in bulimia[76] and depression reinforces these areas as potential targets for treatment development.

Depression subtypes of varying symptom patterns and those existing along a severity gradient have been proposed.[22,77–79] The classes we identified share some similarities but do not correspond completely with traditional subtypes, including the melancholic, anxious and atypical DSM specifiers.[59] For example, hypersomnia was not likely to be endorsed by any class despite being a hallmark feature of atypical depression.

Direct comparisons between classes from this analysis and those described in the DSM are not feasible because the QIDS-SR_{16} does not assess several DSM specifier criteria, such as diurnal variation. Moreover, the value of these specifiers remains to be determined, given persistent controversies over their validity.[80] Numerous factors, including using the QIDS-SR_{16} for indicators and STAR*D's clinical trial setting, must be considered when comparing our findings to others.

Our results should be interpreted with caution. STAR*D data were not originally designed to examine depression subgroups. STAR*D participants had depression, were seeking care, and enrolled in a clinical trial. This might lead to confounding by treatment-seeking status, sampling bias due to the least severely depressed patients not being eligible to participate, and limited generalizability. STAR*D had broader eligibility criteria than most trials and this sample is more likely than other trials to be representative of patients seen in real-world settings.

The measurement of indicators and correlates of class membership might have introduced concerns. This analysis focused on baseline symptoms; how treatment-emergent symptoms might alter the subgroups over time was beyond the scope of this work. The indicators were established by self-report on the QIDS-SR_{16}, which might introduce response bias and lead to incorrect estimates of symptoms. Despite this, the QIDS-SR_{16} measures DSM depression criteria, is widely used in research and clinical care, and has been shown to be strongly correlated with clinician ratings and sensitive to changes in symptoms.[81,82] Furthermore, LCA estimates and adjusts for measurement error.[34]

This analysis is notable in several ways. It is one of the largest to use LCA, which is enhanced by sample sizes > 500,[83] to examine depression subgroups. The models included several high quality indicators with item-response probabilities close to 0 or 1[84] and clinically relevant correlates of depression class membership. Previous work has relied on regression models not well-suited to detecting subgroups. Such approaches to subgroup analysis can be limited by the vulnerability of multiple comparisons to Type I error rates and statistical power varying across the subgroups due to unbalanced sample sizes.[35] These issues limit the possibility of examining higher-order interactions among subgroups.[35]

Our study suggests that sleep disturbances, appetite changes, and psychiatric comorbidities may differentiate subgroups of people with major depression. These latent classes differed by gender in the proportion of men and women belonging to each class. Our study demonstrates the potential of using a person-centered approach to detect subgroups based on not easily observable but clinically important symptom patterns. We demonstrated that LCA can be used to examine heterogeneity and identify subgroups for which treatment strategies can be tailored, a goal of precision medicine[32] and the National Institute of Mental Health's Research Domain Criteria initiative.[85]

Table 2.1: Baseline demographic and clinical characteristics of participants by gender

Characteristic at study entry	Women (n =1,763)	Men (n = 1,009)	*p*-value
Mean age (SD)	40.0 (13.0)	43.0 (12.6)	<0.001
Race, n (%)			
White	1,313 (74.5)	802 (79.5)	0.012
Black or African American	325 (18.4)	149 (14.8)	
Other	125 (7.1)	58 (5.8)	
Hispanic, n (%)	279 (15.8)	85 (8.4)	<0.001
Mean age at onset (SD)	24.3 (14.2)	23.7 (14.3)	<0.001
Mean duration (SD)	15.7 (13.0)	16.4 (13.4)	0.200
Mean number of episodes (SD)	4.6 (7.5)	7.1 (11.4)	<0.001
Mean current depression severity (QIDS-SR_{16}) (SD)	16.5 (16.5)	15.6 (3.9)	<0.001
Psychiatric comorbidities, n (%)			
GAD	467 (26.5)	188 (18.6)	<0.001
PTSD	371 (21.0)	203 (20.1)	0.563
Bulimia	294 (16.7)	68 (6.7)	<0.001
Social phobia	581 (33.0)	288 (28.5)	0.016
Any other[a]	705 (40.0)	463 (45.9)	0.003
QIDS-SR_{16} items, n (%)			
Sleep onset insomnia	1,193 (67.7)	673 (66.7)	0.601
Mid-nocturnal insomnia	1,298 (73.6)	722 (71.6)	0.239
Early morning insomnia	874 (49.6)	493 (48.9)	0.696
Hypersomnia	239 (13.6)	119 (11.8)	0.183
Sad mood	1,528 (86.7)	812 (80.5)	<0.001
Decreased appetite	444 (25.2)	210 (20.8)	0.009
Increased appetite	328 (18.6)	120 (11.9)	<0.001
Decreased weight	242 (13.7)	136 (13.5)	0.855
Increased weight	245 (13.9)	89 (8.8)	<0.001
Impaired concentration	1,121 (63.6)	595 (59.0)	0.018
Negative view of self	976 (55.4)	503 (49.9)	0.006
Suicidal ideation	219 (12.4)	166 (16.5)	0.003
Lack of general interest	1,049 (59.5)	593 (58.8)	0.729
Fatigue	1,327 (75.3)	659 (65.4)	<0.001
Psychomotor slowing	637 (36.2)	310 (30.8)	0.004
Psychomotor agitation	548 (31.1)	351(34.8)	0.043

[a]Any other psychiatric comorbidity includes obsessive compulsive disorder, panic disorder, psychosis, agoraphobia, alcohol abuse/dependence, drug abuse/dependence, somatization, and/or hypochondriasis.

Table 2.2: Latent class prevalences and item-response probabilities of endorsing depression symptoms from a four-class LCA model of baseline QIDS-SR_{16} items by gender

	Mild	**Moderate**	**Severe with Increased Appetite**	**Severe with Insomnia**
Latent Class Prevalence				
Men	37%	24%	13%	26%
Women	27%	21%	22%	31%
Item-Response Probabilities of QIDS-SR_{16} Indicators[1]				
Sleep onset insomnia	0.60	0.51	0.64	0.89
Mid-nocturnal insomnia	0.73	0.58	0.70	0.86
Early morning insomnia	0.45	0.23	0.51	0.72
Hypersomnia	0.07	0.22	0.19	0.08
Sad mood	0.63	0.88	0.95	0.97
Decreased appetite	0.09	0.17	0.00	0.59
Increased appetite	0.08	0.00	0.74	0.00
Decreased weight	0.09	0.04	0.01	0.34
Increased weight	0.07	0.03	0.42	0.03
Impaired concentration	0.25	0.73	0.79	0.82
Negative view of self	0.28	0.56	0.69	0.69
Suicidal ideation	0.07	0.12	0.15	0.22
Lack of general interest	0.27	0.69	0.74	0.77
Fatigue	0.38	0.85	0.87	0.85
Psychomotor retardation	0.12	0.32	0.46	0.32
Psychomotor agitation	0.23	0.22	0.31	0.22

[1]The item-response probabilities for each class were constrained to be equal across gender.

Table 2.3: Associations between baseline demographic and clinical variables and latent class membership

	Moderate		Severe with Increased Appetite		Severe with Insomnia	
	Unadjusted	Adjusted[1]	Unadjusted	Adjusted[1]	Unadjusted	Adjusted[1]
Correlate	**Odds Ratio (95% Confidence Interval)[2]**					
Men						
Age ≥ 45 years	0.32 (0.18-1.47)	0.41 (0.24-0.72)	0.59 (0.35-1.02)	0.66 (0.36-1.21)	0.65 (0.45-0.95)	0.65 (0.43-0.99)
Race						
White vs. Black/Other	2.64 (1.22-5.72)	1.92 (0.87-4.24)	1.91 (0.90-4.04)	1.51 (0.65-3.52)	0.91(0.56-1.49)	1.25 (0.74-2.10)
Psychiatric comorbidities						
GAD	1.04 (0.55-1.98)	1.03 (0.47-2.29)	2.63 (1.29-5.36)	1.33 (0.61-2.92)	3.01 (1.86-4.86)	2.07 (1.24-3.46)
PTSD	0.34 (0.14-0.82)	0.28 (0.09-0.88)	0.90 (0.47-1.72)	0.80 (0.37-1.72)	2.64 (1.66-4.18)	2.00 (1.24-3.24)
Bulimia	1.58 (0.48-5.25)	0.98 (0.22-4.27)	15.43 (6.54-36.40)	12.29 (5.28-28.6)	0.66 (0.17-2.48)	0.45 (0.11-1.85)
Social phobia	1.56 (0.85-2.87)	2.05 (1.12-3.76)	3.74 (2.20-6.37)	3.22 (1.71-6.06)	2.38 (1.54-3.69)	1.47 (0.90-2.38)
Any other[3]	0.78(0.47-1.29)	0.85 (0.50-1.45)	1.70 (0.98-2.95)	1.31 (0.71-2.43)	2.41 (1.62-3.61)	1.75 (1.14-2.68)
Women						
Age ≥ 45 years	0.17 (0.08-0.37)	0.23 (0.12-0.43)	0.69 (0.48-0.98)	0.91 (0.61-1.34)	0.60 (0.44-0.82)	0.79 (0.56-1.12)
Race						
White vs. Black/Other	2.58 (1.25-5.34)	2.35 (1.27-4.34)	1.01 (0.67-1.52)	0.88 (0.56-1.37)	0.48 (0.33-0.71)	0.64 (0.44-0.93)
Psychiatric comorbidities						
GAD	1.04 (0.55-1.98)	1.44 (0.74-2.80)	3.53 (2.18-5.71)	1.74 (1.06-2.85)	4.95 (3.32-7.37)	2.91 (1.91-4.41)
PTSD	0.75 (0.36-1.57)	0.77 (0.38-1.58)	2.04 (1.31-3.19)	1.28 (0.78-2.09)	4.10 (2.71-6.19)	2.30 (1.51-3.51)
Bulimia	1.44 (0.70-2.97)	0.85 (0.38-1.58)	7.93 (4.70-13.40)	5.21 (3.16-8.59)	1.26 (0.74-2.17)	0.82 (0.47-1.41)
Social phobia	2.86 (1.68-4.87)	3.12 (1.82-5.36)	5.49 (3.64-8.28)	3.68 (2.36-5.75)	4.21 (2.86-6.21)	2.41 (1.60-3.63)
Any other[3]	0.48(0.27-0.87)	0.40 (0.21-0.74)	1.94 (1.33-2.84)	1.17 (0.78-1.77)	3.15 (2.25-4.41)	1.69 (1.19-2.41)

[1]Adjusted for age, race, and psychiatric comorbidity.
[2]Reference class = Mild Depression.
[3]Any other psychiatric comorbidity includes obsessive compulsive disorder, panic disorder, psychosis, agoraphobia, alcohol abuse/dependence, drug abuse/dependence, somatization, and/or hypochondriasis.

Supplementary Table 2.1: Fit information for LCA models of baseline depression symptoms without covariates

# of Classes	*df*	AIC	BIC	CAIC	Adjusted BIC	Entropy	G^2	% of seeds associated with best fitted model
Base Model Without Covariates								
2	65502	7720.58	7916.18	7949.18	7811.33	0.60	7654.58	100
3	65485	6918.83	7215.20	7265.20	7056.33	0.64	6818.83	100
4	65468	6767.30	7134.43	7201.43	6921.55	0.62	6603.30	100
5	66045	6622.11	7120.00	7204.00	6853.11	0.70	6454.11	15
6	65434	6548.30	7146.96	7247.96	6826.05	0.65	6346.30	20
7	65417	6494.96	7194.39	7312.39	6819.46	0.66	6258.96	5
Gender as Grouping Variable: Without Measurement Invariance								
2	131005	9828.18	10219.38	10285.38	10009.68	0.60	9696.18	100
3	100971	9039.94	9632.68	9732.68	9314.94	0.64	8839.94	100
4	130937	8881.03	9675.29	9809.29	9249.53	0.63	8613.03	45
5	130903	8768.33	9764.12	9932.12	9230.32	0.69	8432.33	5
6	130869	8701.20	9898.52	10100.52	9256.70	0.68	8297.20	5
7	130835	8656.84	10055.69	10291.69	9305.84	0.65	8184.84	5
Gender as Grouping Variable: With Measurement Invariance								
2	131037	9863.89	10065.42	10099.42	9957.39	0.60	9795.89	100
3	131019	9055.45	9363.67	9415.67	9198.45	0.64	8951.45	100
4	131001	8868.27	9283.18	9353.18	9283.18	0.64	8728.27	30
5	130983	8756.21	8756.21	9277.82	8998.21	0.70	8580.21	10
6	130965	8679.39	9307.69	9413.69	8970.89	0.66	8467.39	30
7	130947	8624.83	9359.82	9483.82	8965.83	0.66	8376.83	5

Supplementary Table 2.2: Latent class prevalences and item-response probabilities of endorsing depression symptoms from a four-class model of baseline QIDS-SR_{16} items

Symptoms	Mild	Moderate	Severe with Increased Appetite	Severe with Insomnia
Latent Class Prevalence				
	30%	23%	16%	31%
Item-Response Probabilities of QIDS-SR_{16} Indicators				
Sleep onset insomnia	0.61	0.48	0.68	0.88
Mid-nocturnal insomnia	0.74	0.56	0.73	0.85
Early morning insomnia	0.46	0.21	0.57	0.71
Hypersomnia	0.07	0.24	0.17	0.09
Sad mood	0.63	0.89	0.95	0.98
Decreased appetite	0.09	0.13	0.00	0.58
Increased appetite	0.08	0.09	0.74	0.00
Decreased weight	0.09	0.04	0.09	0.33
Increased weight	0.07	0.04	0.48	0.05
Impaired concentration	0.24	0.74	0.80	0.82
Negative view of self	0.28	0.55	0.71	0.68
Suicidal ideation	0.07	0.11	0.16	0.21
Lack of general interest	0.27	0.68	0.76	0.77
Fatigue	0.39	0.84	0.88	0.86
Psychomotor retardation	0.12	0.31	0.48	0.51
Psychomotor agitation	0.23	0.20	0.33	0.51

CHAPTER III

GENDER DIFFERENCES IN CHANGES IN DEPRESSION SUBGROUPS IN STAR*D - A LATENT TRANSITION ANALYSIS

Abstract

Objective: To characterize gender differences in latent statuses of major depression and changes in these statuses among adults receiving citalopram in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial.

Method: Latent transition analysis was applied to data from 387 men and 755 women who completed baseline and week 12 study visits in level 1 of STAR*D. Items from the self-report version of the Quick Inventory of Depressive Symptomatology were used as indicators of latent depression status.

Results: Four statuses were identified for each gender at baseline and week 12. Baseline statuses for men were Mild_{Men} (10%), $\text{Moderate}_{\text{Men}}$ (53%), Severe with Psychomotor Slowing (20%), and Severe with Psychomotor Agitation (17%). For men at week 12, the statuses were $\text{Symptom Resolution}_{\text{Men}}$ (41%), Mild_{Men} (36%), $\text{Moderate}_{\text{Men}}$ (18%), and Severe with Psychomotor Slowing (5%). Baseline statuses for women were $\text{Mild}_{\text{Women}}$ (21%), $\text{Moderate}_{\text{Women}}$ (30%), Severe with Increased Appetite (16%), and Severe with Decreased Appetite (34%). For women, week 12 statuses were $\text{Symptom Resolution}_{\text{Women}}$ (65%), $\text{Mild}_{\text{Women}}$ (23%), $\text{Moderate}_{\text{Women}}$ (9%), and Severe with Psychomotor Disturbances (3%). Men in the Mild status at baseline were most likely to transition to Symptom Resolution (probability = 69%). Men in the Severe with Psychomotor Agitation status did not transition to Symptom Resolution. Women in the Moderate status had the greatest chance of moving to Symptom Resolution (87%). Women in the Severe with Decreased Appetite status had the lowest chance of transitioning to Symptom Resolution (46%).

Conclusions: Depression severity and psychomotor disturbances distinguished depression statuses for men whereas severity, appetite, and insomnia distinguished statuses for women. After treatment, depression severity characterized statuses for men but psychomotor and sleep disturbances characterized women. This work highlights the need to consider symptoms and not focus solely on summary rating scores when treating depression.

Introduction

Heterogeneity in major depression is well established, with more than 1,400 combinations of DSM criteria symptoms possible[12] and considerable differences in illness course, prominent symptoms, and treatment response.[13] Depression subtypes based on symptom patterns have been proposed but the extent to which these subtypes or even individual symptoms change over time is not clear. It has been seen that individual symptoms[29,30] and subtypes[31] can change throughout depressive episodes but few studies have been conducted and treatment is rarely considered. Research about the longitudinal stability of subtypes, including transitions between subtypes in response to treatment, could inform efforts to address depression symptom heterogeneity in personalizing treatment strategies, a goal of precision medicine and the National Institute of Mental Health's Research Domain Criteria (RDoC) initiative.[32]

Latent transition analysis (LTA) is one potential method for efficiently determining depression subgroups by incorporating numerous patterns of depression symptoms and examining how these subgroups change over time. LTA is a longitudinal extension of latent class analysis, a finite mixture modeling method that posits that there are mutually exclusive and exhaustive "statuses," or groups, of people that can be distinguished by values of an unobserved variable.[34] This latent variable is comprised of observed indicator variables of characteristics such as depression symptoms. LTA can be used to model changes in the qualitative nature of these statuses, in the prevalence of each status at each time, and in the individuals' membership in these statuses over time.

To our knowledge, only two other studies to date have used LTA to look at the stability of types of depression over time and treatment was unknown in both studies.[86,87]

This study sought to use LTA to characterize statuses of major depression for men and women and to examine changes in statuses after receiving citalopram treatment in level 1 of the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial. The objectives were to: 1) characterize patterns of depression symptoms at the beginning of a treatment trial; 2) examine changes in these patterns, including changes in the descriptive nature of each status and how participants move between statuses, over 12 weeks of treatment; and 3) examine gender differences in these statuses.

Methods

Study Participants

The publicly available, de-identified dataset from level 1 of STAR*D was used. STAR*D was a pragmatic clinical trial originally designed to assess the effectiveness of a variety of pharmacological and psychosocial treatments for moderate-to-severe non-psychotic major depression.[42] Between July 2001 and April 2004, 4,041 participants were enrolled from 18 primary care and 23 outpatient psychiatric sites. In level 1, participants received citalopram for up to 14 weeks. Study visits were conducted at 2, 4, 6, 9, and 12 weeks with an optional visit at week 14. Of the evaluable sample of participants who scored ≥ 14 on the baseline 17-item Hamilton Rating Scale of Depression (HRSD) and completed ≥ one post-baseline visit,[45] almost 30% achieved remission (HRSD ≤ 7 or last observed 16-item Quick Inventory of Depressive Symptomatology ($QIDS\text{-}SR_{16}$) score ≤ 5) and 47% experienced response (≥ 50% reduction in baseline $QIDS\text{-}SR_{16}$ score).[33]

Additional details about STAR*D have been described elsewhere.[33,42]

This analysis sought to quantify transitions in depression statuses over a 12-week period. As such, participants must have completed baseline and week 12 study visits. While LTA can theoretically accommodate some missing data for the indicator variables used to define the statuses at the different time points, we used a complete cases analysis because we were unable to achieve model convergence when all participants, regardless of whether they completed the week 12 assessment, were included. Additionally, 56 participants were excluded because they had missing data on key variables of interest (e.g. QIDS-$SR_{16,}$ etc.). The remaining 1,142 participants comprised the study sample.

STAR*D participants provided written informed consent after receiving a complete description of the study at enrollment. The protocol was originally approved and monitored by the institutional review boards at the trial's national coordinating center, the data coordinating center, clinical sites, and the Data, Safety, and Monitoring Board of the National Institute of Mental Health. The Institutional Review Board at the University of Massachusetts Medical School determined that this secondary analysis was not human subject research as defined by Department of Health and Human Services and Food and Drug Administration regulations.

Measures

Indicators of Latent Status Membership

The individual QIDS-SR_{16} items collected at baseline and week 12 were used as indicators of latent depression status. The QIDS-SR_{16} measures depression severity and contains 16 items corresponding to the nine DSM-IV symptom criterion domains for

major depression.[46] All items except weight change reflect the past seven days whereas the weight change items reflect the last 14 days. The individual item scores range from 0-3, with any item score ≥ 2 indicating that the symptom meets the DSM-IV threshold for major depression.[44] Thus, for this analysis, all items were dichotomized with individual item scores ≤ 1 denoting the absence of a DSM-IV criterion symptom. Scores ≥ 2 signify the presence of a criterion symptom. The QIDS-SR_{16} instructions stipulate that only one item on appetite increase or decrease and weight increase or decrease should be completed. We included four separate indicator variables to better capture the direction changes in appetite and weight.

Gender Differences in Latent Statuses

Gender differences have been documented in many aspects of depression, including rates, severity, course, risk factors, and symptoms. Women experience depression more often and more severely than men.[17–20] Women also seem to be more likely to experience somatic, atypical, and anxiety symptoms[21–23] and to have more transitions between subtypes.[87] Because of these established differences, we hypothesized that men and women would not experience depression and treatment response similarly and sought to evaluate if depression statuses at each time were qualitatively the same for men and women.

Analysis

The statistical analysis was conducted in three phases: 1) describing the overall sample and comparing characteristics of those included and excluded from the analysis; 2) developing the basic latent status model for transitions between baseline and week 12

of the trial; and 3) determining if the depression statuses differed for men and women. First, we compared demographic and clinical characteristics of those meeting our eligibility criteria to those excluded from the analysis. Because trivial differences were likely to achieve statistical significance owing to the large sample size, we considered 5% absolute differences in the prevalence estimates between those included and excluded to be noteworthy. For the second phase of the analysis, we fit multiple basic LTA models, varying the number of statuses. Selecting the optimal number of latent statuses was informed by fit statistics such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) and the interpretability of the statuses. The relative fit of the models was emphasized over traditional hypothesis testing because of issues arising from LTA models having very large degrees of freedom and extreme sparseness.[34] After the optimal number of latent statuses was selected for the basic LTA model, the assumption of measurement invariance of the qualitative nature of the statuses was tested across time. This was done using the difference G^2 test,[34] AIC and BIC to compare two nested models: 1) a model in which the item-response probabilities were allowed to vary between baseline and week 12; and 2) a model in which the item-response probabilities were constrained for both time points. The measurement invariance assumption was not supported and the analysis proceeded under the assumption that there were differences in the depression statuses at baseline and week 12.

In the third phase of analysis, LTA models with gender as a grouping variable were fit. The optimal number of latent statuses was selected using a similar strategy as described in phase two of the analysis. Measurement invariance between men and women

and across time was then tested using the G^2 difference test to compare several nested models: 1) a model in which the item-response probabilities were free to vary between men and women and across times; 2) a model in which the item-response probabilities were constrained to be equal across times but were allowed to vary by gender; 3) a model in which item-response probabilities were constrained to be equal in men and women but were allowed to vary by time; and 4) a model in which item-response probabilities were constrained to be equal across gender and times. Models 2-4 were compared to model 1 using separate tests. Measurement invariance could not be established across genders and times and thus separate LTA models for men and women were ultimately investigated. The analysis described above for the basic LTA of all participants was conducted separately for each gender.[34] Analyses were conducted using PROC LTA[53,88] in SAS 9.3 (SAS Institute, Inc., Cary, NC).

Results

Overall Sample

Participants who completed a baseline and week 12 study visit were mostly women, were in their early- to mid-40's, white, and had moderate to severe depression (Table 3.1). Men who were included in this analysis were similar to those who were excluded. When compared to included women, the women who were excluded were two years younger on average. The excluded women were also more likely to be black or African American, and more likely to have comorbid anxiety disorders such as generalized anxiety disorder, PTSD, and social phobia at baseline when differences in frequency ≥ 5% were considered.

The fit statistics for LTA models in the first phase of analysis supported a four-status model (Supplementary Table 3.1). The G^2 difference test of the measurement invariance hypothesis of the item-response probabilities across time was significant (G_2^2-G_1^2 =252.67, df = 64, p <0.0001) and thus it could not be assumed that the latent statuses were qualitatively the same at baseline and week 12. The statuses at baseline were primarily distinguished by depression severity and were labeled Mild Depression, Moderate Depression, Severe Depression with Insomnia, and Severe Depression with Increased Appetite (Supplementary Table 3.2). Thirty percent of participants were likely to be in the Severe with Insomnia status, 28% in the Mild status, 25% in the Moderate status, and 18% in the Severe with Increased Appetite status. Members of all statuses were likely to endorse having mid-nocturnal insomnia and sad mood. All participants except for members of the Mild status were likely to report having impaired concentration, negative self-view, lack of general interest, and fatigue. The statuses at week 12 were still mainly differentiated by severity but with the statuses identified as Symptom Resolution, Mild Depression with Insomnia, Moderate Depression, and Severe Depression with Insomnia. The symptom pattern of the Severe with Insomnia status at week 12 was similar to that of baseline except for having higher probabilities of the insomnias, impaired concentration, and negative self-view. The majority of participants at week 12 were likely to belong to the Symptom Resolution status (61%) and the fewest were in the Severe with Insomnia status (6%).

Men

On average, the men included in this analysis were likely to be middle-aged, white, and have a QIDS-SR_{16} score indicative of moderate depression at baseline (Table 3.1). Anxiety disorders were the most common psychiatric comorbidity, with social phobia being the most prevalent. The majority of men were experiencing sleep-onset insomnia, mid-nocturnal insomnia, sad mood, impaired concentration, lack of interest, and fatigue at baseline but mid-nocturnal insomnia was the only depression symptom experienced by a majority by week 12 (Table 3.2).

A four-status LTA model fit the data best at both baseline and week 12. At baseline, the statuses were Mild $\text{Depression}_{\text{Men}}$, Moderate $\text{Depression}_{\text{Men}}$, Severe Depression with Psychomotor Slowing, and Severe Depression with Psychomotor Agitation (Table 3.3). At week 12, the Mild_{Men}, $\text{Moderate}_{\text{Men}}$, and Severe with Psychomotor Slowing statuses were still present but a Symptom $\text{Resolution}_{\text{Men}}$ status also emerged. The G^2 difference test suggested that measurement invariance of the overall item-response probabilities across time could not be assumed, meaning that the statuses were qualitatively different at each time (G_2^2-G_1^2 = 117.53, df = 64, $p < 0.0001$) (Supplementary Table 3.3). After thoroughly considering the interpretation of each individual status at each time, however, measurement invariance was assumed for the Mild_{Men}, $\text{Moderate}_{\text{Men}}$, and Severe with Psychomotor Slowing statuses between baseline and week 12. At baseline, the most men were likely to belong to the $\text{Moderate}_{\text{Men}}$ status (53% of men) (Figure 3.1). The least prevalent status was Mild_{Men} (10% of men). Members of the Mild_{Men} status at baseline were the most likely to transition to the

Symptom Resolution$_{Men}$ status, with a 69% chance of making this transition. None of those in the Severe with Psychomotor Agitation status transitioned to the Symptom Resolution$_{Men}$ status at week 12. Men in the Severe with Psychomotor Slowing statuses were least likely to transition to a less severe status, with a 23% of remaining in the Severe Psychomotor Slowing status at week 12. Conversely, men in the Mild$_{Men}$ status at baseline were the least likely to move into a more severe status, with none transitioning to the Moderate$_{Men}$ or Severe with Psychomotor Slowing status. Overall, 41% of men were likely to be in the Symptom Resolution$_{Men}$ status at week 12.

Women

At baseline, the women included in this analysis were mostly in their early 40s, white, and severely depressed (Table 3.1). Anxiety disorders were the most frequently occurring psychiatric comorbidities at baseline. Social phobia was the most common comorbidity. Women who were excluded were more likely than those included to be black or African-American, have GAD, have PTSD, or have social phobia. Of the criterion depression symptoms, the majority of included women were experiencing sleep-onset insomnia, mid-nocturnal insomnia, sad mood, impaired concentration, negative self-view, lack of interest, and fatigue at baseline (Table 3.2). Mid-nocturnal insomnia was the only symptom still experienced by most women by week 12.

An LTA model with four statuses most appropriately described the baseline and week 12 data. At baseline, the four statuses were Mild$_{Women}$, Moderate$_{Women,}$ Severe with Increased Appetite, and Severe with Decreased Appetite (Table 3.4). The statuses at week 12 were Symptom Resolution$_{Women}$, Mild$_{Women}$, Moderate$_{Women}$, and Severe with

Psychomotor Disturbances. The qualitative descriptions of the overall statuses appeared to change over time (G_2^2-G_1^2 = 168.95, df = 64, $p < 0.0001$) (Supplementary Table 3.3) but the item-response probabilities for the Mild statuses at each time were similar. As such, measurement invariance was imposed on these statuses. The same was true for the Moderate statuses at baseline and week 12.

The baseline Severe with Decreased Appetite status was the most prevalent (34%), followed by the Moderate status (30%) (Figure 3.2). The Severe with Increased Appetite status was the least prevalent (16%). The Symptom Resolution$_{\text{Women}}$ status was the most prevalent (65%) at week 12 while the Severe with Psychomotor Disturbances was the least common (3%). Women who were most likely to belong to the Moderate status at baseline had the highest chance of transitioning to the Symptom Resolution$_{\text{Women}}$ status at week 12 (87%). Women in the Severe with Increased Appetite and Severe with Decreased Appetite statuses were also more likely to transition to the Symptom Resolution status than to any other status (transition probability = 48% and 46%, respectively). The probability that women in the Mild$_{\text{Women}}$ and Moderate$_{\text{Women}}$ statuses at baseline transitioned into the Severe with Psychomotor Disturbances group at week 12 was zero. While women in the Mild$_{\text{Women}}$ status at baseline had a 26% chance of remaining in that status at week 12, women in the Moderate$_{\text{Women}}$ status at baseline had a 13% of staying in that status at week 12. Women in the Severe with Increased Appetite status and those in the Severe with Decreased Appetite status had a 3% and 8% chance of moving to the Severe with Psychomotor Disturbances status, respectively.

Discussion

The primary objective of this study was to use LTA to explore changes in depression statuses for men and women through 12 weeks of citalopram treatment. Four different statuses were identified at each time point for each gender. Depression severity and psychomotor disturbances were common distinguishing features of the statuses for both men and women. The majority of men and of women were likely to transition into statuses that were characterized by fewer dominant symptoms, an improvement obscured when defining treatment success by looking only at those meeting criteria for complete symptom remission.

In men, we found that severe depression was further differentiated by psychomotor symptoms. This finding agrees with previous work noting psychomotor disturbances in depression.[89] Despite such focus on these psychomotor symptoms and the assumption that they are core features of melancholic depression, it remains unclear if both symptoms are core features of melancholia or are shared across depression subtypes.[90] Agitation is also associated with mood-switching in unipolar depression[91] and with comorbid substance dependence,[92] highlighting the need to determine how to improve treatment to resolve specific symptoms. After 12 weeks of treatment, psychomotor disturbances no longer differentiated depression subgroups in men but we found that men who started in the Severe with Psychomotor Slowing group had the lowest chances of improving. This is consistent with previous work demonstrating that psychomotor retardation is related to greater depression severity[93] and poor response to citalopram.[94] Psychomotor retardation appears to involve the hypothalamic-pituitary-

adrenal (HPA) axis, basal ganglia, and prefrontal cortex[95] and is associated with differential treatment response, with improvements seen in depression with psychomotor slowing when treated with fluoxetine and sertraline but not with citalopram.[96] This suggests potential areas for targeted treatment.

In women, appetite changes were influential in distinguishing statuses at baseline. This is similar to our previous findings that increased appetite was associated with separate severe depression subgroups in a latent class analysis of STAR*D participants at baseline (Chapter II). That both increased and decreased appetite symptoms were prominent at baseline is in line with melancholic depression being associated with diminished appetite and atypical depression with hyperphagia. These behaviors have been proposed to result from distinct dysregulations of the stress response system, particularly related dysfunction in activation of corticotropin-releasing hormone (CRH).[97] It is encouraging, however, that the subgroups experiencing appetite changes were highly likely to transition to either the Symptom Resolution$_{Women}$ or Mild$_{Women}$ statuses. That insomnia was still experienced by almost a third of women at week 12 is consistent with previous research. Insomnia is known to play a large role in depression, especially for women,[60] and is associated with depression recurrence[57] and inadequate treatment response.[98] Insomnia is also a side effect of many antidepressants. Considering residual symptoms such as insomnia may be necessary for improving depression treatment since residual symptoms increase risk of relapse[44] and since doing so may elucidate important clues for how to target symptoms to produce overall depression remission and improved functioning.[99]

While this analysis used data from the largest and longest depression treatment trial to examine latent depression statuses and changes in statuses membership over time, our results do have several limitations. This is a post-hoc analysis of data that was not originally collected for the purposes of performing such subgroup analyses. Furthermore, this complete case analysis included not only individuals who were seeking care in primary care or psychiatric outpatient settings and enrolled in a clinical trial but was limited to those who attended both the baseline and week 12 study visits. While there did not appear to be substantial differences in sociodemographic and clinical characteristics between those who were included and excluded from our analysis, biased estimates from the complete case analysis are still possible and generalizability of results might be limited. STAR*D, however, had broader inclusion criteria than most clinical trials and this sample is more representative of people seeking outpatient depression treatment than are participants from other clinical trials.[40] Overall, level 1 of STAR*D is a strong source of observational data where, unlike in other naturalistic studies, treatment is known.

Psychomotor disturbances and the other depression symptoms used as indicators of the latent depression status variable were drawn from the self-report QIDS and not confirmed through clinical observation. This could introduce response bias, produce inaccurate estimates of symptoms, and limit comparison to other studies of depression, particularly those using objective assessments of psychomotor disturbances. The timing of the assessment and the length of the trial must be considered when interpreting our findings. The QIDS-SR_{16} inquires about symptoms from the previous 1-2 weeks, which might not be adequate to capture all symptom changes, especially those related to longer

processes such as weight changes. Furthermore, although 12 weeks is not an uncommon length of time for a clinical trial, it might not be sufficient to identify all the transitions that might occur with antidepressant treatment.

Despite these considerations, this analysis is noteworthy in several ways. Although only complete cases were used, this sample is the largest in which LTA has been used to examine at the stability of depression subgroups during known antidepressant treatment. LTA models consolidate large arrays of contingency table data representing multidimensional constructs, such as depression, into meaningful subgroups. This efficiency is especially valuable when examining the heterogeneity that can arise from combinations of 16 depression symptoms at two time points. One study on the frequency of symptom patterns observed in STAR*D reported 1,030 unique symptom profiles at baseline, the majority of which were endorsed by five or fewer participants.[100] Because LTA also allows for measurement error, it does not penalize models when participants' symptoms do not clearly suggest membership in one particular status.

Our results indicate that men and women experienced different patterns of depression symptoms during citalopram treatment in level 1 of STAR*D and that these patterns were not stable over time. While there has been continual interest in elucidating subgroups of depression, the stability of subgroups has received far less attention. Our study demonstrates the potential of using the person-centered approach of LTA to detect subgroups comprised of important symptom patterns and to examine changes in these subgroups during treatment. LTA can be used to examine heterogeneity and identify subgroups, work that can eventually inform the development of tailored treatments.

Table 3.1: Baseline demographic and clinical characteristics of men and women by study inclusion status

	Men			Women		
Characteristic	**Included N = 387**	**Excluded N = 657**	***p*-value**	**Included N = 755**	**Excluded N = 1,077**	***p*-value**
Age at study entry, mean (SD)	44.6 (12.5)	42.3 (12.6)	0.004	41.4 (13.2)	39.4 (12.9)	0.001
Race, n (%)						
White	303(78.3)	519 (79.0)	0.871	581 (77.0)	776 (72.1)	0.008
Black or African American	62 (16.0)	98 (14.9)		117 (15.5)	229 (21.3)	
Other	22 (5.7)	40 (6.1)		57 (7.6)	72 (6.7)	
Hispanic, n (%)	31 (8.0)	58 (8.8)	0.648	103 (13.6)	182 (16.9)	0.058
Age at onset, mean (SD)	27.3 (14.7)	26.5 (14.2)	0.394	24.8 (14.6)	24.3 (14.1)	0.409
Number of depressive episodes before baseline, mean (SD)	7.9 (12.8)	6.5 (10.3)	0.088	4.5 (6.4)	4.7 (8.3)	0.645
Depression severity (QIDS-SR_{16}) at baseline, mean (SD)	15.3 (3.9)	15.8 (4.0)	0.059	16.1 (4.0)	16.8 (4.0)	<0.001
Psychiatric comorbidities, n (%)						
GAD	74 (19.3)	118 (18.3)	0.689	171 (22.9)	309 (29.2)	0.003
PTSD	78 (20.4)	127 (19.8)	0.812	115 (15.5)	261 (24.8)	<0.001
Bulimia	26 (6.7)	44 (6.8)	0.973	122 (16.4)	178 (16.8)	0.822
Social phobia	110 (28.8)	182 (28.2)	0.831	211 (28.5)	384 (36.3)	<0.001
OCD	44 (11.5)	83 (12.8)	0.518	62 (8.3)	125 (11.8)	0.017
Panic disorder	42 (11.0)	75 (11.6)	0.753	89 (11.9)	164 (15.5)	0.033
Psychosis	48 (12.5)	103 (15.9)	0.138	87 (11.7)	159 (15.0)	0.041
Agoraphobia	35 (9.1)	79 (12.2)	0.129	75 (10.2)	144 (13.7)	0.026
Alcohol abuse/dependence	67 (17.5)	118 (18.2)	0.789	65 (8.7)	92 (8.7)	0.991
Drug abuse/dependence	38 (9.9)	68 (10.5)	0.749	34 (4.6)	68 (6.4)	0.088
Somatization disorder	3 (0.8)	7 (1.1)	0.636	19 (2.6)	40 (3.8)	0.143
Hypochondriasis	6 (1.6)	29 (4.5)	0.012	31 (4.2)	60 (5.7)	0.148

Table 3.2: Frequency of QIDS-SR_{16} items at baseline and week 12 for men and women included in this analysis

	Men (n= 387)		**Women (n= 755)**	
	Baseline	**Week 12**	**Baseline**	**Week 12**
QIDS-SR_{16} items	N (%)	N (%)	N (%)	N (%)
Sleep-onset insomnia	246 (63.6)	125 (32.3)	488 (64.6)	206 (27.3)
Mid-nocturnal insomnia	271 (70.0)	237 (61.2)	541 (71.7)	428 (56.7)
Early morning insomnia	180 (46.5)	95 (24.6)	365 (48.3)	148 (19.6)
Hypersomnia	38 (9.8)	37 (9.6)	114 (15.1)	61 (8.1)
Sad mood	312 (80.6)	93 (24.0)	641 (84.9)	137 (18.2)
Decreased appetite	80 (20.7)	22 (5.7)	165 (21.9)	31 (4.1)
Increased appetite	41 (10.6)	25 (6.5)	146 (19.3)	51 (6.8)
Decreased weight	52 (13.4)	11 (2.8)	89 (11.8)	22 (2.9)
Increased weight	27 (7.0)	18 (4.7)	92 (12.2)	34 (4.5)
Impaired concentration	220 (56.9)	77 (19.9)	454 (60.1)	106 (14.0)
Negative view of self	179 (46.3)	65 (16.8)	417 (55.2)	110 (14.6)
Suicidal ideation	51 (13.2)	19 (4.9)	84 (11.1)	29 (3.8)
Lack of interest	226 (58.4)	70 (18.1)	444 (58.8)	135 (17.9)
Fatigue	244 (63.1)	69 (17.8)	563 (74.6)	158 (20.9)
Psychomotor retardation	126 (32.6)	46 (11.9)	263 (34.8)	64 (8.5)
Psychomotor agitation	130 (33.6)	59 (15.3)	221 (29.3)	82 (10.9)

Table 3.3: Item-response probabilities from a four-status LTA of QIDS-SR_{16} indicators for men

	Men (n=387)							
	Baseline Latent Statuses				Week 12 Latent Statuses			
QIDS-SR_{16} items	**$Mild_{Men}$ [1]**	**$Moderate_{Men}$ [2]**	**Severe with Psychomotor Slowing [3]**	**Severe with Psychomotor Agitation**	**Symptom $Resolution_{Men}$**	**$Mild_{Men}$ [1]**	**$Moderate_{Men}$ [2]**	**Severe with Psychomotor Slowing [3]**
Sleep onset insomnia	0.35	0.56	0.82	0.83	0.13	0.35	0.56	0.82
Mid-nocturnal insomnia	0.71	0.69	0.83	0.76	0.39	0.71	0.69	0.83
Early morning insomnia	0.35	0.34	0.66	0.86	0.00	0.35	0.34	0.66
Hypersomnia	0.05	0.15	0.12	0.00	0.09	0.05	0.15	0.12
Sad mood	0.11	0.84	0.89	0.89	0.06	0.11	0.84	0.89
Decreased appetite	0.05	0.07	0.44	0.49	0.00	0.05	0.07	0.44
Increased appetite	0.07	0.17	0.08	0.00	0.00	0.07	0.17	0.08
Decreased weight	0.04	0.07	0.16	0.34	0.01	0.04	0.07	0.16
Increased weight	0.05	0.10	0.06	0.00	0.03	0.05	0.10	0.06
Impaired concentration	0.59	0.51	0.93	0.59	0.01	0.59	0.51	0.93
Negative self-view	0.41	0.47	0.76	0.41	0.02	0.41	0.47	0.76
Suicidal ideation	0.03	0.15	0.24	0.03	0.02	0.03	0.15	0.24
Lack of general interest	0.72	0.51	0.81	0.72	0.00	0.72	0.51	0.81
Fatigue	0.63	0.57	0.96	0.63	0.00	0.63	0.57	0.96
Psychomotor retardation	0.27	0.21	0.86	0.27	0.00	0.27	0.21	0.86
Psychomotor agitation	0.62	0.26	0.45	0.62	0.00	0.62	0.26	0.45

[1]The the item-response probabilities for the $Mild_{Men}$ statuses were constrained to be equal at baseline and week 12.
[2]The item-response probabilities for the $Moderate_{Men}$ statuses were constrained to be equal at baseline and week 12.
[3]The item-response probabilities for the Severe with Pychomotor Slowing statuses were constrained to be equal at baseline and week 12.

Table 3.4: Item-response probabilities from four-status LTA of QIDS-SR_{16} indicators for women

	Women (n=775)							
	Baseline Latent Statuses				**Week 12 Latent Statuses**			
QIDS-SR_{16} items	**$Mild_{Women}$ [1]**	**$Moderate_{Women}$ [2]**	**Severe with Increased Appetite**	**Severe with Decreased Appetite**	**Symptom $Resolution_{Women}$**	**$Mild_{Women}$ [1]**	**$Moderate_{Women}$ [2]**	**Severe with Psychomotor Disturbances**
Sleep onset insomnia	0.54	0.46	0.69	0.86	0.12	0.54	0.46	0.82
Mid-nocturnal insomnia	0.77	0.56	0.73	0.82	0.48	0.77	0.56	0.87
Early morning insomnia	0.47	0.24	0.58	0.66	0.06	0.47	0.24	0.74
Hypersomnia	0.01	0.30	0.22	0.10	0.06	0.01	0.30	0.13
Sad mood	0.44	0.79	0.99	0.97	0.03	0.44	0.79	1.00
Decreased appetite	0.07	0.10	0.00	0.51	0.01	0.07	0.10	0.28
Increased appetite	0.08	0.17	0.85	0.00	0.04	0.08	0.17	0.22
Decreased weight	0.06	0.05	0.00	0.28	0.01	0.06	0.05	0.04
Increased weight	0.07	0.04	0.55	0.03	0.03	0.07	0.04	0.20
Impaired concentration	0.18	0.65	0.72	0.75	0.02	0.18	0.65	1.00
Negative self-view	0.22	0.52	0.71	070	0.03	0.22	0.52	0.87
Suicidal ideation	0.09	0.05	0.16	0.16	0.00	0.09	0.05	0.26
Lack of general interest	0.25	0.55	0.77	0.77	0.06	0.25	0.55	0.83
Fatigue	0.33	0.79	0.89	0.87	0.06	0.33	0.79	0.95
Psychomotor retardation	0.11	0.32	0.49	0.49	0.00	0.11	0.32	0.57
Psychomotor agitation	0.21	0.21	0.27	0.46	0.02	0.21	0.21	0.53

[1]The item-response probabilities for the $Mild_{Women}$ statuses were constrained to be equal at baseline and week 12.
[2]The item-response probabilities for the $Moderate_{Women}$ statuses were constrained to be equal at baseline and week 12.

Figure 3.1: Latent status prevalences and probabilities of transitioning in status membership from baseline to week 12 for men

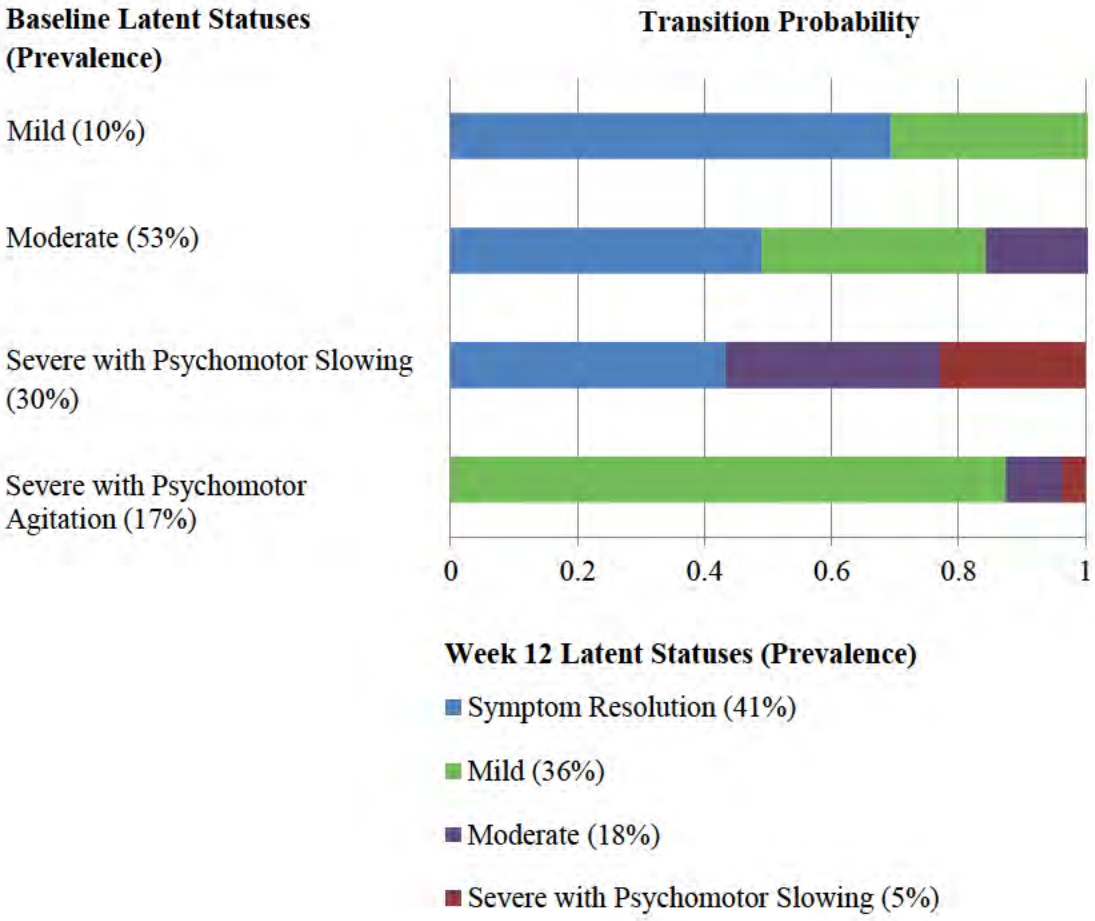
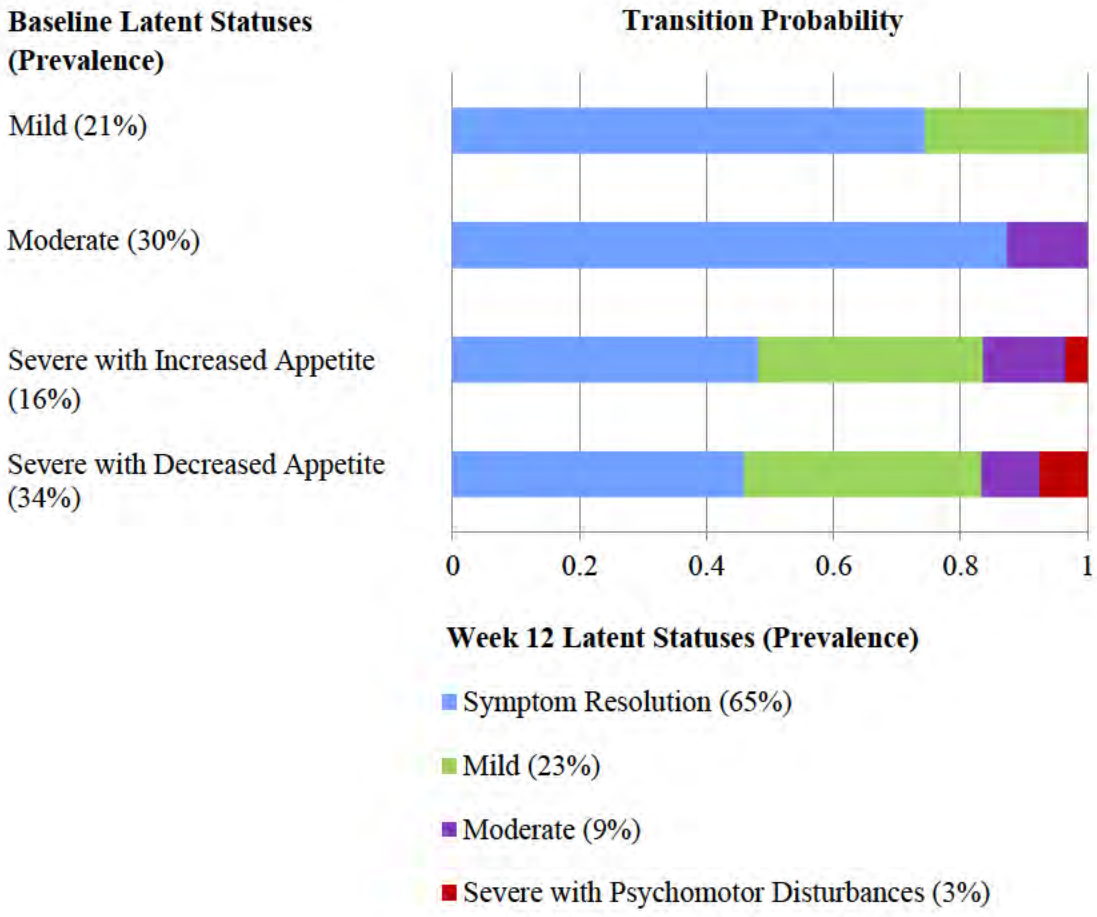


Figure 3.2: Latent status prevalences and probabilities of transitioning in status membership from baseline to week 12 for women



Supplementary Table 3.1: Fit indices for basic LTA models of QIDS-SR_{16} at baseline and 12 weeks

# of Statuses	*df*	G^2	AIC	BIC
No Measurement Invariance				
2	4294967228	15587.75	15721.75	16059.46
3	4294967191	15039.40	15247.40	15771.62
4	4294967152	14838.00	15124.00	15844.79
5	4294967111	14645.86	15013.86	15941.32
6	4294967068	14518.52	14972.52	16116.72
7	4294967023	14392.01	14936.01	16307.03
Measurement Invariance Rho parameters constrained to be equal across time				
2	4294967260	16044.22	16114.22	16290.64
3	4294967239	15488.38	15600.38	15882.65
4	4294967216	15081.96	15239.96	15638.16
5	4294967191	14917.95	15125.95	15650.17
6	4294967164	14789.00	15051.00	15711.31
7	4294967135	14685.09	15005.09	15811.57

Supplementary Table 3.2: Latent status prevalences and item-response probabilities from four-status LTA of QIDS-SR_{16} indicators without constraints on any parameters

	Baseline Latent Statuses				Week 12 Latent Statuses			
	Mild	**Moderate**	**Severe with Insomnia**	**Severe with Increased Appetite**	**Symptom Resolution**	**Mild with Insomnia**	**Moderate**	**Severe with Insomnia**
Latent Status Prevalences								
	0.28	0.25	0.30	0.18	0.61	0.24	0.10	0.06
Item-Response Probabilities of QIDS-SR_{16} Indicators								
Sleep onset insomnia	0.56	0.49	0.87	0.30	0.11	0.56	0.38	0.89
Mid-nocturnal insomnia	0.73	0.57	0.83	0.69	0.47	0.77	0.63	0.89
Early morning insomnia	0.41	0.24	0.72	0.51	0.07	0.48	0.18	0.74
Hypersomnia	0.04	0.23	0.08	0.23	0.06	0.04	0.32	0.13
Sad mood	0.63	0.85	0.95	0.94	0.04	0.27	0.60	0.95
Decreased appetite	0.06	0.22	0.49	0.00	0.01	0.07	0.04	0.38
Increased appetite	0.08	0.00	0.00	0.78	0.03	0.08	0.25	0.11
Decreased weight	0.06	0.09	0.28	0.00	0.01	0.07	0.03	0.06
Increased weight	0.06	0.01	0.03	0.42	0.03	0.07	0.08	0.10
Impaired concentration	0.23	0.66	0.78	0.74	0.02	0.19	0.52	0.93
Negative self-view	0.24	0.57	0.68	0.63	0.03	0.22	0.45	0.72
Suicidal ideation	0.09	0.08	0.18	0.12	0.01	0.06	0.12	0.18
Lack of general interest	0.22	0.69	0.77	0.71	0.05	0.23	0.55	0.70
Fatigue	0.35	0.81	0.84	0.88	0.04	0.25	0.72	0.80
Psychomotor retardation	0.09	0.30	0.54	0.46	0.00	0.13	0.35	0.54
Psychomotor agitation	0.19	0.20	0.54	0.26	0.02	0.27	0.18	0.53

Supplementary Table 3.3. Fit indices for LTA models of QIDS-SR_{16} examining gender differences at baseline and week 12

# of Statuses	*df*	G^2	AIC	BIC
Gender as a Grouping Variable				
No Measurement Invariance				
2	8589934459	16857.52	17125.52	17800.95
3	8589934389	16312.55	16728.55	17776.98
4	8589934317	16051.45	16623.45	18065.04
5	8589934243	15803.81	16539.81	18394.73
6	8589934167	15564.86	16472.86	18761.26
7	8589934089	15414.68	16502.68	19244.73
Measurement Invariance: Rho parameters constrained to be equal across times				
2	8589934523	17357.47	17497.47	17850.31
3	8589934485	16764.95	16988.95	17553.49
4	8589934445	16344.33	16660.33	17456.73
5	8589934403	16181.53	16597.53	17645.96
6	8589934359	15999.19	16523.19	17843.81
7	8589934313	15898.41	16538.41	18151.38
Measurement Invariance: Rho parameters constrained to be equal across time and across genders				
2	8589934555	17432.95	17508.95	17700.49
3	8589934533	16865.54	16993.54	17316.13
4	8589934509	16443.64	16631.64	17105.45
5	8589934483	16278.25	16534.25	17179.44
6	8589934455	16129.92	16461.92	17298.65
7	Model did not converge			
Men Only				
No Measurement Invariance				
2	4294967228	6159.37	6293.37	6558.59
3	4294967191	6015.72	6223.72	6635.39
4	4294967152	5884.49	6170.49	6736.55
5	4294967111	5793.33	6161.33	6889.68
6	4294967068	5643.23	6097.23	6995.79
7	4294967023	5624.41	6168.41	7245.10
Measurement Invariance: Rho parameters constrained to be equal across times				
2	4294967260	6320.06	6390.06	6528.60
3	4294967239	6113.50	6225.50	6447.17
4	4294967216	6002.34	6160.34	6473.05

# of Statuses	*df*	G^2	AIC	BIC
5	4294967191	5963.18	6171.18	6582.86
6	4294967164	5851.89	6113.89	6632.44
7	4294967135	5823.23	6143.23	6776.58
Women Only				
No Measurement Invariance				
2	4294967228	10698.15	10832.15	11142.14
3	4294967191	10296.83	10504.83	10986.01
4	4294967152	10166.95	10452.95	11114.57
5	4294967111	10010.48	10378.48	11229.80
6	4294967068	9891.20	10345.20	11395.47
7	4294967023	9791.08	10335.08	11593.55
Measurement Invariance: Rho parameters constrained to be equal across times				
2	4294967260	11037.41	11107.41	11269.34
3	4294967239	10651.46	10763.46	11022.55
4	4294967216	10341.99	10499.99	10865.50
5	4294967191	10440.05	10648.05	11131.95
6	4294967164	10150.84	10412.84	11018.94
7	4294967135	10046.69	10366.69	11106.96

CHAPTER IV

FUNCTIONAL IMPAIRMENT AND CHANGES IN DEPRESSION SUBTYPES FOR WOMEN IN STAR*D - A LATENT TRANSITION ANALYSIS

Abstract

Objective: To characterize the association between functional impairment and major depression subgroups at baseline and to characterize changes in depression subgroups by level of functional impairment in women receiving citalopram in the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial.

Method: We identified 755 women who completed baseline and week 12 study visits in level 1 of STAR*D. Indicators used to define the latent depression subgroups were self-reported Quick Inventory of Depressive Symptomatology. The Work and Social Adjustment Scale was used to classify women as having normal/significant or major functional impairment at baseline. A latent transition analysis model with level of functional impairment used as a grouping variable provided estimates of the prevalence of latent depression status membership and transition probabilities.

Results: Sixty-nine percent of women were classified as having major functional impairment at baseline. Four-status LTA models of latent depression statuses fit the data best. Regardless of functional impairment level, the depression statuses were differentiated by severity, appetite changes, psychomotor disturbances, and insomnia. Sixty-seven percent of women with major impairment belonged to severe depression statuses at baseline and 5% at week 12. Among women with normal/significant functional impairments, 37% belonged to a severe depression at baseline nit no one belonged to such a status at week 12. Regardless of functional impairment level, the majority of women were likely to transition to a Symptom Resolution status at week 12 (67% of women with normal/significant functional impairment; 60% of women with

major impairment). Women with baseline major impairment who were in the Severe with Psychomotor Agitation at the beginning of the study were least likely to transition to the Symptom Resolution status (4% chance).

Conclusions: Level of functional impairment was related to both the kind of depression and the likelihood of moving to a different depression status for women treated with citalopram in level of 1 of STAR*D. These results underscore the need to incorporate not only depression symptoms but also functioning in the assessment and treatment of major depression.

Introduction

Major depression is the second leading cause of disability worldwide[101] and is associated with impairments in daily functioning and quality of life. In 2012, more than 10 million adults in the U.S. had at least one major depressive episode with severe impairment in the past year.[17] The extent of functional impairment seen with depression often exceeds that which is associated with many other common illnesses.[102] Depression has the greatest impact on total work impairment of any chronic health disorder.[103] This includes effects on lifetime employment opportunities due to lower education levels, productivity, presenteeism, absenteeism, and accidents.[104] Almost 80% of people with major depression in the previous year have reported at least some interference with their ability to function with work.[105]

Despite the profound impact of major depression on functional status, clinical trials of major depression rarely consider improvements in functioning as part of treatment success.[106] Remission in trials is usually defined as reaching a specified score as determined by symptom rating scales and response is generally defined as experiencing a certain decrease in symptom rating score. When functioning is incorporated in trial designs, it typically is considered as a secondary outcome. As such, studies may not be powered to detected differences in improvements in functional status. This lack of emphasis of improved functional status in clinical studies of major depression seems paradoxical for several reasons. First, it occurs despite the inclusion of impaired functioning in the diagnostic criteria for major depression.[59] Second, depression

treatment guidelines recommend assessment of functional impairment. Lastly, treatment guidelines encourage interventions to target improved functioning.[107]

The reliance on symptom ratings to assess improvements in depression is problematic as it may fail to capture changes in functional impairment. Discordance often exists between patients' symptoms and their level of functioning.[106,108] This discordance is not unidirectional. Some people with major depression symptoms rated as mild/moderate can have major deficits in functioning. Conversely, others might have severe symptoms but are able to function normally.[109] In one study of people who were being treated for major depression, half of those with residual depression symptoms but who reported normal functioning considered themselves to be in remission from their depression.[108] Obtaining remission as defined in most studies does not guarantee normal functioning. Subthreshold depression symptoms are associated with functional impairment.[102] This is particularly concerning because most people who achieve remission when treated for depression still experience residual symptoms.[44]

Women and men experience certain aspects of depression differently and these differences appear to extend to functioning. Not only do women with major depression experience greater depression severity than men but women also have a greater burden of depressive disorders when compared to men.[101,110] More women than men experience a major depressive episode with severe impairment, e.g., the episode severely impacts their ability to function in the domains of home management, work, close relationships with others, and/or social life.[17] Women and men also seem to differ in the roles in which

functioning is impaired, with women experiencing more physical limitations and men having impaired social relationships.[24]

While it has been seen that the functional impairment associated with major depression is more pronounced in women, there is a dearth of information about how transitions between depression subtypes during treatment may be different as a function of baseline functional impairment. Better understanding of the association between functional impairment and transitions in depression subtypes could influence treatment.[102,111] Latent transition analysis allows us to explore differences in latent depression statuses by level of functional impairment, including changes in these depression statuses after antidepressant treatment.

The overall objective of this study was to examine differences in functional impairment in latent statuses of depression in women participating in level 1 of STAR*D. Specifically, the aims were to 1) characterize the association between functional impairment and major depression subgroups at baseline; and 2) characterize changes in depression subgroups by level of baseline impairment at the end of 12 weeks of citalopram treatment.

Methods

Study Participants

This analysis used the limited use, de-identified dataset of participants from level 1 of STAR*D. This dataset is publicly available from the National Institute of Mental Health (NIMH). STAR*D was a large pragmatic clinical trial originally designed to evaluate the effectiveness of different pharmacological and psychosocial treatments for

real world patients with moderate to severe non-psychotic major depression.[42] Eighteen primary care and 23 outpatient psychiatric sites enrolled 4,041 participants who were seeking depression treatment from July 2001-April 2004. Participants in level 1 all received citalopram for up to 14 weeks. After enrollment, study visits were conducted at 2, 4, 6, 9, and 12 weeks with an optional visit at week 14. Of the evaluable sample of 2,876 participants who had a score greater than or equal to 14 on the Hamilton Rating Scale of Depression (HRSD) and who completed at least one post-baseline visit, 28% achieved remission as defined by an HRSD score less than or equal to 7. Thirty-three percent of participants achieved remission when it was defined as an observed self-report Quick Inventory of Depressive Symptomatology (QIDS-SR_{16}) score less than or equal to 5.[33] Approximately 47% of the evaluable sample participants achieved response as defined by at least a 50% reduction in baseline QIDS-SR score. STAR*D has been described in further detail elsewhere.[112]

Because our overall goal was to examine how functional impairment is related to transitions in depression subgroups over the 12-week treatment period, participants could not be missing all QIDS-SR_{16} items at baseline and week 12 or the Work and Social Adjustment Scale (WSAS) at baseline to be eligible for this study (n = 755).

STAR*D participants provided written informed consent after receiving a complete description of the study at enrollment. The protocol was originally approved and monitored by the institutional review boards at the trial's national coordinating center, the data coordinating center, clinical sites, and the NIMH Data, Safety, and

Monitoring Board. The institutional review board at the University of Massachusetts Medical School determined that this secondary analysis was not human subject research.

Measures

Indicators of Latent Status Membership

The 16 individual QIDS-SR_{16} items collected at baseline and week 12 were used as the observed indicators of latent depression status. The QIDS-SR_{16} measures overall depression severity and the items correspond to the nine DSM-IV criterion symptoms for major depressive disorder.[46] Although the QIDS-SR16 instructions specify that only one item on appetite increase or decrease and weight increase or decrease should be completed, we included these items as four separate indicator variables to capture the direction of appetite and weight changes. Each item except those pertaining to weight changes reflects the previous seven days. The increased and decreased weight items inquire about changes in the last 14 days. The score for each item ranges from 0-3, with a score ≥ 2 reflecting that the symptom meets the DSM-IV threshold for the presence of a criterion symptom. Accordingly, for this analysis, the items were dichotomized so that a score ≤ 1 indicated the absence of a criterion symptom while a score ≥ 2 indicated the presence of a criterion symptom.[44]

Functional Impairment

Depression-specific functional impairment was measured with the Work and Social Adjustment Scale (WSAS). The WSAS is a 5-item self-report scale assessing work, home management, social activities, private leisure activities, and close relationships.[113] STAR*D participants completed the WSAS via interactive voice

response system (IVR) calls at baseline, week 6, and week 12/study exit. Participants were asked to rate how much their depression specifically impaired these domains. Each item is scored from 0 (no impairment) to 8 (very severe impairment). WSAS total scores greater than 20 indicate major impairment, scores of 10-20 represent significant functional impairment, and scores less than 10 are considered to be within normal ranges of functioning. Only 5% of women had a baseline WSAS score of 0-9 so WSAS scores were dichotomized as normal/significant functional impairment (WSAS = 0-20) and major functional impairment (WSAS ≥ 21).

Analysis

The statistical analysis was conducted in two parts: 1) characterizing the sociodemographic and clinical characteristics of the overall sample by level of baseline functional impairment; and 2) evaluating differences in latent depression statuses throughout treatment by level of baseline functional impairment. First, we calculated descriptive statistics to compare demographic and clinical characteristics of the women categorized as having major functional impairment or normal/significant functional impairment. Since small differences were likely to achieve statistical significance due to the sample size, we considered only absolute differences ≥ 5% in the prevalence estimates between the functional impairment groups to be notable.

In the second part of our analysis, the association of baseline functional impairment and depression subgroup was examined by fitting LTA models with categorical WSAS scores as a grouping variable. Models with all parameters freed to vary and the number of statuses ranging from two to seven were fit first. The selection of

the optimal number of statuses was informed by the interpretability of each status and fit statistics such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). The relative fit and parsimony of the models was emphasized in addition to the formal fit statistics because LTA models can have very large degrees of freedom and extreme sparseness, which can skew the fit statistics.[34] Models where measurement invariance was imposed on the item-response probabilities between groups and over time were then explored to see if the depression statuses differed by level of functional impairment. Measurement invariance was formally tested using the G^2 difference test to compare several nested models: 1) models where all parameters were allowed to vary between functional impairment group and from baseline to week 12; 2) models where the parameters were constrained to be equal between the two functional impairment groups but were allowed to vary over time; and 3) models in which the parameters were constrained to be equal between the functional impairment groups and at baseline and week 12.[34] Models 2 and 3 were compared to model 1 using separate G^2 difference tests. In this second phase, we used one LTA model to see if the qualitative nature of the latent depression statuses were different between the baseline impairment level and to see how the transitions in status membership were different for each impairment group. As we constructed the LTA model, we were able to constrain some of the item-response probabilities to reduce sources of heterogeneity and improve the interpretability of the model. Analyses were conducted using PROC LTA[53,88] in SAS 9.3 (SAS Institute, Inc., Cary, NC).

Results

Sixty-nine percent of women had major functional impairment at baseline as measured by the WSAS. Women with major functional impairment at baseline were more likely than women with normal/significant impairment to be younger at depression onset, to have severe depression, and to have lower physical and mental functioning scores, and to have lower quality of life enjoyment and satisfaction (Table 4.1). Both groups had high rates of having a psychiatric comorbidity but women with major functional impairment were more likely to have GAD, PTSD, bulimia, social phobia, psychosis, agoraphobia, and drug abuse/dependence. At the beginning of STAR*D, the majority of both groups of women were likely to be experiencing sleep-onset insomnia, mid-nocturnal insomnia, sad mood, and fatigue (Table 4.2). The women with major impairment were more likely to have sleep-onset insomnia, sad mood, decreased appetite, weight changes, impaired concentration, negative self-view, lack of general interest, fatigue, psychomotor retardation, and psychomotor agitation.

When functional impairment was considered as a grouping variable in the LTA models, a four-status model fit the data best (Supplementary Table 4.1). Formal G^2 difference tests indicated that measurement invariance for all statuses across the impairment groups (G_2^2-G_1^2 =171.34, df = 128, p = 0.006) or across groups and time could not be assumed (G_2^2-G_1^2 =355.28, df = 192, p <0.0001). After careful consideration of the item-response probabilities of the indicator variables that produce the description of each status, however, measurement invariance was imposed on several latent statuses due to qualitative similarities: the Moderate statuses at baseline and week 12 for women in the

normal/significant and major impairment groups; the Severe with Increased Appetite statuses at baseline for women in both impairment groups; the Symptom Resolution statuses at week 12 for both impairment groups; and the Insomnias Only statuses at week 12 for both impairment groups (Table 4.3). Imposing measurement invariance in this way is desirable in aiding model fitting and enhancing the interpretability of the statuses.

For the women in the normal/significant impairment group, the statuses at baseline were Mild Depression, Moderate Depression, Severe Depression with Increased Appetite, and Severe Depression with Insomnias (Table 4.3). The statuses at week 12 were Symptom Resolution, Mid-Nocturnal Insomnia Only, All Insomnias Only, and Moderate Depression. Mid-nocturnal insomnia and sad mood were the only symptoms highly likely to be endorsed in every status at baseline whereas only sleep-onset insomnia was likely to be endorsed by women in all the statuses at week 12 except for those in the Symptom Resolution status. At baseline, the Mild, Moderate, and Severe with Insomnias statuses were almost all equally most prevalent, with 31-32% of women likely to belong to these statuses (Figure 4.1). With a prevalence of 5%, Severe with Increased Appetite was the least prevalent at baseline. The majority of women (67%) were likely to be in the Symptom Resolution status after treatment. The fewest women (8%) were likely to belong to the Moderate status. Women in the Moderate status at baseline moved to the Symptom Resolution status at week 12. The women in the Severe with Insomnias status were the least likely to move to the Symptom Resolution status (32% chance). These women were more likely to transition to the All Insomnias Only status (34% chance).

For the women with major functional impairment, the statuses at baseline were Moderate Depression, Severe Depression with Decreased Appetite, Severe Depression with Increased Appetite, and Severe Depression with Psychomotor Agitation (Table 4.3). The statuses at week 12 were Symptom Resolution, Depression with All Insomnias Only, Moderate Depression, and Severe Depression with Psychomotor Disturbances. Women in all statuses were likely to be experiencing mid-nocturnal insomnia, sad mood, and fatigue at baseline. Sleep-onset insomnia and mid-nocturnal insomnia were likely to be endorsed by all statuses except Symptom Resolution at week 12. The Severe with Decreased Appetite status was the most common at baseline, with a prevalence of 36% (Figure 4.2). At 12% of women, the Severe with Psychomotor Agitation status was the least prevalent. These women had the lowest chance of transitioning to the Symptom Resolution status at week 12 (4%) and were most likely to move to the All Insomnias Only status (62% chance). Those in the Moderate status had the greatest chance (86%) of transitioning to the Symptom Resolution status. The majority of women were likely to belong to the Symptom Resolution status (60%) at week 12 while only 5% were likely to be in Severe with Psychomotor Disturbances. Women in the statuses distinguished by appetite changes were the only ones with a chance of transitioning to Severe with Psychomotor Disturbances, the only status not distinguished by having fewer prominent symptoms, at 13% and 5%, respectively.

Discussion

The overall goal of this study was to explore how latent depression statuses differ by baseline functional impairment and to describe how the qualitative nature of these

depression statuses differed by baseline level of functional impairment for women being treated with citalopram. Using level of functional impairment in a multiple-group LTA model demonstrated that depression types for women in STAR*D differed by level of impairment and that baseline impairment influenced changes in depression type during citalopram treatment. Women with major functional impairment at baseline had more Severe Depression statuses at both time points when compared to women with normal/significant functional impairment. This is expected since greater depressive symptom severity was seen to be correlated with lower functioning in a previous analysis of a subset of men and women participating in STAR*D.[114]

The types of depression experienced by women in both functional impairment groups were similar in a few ways but the statuses for those with major impairment were characterized by more severe depression throughout the study. The Moderate, Severe with Increased Appetite, Symptom Resolution, and All Insomnias Only statuses were common to both groups but the women who started level 1 of STAR*D with major functional impairment had more Severe depression statuses at both times. The Severe statuses for these women were marked by prominent symptoms related to decreased and increased appetite and psychomotor disturbances. The women with normal/significant impairment had Severe statuses distinguished by increased appetite and insomnias and these Severe statuses were only present at baseline. Beyond differences in the descriptive nature of the Severe statuses, the prevalences of these statuses also differed by functional impairment group, with more than three times as many women with major impairment likely to be in the Severe with Increased Appetite baseline than in the mild/significant

functional impairment group. Women with baseline major impairment also had lower probabilities of transitioning to a status differentiated by endorsing fewer symptoms when compared to women with baseline normal/significant impairment.

It is not surprising that the latent depression statuses for women with major functional impairment would be characterized by the endorsement of more depression symptoms than those of the women with normal/significant functional impairment since greater depression symptom severity has been seen to be related to reduced quality of life and functioning in all STAR*D participants.[115,116] Additionally, women with major functional impairment in this analysis had more comorbid anxiety disorders than women with normal/significant impairment and comorbid anxiety disorders appear to increase the risk of low health-related quality of life across numerous domains for women.[117] Rates of improvement in quality of life are lower for people with chronic major depression and those with comorbid psychiatric disorders. Furthermore, specific anxiety disorders may differentially impact domains of functioning, e.g., social phobia would impair social functioning,[111] but examining individual psychiatric comorbidities was beyond the scope of this analysis.

Both groups of women experienced depression statuses distinguished by combinations of insomnia symptoms at both baseline and week 12. Almost a third of the women with normal/significant impairment were in the Severe with Insomnias status at baseline. These women had the lowest chances of transitioning to the Symptom Resolution status after treatment. While the women with major impairment did not have a baseline depression status distinguished by insomnia, almost a quarter of them had

moved into the All Insomnias Only status at week 12. The prominence of insomnia is in line with the associations between insomnia, sleepiness, fatigue and functioning in depression.[102,109] These insomnia symptoms likely warrant further attention in treatment approaches for depression since they are common residual symptoms[44] and treatment side effects. Insomnia also increases risk of depression recurrence[57,65] and lack of treatment response.[66]

Our results should be interpreted in the context of several limitations. This was a post-hoc analysis of clinical trial data that were not originally collected for such subgroup analyses. These LTA models were unable to address all of the known correlates of reduced functioning in depression because of issues of model convergence when trying to fit complex latent variable models. Age, race, education, marital status, employment status, medical comorbidities, and health insurance coverage have been observed to be related to baseline functioning in a separate subsample of STAR*D participants.[114] This analysis also only examined functioning as captured by the WSAS, which does not cover all domains of health-related quality of life. Furthermore, 21 women were excluded from this analysis because they did not complete the IVR call during which the functioning assessments were completed and these women might differ on level of functioning compared to the women who were able to complete the call. Lastly, this analysis focused only on functional impairment observed at baseline. It may be important to consider functioning longitudinally since improvements in functioning have been demonstrated with antidepressant treatment and such improvements can lag behind symptom improvements.[118]

Despite these limitations, our analysis is noteworthy for several reasons. This is one of the first analyses to use LTA to examine how functional impairment is related to depression subtypes and changes in these subtypes following antidepressant treatment. LTA allowed us to efficiently discern depression subtypes among women during a clinical trial and to examine the association between functional impairment and changes in depression symptoms following treatment with citalopram. Although data sparseness limited our ability to examine some factors of potential interest, this is still one of the largest samples to which LTA has been applied. While methods for power calculations in LTA are still being developed, it has been suggested that sample sizes of 300 people or more are sufficient.[119] Although STAR*D enrolled treatment-seeking outpatients and thus our results have limited generalizability, it is still the largest and longest depression treatment study and is considered more representative of people with depression who are seen outside of idealized research setting than most trials.

Our results highlight the importance of looking beyond summary rating scores of depression symptoms when studying depression heterogeneity during treatment. For women in STAR*D, level of functional impairment was related to the likelihood of moving to a depression status differentiated by endorsing fewer symptoms and thus treatment strategies may want to consider not only symptom severity but also degree of functional impairment. Assessment and treatment of major depression should not focus exclusively on symptoms but also incorporate domains of functioning. Doing so could reduce the substantial disability and burden associated with depression.

Table 4.1: Baseline demographic and clinical characteristics of women participating in STAR*D level 1 by baseline functional impairment

Characteristic	Normal/significant functional impairment[1] (n = 231)	Major functional impairment[2] (n = 524)	*p*-value
Age at study entry, mean (SD)	42.5 (14.1)	41.0 (12.7)	0.141
45 years of age or older at study entry, n (%)	100 (43.3)	215 (41.0)	0.562
Age at onset, mean (SD)	27.5 (15.5)	23.6 (14.0)	<0.001
Race, n (%)			
White	184 (79.7)	397 (75.8)	0.227
Black or African American	28 (12.1)	89 (17.0)	
Other	19 (8.2)	38 (7.3)	
Hispanic, n (%)	33 (14.3)	70 (13.4)	0.732
Number of depressive episodes before baseline, mean (SD)	4.3 (6.7)	4.6 (6.3)	0.581
Depression severity (QIDS-SR_{16}), mean (SD)	13.6 (3.7)	17.2 (3.6)	<0.001
SF-12 PCS total score,[3] mean (SD)	52.6 (10.8)	48.4 (12.0)	<0.001
SF-12 MCS total score,[4] mean (SD)	28.8 (8.7)	24.0 (7.1)	<0.001
Q-LES-Q total score,[5] mean (SD)	50.8 (11.6)	36.8 (12.3)	<0.001
Psychiatric comorbidity, n (%)			
Any psychiatric comorbidity	115 (49.8)	349 (66.6)	<0.001
Generalized anxiety disorder	30 (13.1)	141 (27.2)	<0.001
Post-traumatic stress disorder	17 (7.5)	98 (19.1)	<0.001
Bulimia	28 (12.2)	94 (18.2)	0.042
Social phobia	45 (20.0)	166 (32.2)	<0.001
OCD	13 (5.7)	49 (9.5)	0.084
Panic disorder	13 (5.7)	76 (14.6)	<0.001
Psychosis	17 (7.4)	70 (13.5)	0.017
Agoraphobia	9 (4.0)	66 (12.9)	<0.001
Alcohol abuse/dependence	18 (7.9)	47 (9.0)	0.600
Drug abuse/dependence	3 (1.3)	31 (6.0)	0.005
Somatization disorder	3 (1.3)	16 (3.1)	0.154
Hypochondriasis	10 (4.4)	21 (4.1)	0.843

[1]Normal/significant functional impairment = WSAS total score ≤ 20 at baseline.
[2]Major functional impairment = WSAS total score ≥ 21 at baseline.
[3]Short Form-12 Health Survey Physical Component Summary. A higher score indicates better functioning.
[4]Short Form-12 Health Survey Mental Component Summary. A higher score indicates better functioning.
[5]Quality of Life Enjoyment and Satisfaction Questionnaire. A higher score indicates better quality of life.

Table 4.2: Frequency of baseline QIDS-SR_{16} indicators by baseline functional impairment for women participating in STAR*D level 1

QIDS-SR_{16} item	**Normal/ significant functional impairment[1] (n = 231) N (%)**	**Major functional impairment[2] (n = 524) N (%)**	***p*-value**
Sleep-onset insomnia	134 (58.0)	354 (67.6)	0.011
Mid-nocturnal insomnia	162 (70.1)	379 (72.3)	0.537
Early morning insomnia	100 (43.3)	265 (50.6)	0.065
Hypersomnia	26 (11.3)	88 (16.8)	0.050
Sad mood	169 (73.2)	472 (90.1)	<0.001
Decreased appetite	35 (15.2)	130 (24.8)	0.003
Increased appetite	35 (15.2)	111 (21.2)	0.053
Decreased weight	18 (7.8)	71 (13.6)	0.024
Increased weight	15 (6.5)	77 (14.7)	0.002
Impaired concentration	89 (38.5)	365 (69.7)	<0.001
Negative self-view	95 (41.1)	322 (61.5)	<0.001
Suicidal ideation	21 (9.1)	63 (12.0)	0.238
Lack of general interest	91 (39.4)	353 (67.4)	<0.001
Fatigue	131 (56.7)	432 (82.4)	<0.001
Psychomotor retardation	43 (18.6)	220 (42.0)	<0.001
Psychomotor agitation	54 (23.4)	167 (31.9)	0.018

[1]Normal/significant functional impairment = WSAS total score ≤ 20 at baseline.
[2]Major functional impairment = WSAS total score ≥ 21 at baseline.

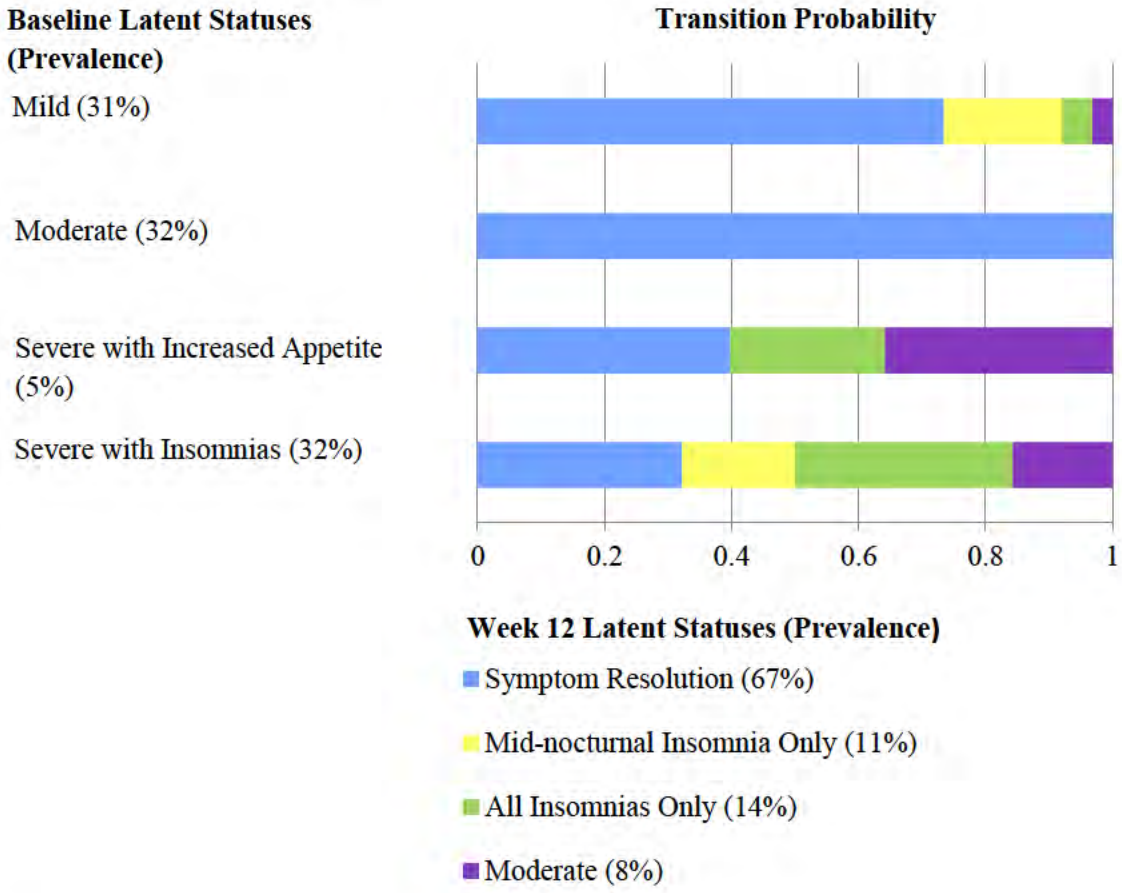
Table 4.3: Item-response probabilities from a four-status LTA of QIDS-SR_{16} indicators with baseline functional impairment as a grouping variable[1]

Normal/Significant Functional Impairment at Baseline[2] (n = 231)								
	Baseline Latent Statuses				Week 12 Latent Statuses			
QIDS-SR_{16} items	**Mild**	**Moderate**	**Severe with Increased Appetite**	**Severe with Insomnias**	**Symptom Resolution**	**Mid-Nocturnal Insomnia Only**	**All Insomnias Only**	**Moderate**
Sleep onset insomnia	0.41	0.45	0.63	0.81	0.12	0.23	0.59	0.45
Mid-nocturnal insomnia	0.78	0.59	0.69	0.77	0.46	1.00	0.76	0.59
Early morning insomnia	0.39	0.28	0.53	0.65	0.05	0.40	0.50	0.28
Hypersomnia	0.00	0.22	0.27	0.08	0.06	0.00	0.05	0.22
Sad mood	0.53	0.74	0.99	0.89	0.04	0.00	0.26	0.74
Decreased appetite	0.03	0.10	0.00	0.37	0.01	0.05	0.08	0.10
Increased appetite	0.07	0.14	0.85	0.10	0.03	0.02	0.10	0.14
Decreased weight	0.00	0.05	0.00	0.18	0.01	0.08	0.07	0.05
Increased weight	0.05	0.05	0.54	0.01	0.01	0.31	0.07	0.05
Impaired concentration	0.07	0.53	0.78	0.49	0.02	0.00	0.24	0.53
Negative self-view	0.14	0.43	0.73	0.64	0.03	0.05	0.24	0.43
Suicidal ideation	0.05	0.07	0.13	0.17	0.01	0.00	0.07	0.07
Lack of general interest	0.09	0.50	0.83	0.57	0.06	0.00	0.28	0.50
Fatigue	0.28	0.72	0.92	0.62	0.06	0.00	0.31	0.72
Psychomotor retardation	0.00	0.24	0.56	0.30	0.00	0.03	0.18	0.24
Psychomotor agitation	0.09	0.15	0.25	0.45	0.02	0.00	0.30	0.15

Major Functional Impairment at Baseline[3] (n = 524)								
	Baseline Latent Statuses				Week 12 Latent Statuses			
QIDS-SR$_{16}$ items	**Moderate**	**Severe with Decreased Appetite**	**Severe with Increased Appetite**	**Severe with Psychomotor Agitation**	**Symptom Resolution**	**All Insomnias Only**	**Moderate**	**Severe with Psychomotor Disturbances**
Sleep onset insomnia	0.45	0.82	0.63	1.00	0.12	0.59	0.45	0.67
Mid-nocturnal insomnia	0.59	0.81	0.69	0.83	0.46	0.76	0.59	0.82
Early morning insomnia	0.28	0.61	0.53	0.72	0.05	0.50	0.28	0.53
Hypersomnia	0.22	0.14	0.27	0.00	0.06	0.05	0.22	0.24
Sad mood	0.74	0.96	0.99	0.95	0.04	0.26	0.74	0.94
Decreased appetite	0.10	0.54	0.00	0.14	0.01	0.08	0.10	0.15
Increased appetite	0.14	0.00	0.85	0.09	0.03	0.10	0.14	0.25
Decreased weight	0.05	0.31	0.00	0.06	0.01	0.07	0.05	0.07
Increased weight	0.05	0.04	0.54	0.12	0.01	0.07	0.05	0.15
Impaired concentration	0.53	0.85	0.78	0.50	0.02	0.24	0.53	0.92
Negative self-view	0.43	0.75	0.73	0.47	0.03	0.24	0.43	1.00
Suicidal ideation	0.07	0.17	0.13	0.09	0.01	0.07	0.07	0.28
Lack of general interest	0.50	0.88	0.83	0.31	0.06	0.28	0.50	0.85
Fatigue	0.72	0.93	0.92	0.65	0.06	0.31	0.72	0.97
Psychomotor retardation	0.24	0.56	0.56	0.23	0.00	0.18	0.24	0.70
Psychomotor agitation	0.15	0.44	0.25	0.55	0.02	0.30	0.15	0.51

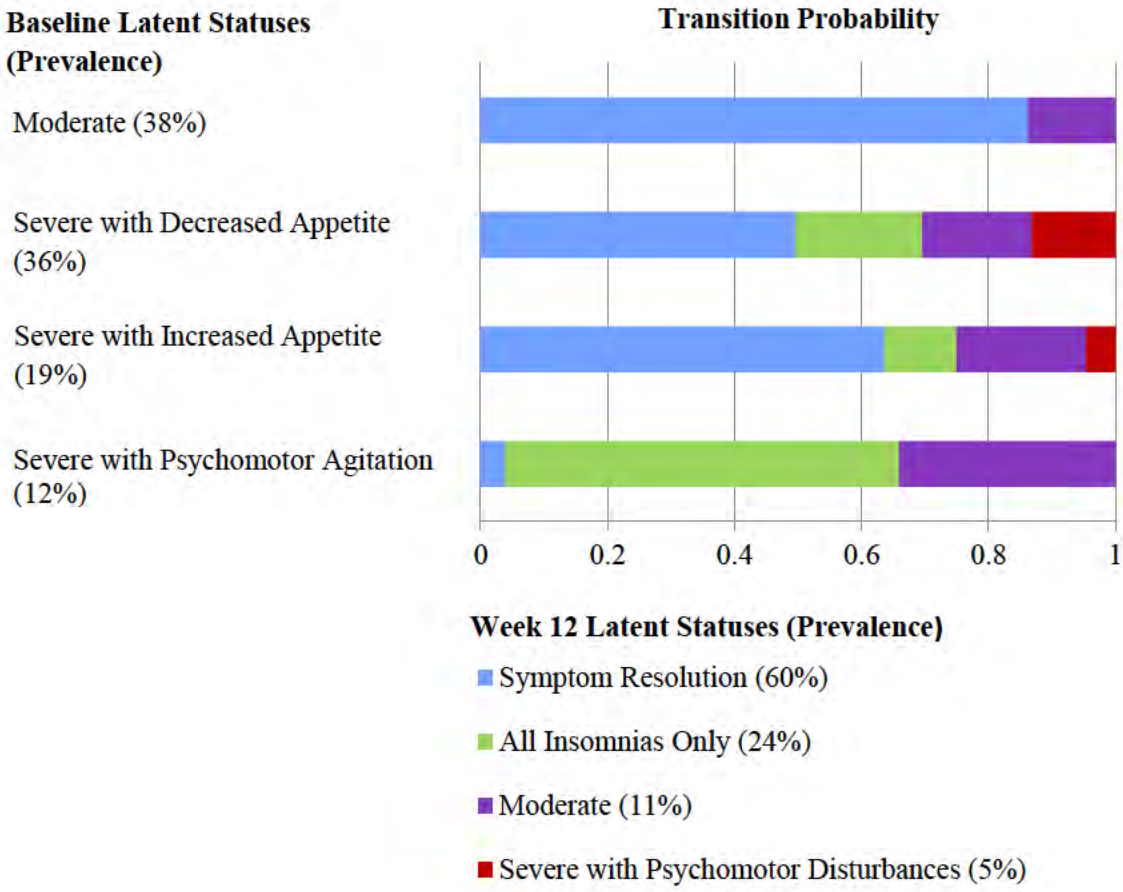
[1]Measurement invariance was imposed on the item-response probabilities describing the Moderate statuses; the Severe with Increased Appetite statuses; the Symptom Resolution statuses; and the Insomnias Only statuses.
[2]Normal/significant functional impairment = WSAS total score ≤ 20 at baseline.
[3]Major functional impairment = WSAS total score ≥ 21 at baseline.

Figure 4.1: Latent status prevalences and probabilities of transitioning in depression status membership from baseline to week 12 for women with baseline normal/significant functional impairment[1]



[1]Measurement invariance was imposed on the item-response probabilities describing the Moderate statuses; the Severe with Increased Appetite statuses; the Symptom Resolution statuses; and the Insomnias Only statuses.

Figure 4.2: Latent status prevalences and probabilities of transitioning in depression status membership from baseline to week 12 for women with baseline major functional impairment[1]



[1]Measurement invariance was imposed on the item-response probabilities describing the Moderate statuses; the Severe with Increased Appetite statuses; the Symptom Resolution statuses; and the Insomnias Only statuses.

Supplementary Table 4.1: Fit statistics for LTA models with baseline WSAS as a grouping variable

# of Statuses	*df*	AIC	BIC	G^2
Basic Model Without Measurement Invariance				
2	8589934459	11585.98	12205.96	11317.98
3	8589934389	11354.67	12317.03	10938.67
4	8589934317	11316.96	12640.20	10744.96
5	8589934592	11313.55	13016.19	10577.55
6	8589934167	11364.18	13464.71	10456.18
7	8589934089	11333.55	13850.49	10245.55
Basic Model with Measurement Invariance on Rho Parameters Across Groups but Not Across Time				
2	8589934523	11615.35	11939.22	11475.35
3	8589934485	11303.47	11821.66	11079.47
4	8589934445	11241.45	11972.47	10925.45
5	8589934403	11246.21	12208.57	10830.21
6	8589934359	11216.49	12428.69	10692.49
7	8589934313	11237.99	12718.54	10597.99
Basic Model with Measurement Invariance on Rho Parameters Across Groups and Time				
2	8589934555	11950.68	12126.50	11874.68
3	8589934533	11550.46	11846.57	11422.46
4	8589934509	11308.87	11743.78	11120.87
5	8589934483	11261.63	11853.85	11005.63
6	8589934455	11215.17	11983.21	10883.17
7	8589934425	11253.14	12215.50	10837.14

CHAPTER V

DISCUSSION AND CONCLUSIONS

The overall purpose of this dissertation was to examine latent variable methods for understanding heterogeneity in major depression. Major depression is based on a phenomenological diagnosis and is associated with extensive variability in etiology, risk factors, and symptom profiles. Despite more than five decades of antidepressant development, medications and psychotherapies fail to help a considerable amount of people with depression.[120] While a substantial amount of previous research has tried to address this heterogeneity with subgroup analyses and various depression subtypes have been proposed, the longitudinal stability of subgroups is rarely evaluated and debate still exists regarding how to successfully treat most people with depression.

Although much still needs to be learned about the pathophysiological mechanisms of the brain responsible for depression,[121] the limitations of traditional methods for subgroup analyses might also be hindering progress towards precision medicine for major depression and other serious mental illnesses. These limitations include the issue of high Type I error due to multiple comparisons and lack of adequate sample sizes and power to detect effects across subgroups or to examine higher-order interactions in subgroups.[35] The work of this dissertation shows that latent class analysis (LCA) and latent transition analysis (LTA) are valuable alternative methods for elucidating depression subtypes based on the many possible patterns of depression symptoms and for examining changes in these symptom patterns after antidepressant treatment in level 1 of STAR*D. Conclusions from the three specific aims of this dissertations and implications for future work in this area are summarized below.

Differences in Latent Depression Subtypes between Men and Women

When latent class analysis was conducted at baseline in Aim 1 (Chapter II), four latent classes of depression were ultimately identified that were consistent for men and women. Men and women experienced the same patterns of depression symptoms but the proportions of men and women likely to be in the Mild Depression, Severe Depression with Increased Appetite, and Severe Depression with Insomnia classes differed. More women were likely to belong to the Severe with Insomnia class than any other class but the most prevalent class for men was the Mild Depression class. The difference between the prevalence of men and women in the Severe with Increased Appetite class prevalence was remarkable, with almost twice as many women in this class. The associations between comorbid anxiety disorders and the odds of membership in each latent class were similar for men and women.

While the qualitative nature of the depression subtypes were the same for men and women when only baseline data was examined, the latent transition analyses performed in Aim 2 (Chapter III) demonstrated that the types of depression experienced by men and women differed when both baseline and week 12 symptoms were included. The latent depression statuses for both men and women were distinguished primarily by severity but psychomotor agitation and retardation further differentiated men's latent depression statuses. Appetite changes, insomnia, and psychomotor disturbances characterized statuses for women. After 12 weeks of citalopram treatment, transition to Symptom Resolution$_{Men}$ was most likely for men in the Mild$_{Men}$ status at baseline but least likely for Men in the Severe Depression with Psychomotor Agitation status. Among

women, those in the $Moderate_{Women}$ status had the greatest chance of moving to Symptom $Resolution_{Women}$ while those in the Severe with Decreased Appetite status had the lowest chance of transitioning to Symptom $Resolution_{Women}$.

The differences between patterns of depression symptoms and changes in these patterns experienced by men and women revealed here emphasize the need to consider the influence on sex and gender when studying and treating depression. [122] Disparities in rates, severity, and course of depression between men and women have previously been observed but the extent to which these differences are due to sex influences (e.g., women's hormone changes precipitating insomnia) or to gender influences (e.g., women being more likely to seek depression treatment) remains unclear.

The Role of Comorbid Psychiatric Disorders

Comorbid psychiatric disorders, especially anxiety disorders, were related to latent depression subgroups in both LCA and LTA models. In Aim I (Chapter II), comorbid GAD, bulimia, and social phobia were related to increased odds of membership in the Severe Depression with the Increased Appetite class at baseline. GAD, PTSD, and social phobia were associated with the Severe Depression with Insomnia class. It was seen in Aim 3 (Chapter IV) that women with baseline major functional impairment were more likely than women with normal/significant impairment to have a psychiatric comorbidity. Low prevalences of some disorders in this sample and missing data reduced our ability to examine some individual psychiatric comorbidities in the LCA and LTA models.

Comorbidity rates between depression and anxiety are generally very high, with 75% of people with a current depressive disorder having a lifetime comorbid anxiety disorder and 67% having a current anxiety disorder.[123] The presence of anxiety with depression is also associated with lower odds of remission during treatment and with delayed treatment response.[124] Given this, the apparent influence of psychiatric comorbidities on depression in this dissertation work is not surprising. What remains unknown, however, is the role of these comorbidities in changes in depression subtypes over time. A latent transition analysis examining individual comorbidities in predicting latent depression statuses at baseline and predicting transitions in statuses throughout treatment would be valuable.

Functional Impairment and Depression Subtypes for Women

Examining LTA models by functional impairment in Aim 3 (Chapter IV) showed that the degree of functional impairment experienced by women at baseline was related to both depression subtypes at baseline and the chances of transitioning to a different depression status at week 12. Almost all women started level 1 of STAR*D with some degree of functional impairment, as assessed by the Work and Social Adjustment Scale. The majority of these women were experiencing major functional impairment at baseline. The multiple-groups LTA model, which is similar to performing stratification in non-latent variable analyses, indicated that some depression statuses differed between the women with baseline major impairment and the women with normal/significant impairment. In particular, unlike those with normal/significant impairment, the women with major functional impairment could belong to statuses at baseline and week 12 that

were distinguished by psychomotor disturbances. The majority of women in both baseline functional impairment groups were likely to transition to the Symptom Resolution status at week 12 but those with baseline major impairment who started in the Severe Depression with Psychomotor Agitation at the beginning of the study were least likely to transition to the Symptom Resolution status.

These results underscore the need to incorporate not only depression symptoms but also functioning in the assessment and treatment of major depression. Sample size issues restricted the LTA models to women and, although women with depression appear to have more functional impairment than men, it would be useful to replicate these analyses in a sample of men since men and women have been observed to experience impairments in different domains of functioning. Additionally, future research on how functional impairment changes during treatment and how it predicts transitions in depression subtypes is warranted.

The Problem with Relying on Summary Rating Scores of Symptom Severity

The results of this dissertation highlight the need to not collapse information about individual depression symptoms and rely solely on summary rating scores of symptom severity when studying and treating depression. Rates of response and remission have been seen to vary in the same population when cut-off scores on different rating scales are used to define treatment response and remission, making it difficult to know when patients are experiencing satisfactory relief from their depression.[125] Furthermore, when exploring predictors and moderators of treatment effects, it is important to consider domains beyond depression symptoms that were not available in

the data analyzed here, such as cognitive measures,[126] that have potential to differentiate people who remit from those who do not. Doing so will become even more important as the National Institute of Mental Health emphasizes clinical trials that focus on the biological processes involved in a psychiatric disease and not only on ameliorating symptoms.[127] Efforts to reduce the burden of depression should also incorporate measures of functional impairment in both predicting treatment effects and in defining response and remission so that successful treatment improves not just symptoms but overall quality of life.

This dissertation demonstrates the potential of LCA and LTA as approaches that can be used to characterize discrete changes among multiple aspects of a disease. Such approaches are particularly important in psychiatry, which is haunted by the absence of known pathophysiologic causes of disorders such as depression[128] and where disorders are currently defined by heterogeneous symptoms which are likely not disease-specific. When the National Comorbidity Survey-Replication (NCS-R) considered 19 different possible DSM diagnoses, only 433 of the 524,288 logically possible disorder profiles were observed.[2] Almost 80% of these observed profiles involved comorbid cases of three or more disorders. Determining specific features that can distinguish groups of people with depression that have differential responses to treatment could ultimately aid in clinical treatment decision-making and alleviate the burden of depression quicker than the current system of selecting the best treatment through trial and error. The results presented here need to be replicated in other samples and for other treatments. Doing so could ultimately provide information about predicting treatment remission.

Beyond improving subgroup detection to inform treatment response in major depression, LCA and LTA should be explored as analytic techniques for informing new approaches to classifying mental disorders. These person-centered statistical methods could be valuable for making sense of the myriad potential biological and psychosocial indicators of serious mental illnesses, as emphasized by the National Institute of Mental Health's Research Domain Criteria initiative.

REFERENCES

1. Whiteford HA, Degenhardt L, Rehm J, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. *Lancet*. 2013;382(9904):1575–86.

2. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. *Arch Gen Psychiatry*. 2005;62(6):617–27.

3. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 4th Edition.Washington, DC: American Psychiatric Association; 1994.

4. Greenberg PE, Fournier A-A, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry*. 2015;76(2):155–162.

5. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*. 2006;3(11):e442.

6. Wang PS, Insel TR. NIMH-funded pragmatic trials: Moving on. *Neuropsychopharmacology*. 2010;35(13):2489–90.

7. National Center for Health Statistics. *Health, United States, 2010: With Special Feature on Death and Dying*. Hyattsville, MD; 2011.

8. 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*. 2005;62(6):629–40.

9. Gaynes BN, Warden D, Trivedi MH, Wisniewski SR, Fava M, Rush AJ. What did STAR*D teach us? Results from a large-scale, practical, clinical trial for patients with depression. *Psychiatr Serv*. 2009;60(11):1439–45.

10. 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*. 2001;58(4):395–401.

11. Simon GE, Perlis RH. Personalized medicine for depression: can we match patients with treatments? *Am J Psychiatry*.2010;167(12):1445–55.

12. Ostergaard SD, Jensen SOW, Bech P. The heterogeneity of the depressive syndrome: when numbers get serious. *Acta Psychiatr Scand*. 2011;124(6):495–6.

13. Rush AJ. The varied clinical presentations of major depressive disorder. *J Clin Psychiatry*. 2007;68 Suppl 8:4–10.

14. Substance Abuse and Mental Health Services Administration. *Results from the 2012 National Survey on Drug Use and Health: Summary of National Findings*. Rockville, MD; 2013.

15. Kessler RC. Epidemiology of women and depression. *J Affect Disord*. 2003;74(1):5–13.

16. Alexopoulos GS. Depression in the elderly. *Lancet*. 2005;365(9475):1961–70.

17. Substance Abuse and Mental Health Services Administration. *Results from the 2012 National Survey on Drug Use and Health: Mental Health Findings*. Rockville, MD; 2013. Available at: http://www.samhsa.gov/data/NSDUH/2k12MH_FindingsandDetTables/2K12MHF/NSDUHmhfr2012.htm.

18. Kendler KS, Gardner CO. Sex differences in the pathways to major depression: a study of opposite-sex twin pairs. *Am J Psychiatry*. 2014;171(4):426–35.

19. Young EA, Kornstein SG, Marcus SM, et al. Sex differences in response to citalopram: a STAR*D report. *J Psychiatr Res*. 2009;43(5):503–11.

20. Marcus SM, Kerber KB, Rush AJ, et al. Sex differences in depression symptoms in treatment-seeking adults: confirmatory analyses from the Sequenced Treatment Alternatives to Relieve Depression study. *Compr Psychiatry*. 2008;49(3):238–46.

21. Silverstein B. Gender differences in the prevalence of somatic versus pure depression: a replication. *Am J Psychiatry*. 2002;159(6):1051–2.

22. 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*. 2010;71(12):1582–9.

23. Schuch JJJ, Roest AM, Nolen WA, Penninx BWJH, de Jonge P. Gender differences in major depressive disorder: results from the Netherlands study of depression and anxiety. *J Affect Disord*. 2014;156:156–63.

24. Gili M, Castro A, Navarro C, et al. Gender differences on functioning in depressive patients. *J Affect Disord*. 2014;166:292–6.

25. Harald B, Gordon P. Meta-review of depressive subtyping models. *J Affect Disord*. 2012;139(2):126–40.

26. Arnow BA, Blasey C, Williams LM, et al. Depression Subtypes in Predicting Antidepressant Response: A Report From the iSPOT-D Trial. *Am J Psychiatry*. 2015:appi.ajp.2015.1.

27. National Institute of Mental Health, National Institutes of Health. *Strategic Plan for Research*. Bethesda, MD; 2013. Available at: http://www.nimh.nih.gov/about/strategic-planning-reports/index.shtml.

28. National Advisory Mental Health Council Workgroup. *From Discovery to Cure: Accelerating the Development of New and Personalized Interventions for Mental Illnesses - Report of the National Advisory Mental Health Council's Workgroup*.; 2010. Available at: http://www.nimh.nih.gov/about/advisory-boards-and-groups/namhc/reports/fromdiscoverytocure.pdf.

29. Oquendo MA, Barrera A, Ellis SP, et al. Instability of symptoms in recurrent major depression: a prospective study. *Am J Psychiatry*. 2004;161(2):255–61.

30. Fournier JC, DeRubeis RJ, Hollon SD, Gallop R, Shelton RC, Amsterdam JD. Differential change in specific depressive symptoms during antidepressant medication or cognitive therapy. *Behav Res Ther*. 2013;51(7):392–8.

31. Angst J, Gamma A, Benazzi F, Ajdacic V, Rössler W. Melancholia and atypical depression in the Zurich study: epidemiology, clinical characteristics, course, comorbidity and personality. *Acta Psychiatr Scand Suppl*. 2007;(433):72–84.

32. Insel TR. The NIMH Research Domain Criteria (RDoC) Project: Precision medicine for psychiatry. *Am J Psychiatry*. 2014;171(4):395–7.

33. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. *Am J Psychiatry*. 2006;163(1):28–40.

34. Collins LM, Lanza ST. *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*. Hoboken, NJ: John Wiley & Sons; 2010.

35. Lanza ST, Rhoades BL. Latent Class Analysis: An Alternative Perspective on Subgroup Analysis in Prevention and Treatment. *Prev Sci*. 2011;14(2):157–68.

36. Jiang Y, Zack MM. A Latent Class Modeling Approach to Evaluate Behavioral Risk Factors and Health-Related Quality of Life. *Prev Chronic Dis*. 2011;8(6):A137.

37. Ryoo JH, Wu C, McCormick C. Characterizing changing classifications: Practical illustrations of latent transition analysis (LTA). In: Nebraska Center of Research on Children, Youth, Families and Schools (CYFS). Available at: http://r2ed.unl.edu/presentations/2012/SRM/033012_RyooWu/033012_RyooWu.pdf. Accessed April 1, 2015.

38. Kessler RC, Demler O, Frank RG, et al. Prevalence and treatment of mental disorders, 1990 to 2003. *N Engl J Med*. 2005;352(24):2515–23.

39. Kendler KS, Eaves LJ, Walters EE, Neale MC, Heath AC, Kessler RC. The identification and validation of distinct depressive syndromes in a population-based sample of female twins. *Arch Gen Psychiatry*. 1996;53(5):391–9..

40. 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*. 2009;166(5):599–607.

41. Van der Lem R, van der Wee NJA, van Veen T, Zitman FG. The generalizability of antidepressant efficacy trials to routine psychiatric out-patient practice. *Psychol Med*. 2011;41(7):1353–63.

42. Rush AJ, Fava M, Wisniewski SR, et al. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. *Control Clin Trials*. 2004;25(1):119–42.

43. Nierenberg AA, Fava M, Trivedi MH, et al. A comparison of lithium and T(3) augmentation following two failed medication treatments for depression: a STAR*D report. *Am J Psychiatry*. 2006;163(9):1519–30; quiz 1665.

44. Nierenberg AA, Husain MM, Trivedi MH, et al. Residual symptoms after remission of major depressive disorder with citalopram and risk of relapse: a STAR*D report. *Psychol Med*. 2010;40(1):41–50.

45. Trivedi MH, Fava M, Wisniewski SR, et al. Medication augmentation after the failure of SSRIs for depression. *N Engl J Med*. 2006;354(12):1243–52.

46. Rush AJ, Trivedi MH, Ibrahim HM, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry*. 2003;54(5):573–83.

47. Bühler J, Seemüller F, Läge D. The predictive power of subgroups: An empirical approach to identify depressive symptom patterns that predict response to treatment. *J Affect Disord*. 2014;163:81–7.

48. Zimmerman M. The Psychiatric Diagnostic Screening Questionnaire Manual. 2002.

49. Rush AJ, Zimmerman M, Wisniewski SR, et al. Comorbid psychiatric disorders in depressed outpatients: demographic and clinical features. *J Affect Disord*. 2005;87(1):43–55.

50. Zimmerman M, Mattia JI. A self-report scale to help make psychiatric diagnoses: the Psychiatric Diagnostic Screening Questionnaire. *Arch Gen Psychiatry*. 2001;58(8):787–94.

51. Akaike H. Factor analysis and AIC. *Psychometrika*. 1987;52:317–332.

52. Schwartz G. Estimating the dimension of a model. *Ann Stat*. 1978;6:461–464.

53. Lanza ST, Dziak JJ, Huang L, Wagner A, Collins LM. PROC LCA & PROC LTA Users' Guide. 2014. Available at: http://methodology.psu.edu/.

54. Lanza ST, Collins LM, Lemmon DR, Schafer JL. PROC LCA: A SAS Procedure for Latent Class Analysis. *Struct Equ Modeling*. 2007;14(4):671–694.

55. Muthén L, Muthén B. Mplus User's Guide. Seventh Edition. Available at: http://www.statmodel.com/download/usersguide/Mplus user guide Ver_7_r6_web.pdf.

56. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. 1960;23:56–62.

57. Ohayon MM, Roth T. Place of chronic insomnia in the course of depressive and anxiety disorders. *J Psychiatr Res*. 2003;37(1):9–15.

58. Ellis JG, Perlis ML, Bastien CH, Gardani M, Espie CA. The natural history of insomnia: acute insomnia and first-onset depression. *Sleep*. 2014;37(1):97–106.

59. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. Washington, DC: American Psychiatric Association; 2013.

60. Krishnan V, Collop NA. Gender differences in sleep disorders. *Curr Opin Pulm Med*. 2006;12(6):383–9.

61. Taylor DJ, Lichstein KL, Durrence HH, Reidel BW, Bush AJ. Epidemiology of insomnia, depression, and anxiety. *Sleep*. 2005;28(11):1457–64.

62. Roth T, Jaeger S, Jin R, Kalsekar A, Stang PE, Kessler RC. Sleep problems, comorbid mental disorders, and role functioning in the national comorbidity survey replication. *Biol Psychiatry*. 2006;60(12):1364–71.

63. Sivertsen B, Krokstad S, Øverland S, Mykletun A. The epidemiology of insomnia: associations with physical and mental health. The HUNT-2 study. *J Psychosom Res*. 2009;67(2):109–16.

64. Manber R, Edinger JD, Gress JL, San Pedro-Salcedo MG, Kuo TF, Kalista T. Cognitive behavioral therapy for insomnia enhances depression outcome in patients with comorbid major depressive disorder and insomnia. *Sleep*. 2008;31(4):489–95.

65. Perlis ML, Giles DE, Buysse DJ, Tu X, Kupfer DJ. Self-reported sleep disturbance as a prodromal symptom in recurrent depression. *J Affect Disord*. 1997;42(2-3):209–12.

66. Emslie GJ, Kennard BD, Mayes TL, et al. Insomnia moderates outcome of serotonin-selective reuptake inhibitor treatment in depressed youth. *J Child Adolesc Psychopharmacol*. 2012;22(1):21–8.

67. Konttinen H, Kiviruusu O, Huurre T, Haukkala A, Aro H, Marttunen M. Longitudinal associations between depressive symptoms and body mass index in a 20-year follow-up. *Int J Obes*. 2014;38(5):668–74.

68. Luppino FS, de Wit LM, Bouvy PF, et al. Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Arch Gen Psychiatry*. 2010;67(3):220–9.

69. Khambaty T, Stewart JC, Muldoon MF, Kamarck TW. Depressive symptom clusters as predictors of 6-year increases in insulin resistance: data from the Pittsburgh Healthy Heart Project. *Psychosom Med*. 2014;76(5):363–9.

70. Ryan JP, Sheu LK, Critchley HD, Gianaros PJ. A neural circuitry linking insulin resistance to depressed mood. *Psychosom Med*. 2012;74(5):476–82.

71. Golden SH, Lazo M, Carnethon M, et al. Examining a bidirectional association between depressive symptoms and diabetes. *JAMA*. 2008;299(23):2751–9.

72. Renn BN, Feliciano L, Segal DL. The bidirectional relationship of depression and diabetes: a systematic review. *Clin Psychol Rev*. 2011;31(8):1239–46.

73. Stetler C, Miller GE. Depression and hypothalamic-pituitary-adrenal activation: a quantitative summary of four decades of research. *Psychosom Med*. 2011;73(2):114–26. doi:10.1097/PSY.0b013e31820ad12b.

74. Zimmermann U, Kraus T, Himmerich H, Schuld A, Pollmächer T. Epidemiology, implications and mechanisms underlying drug-induced weight gain in psychiatric patients. *J Psychiatr Res*. 2003;37(3):193–220.

75. Hryhorczuk C, Sharma S, Fulton SE. Metabolic disturbances connecting obesity and depression. *Front Neurosci*. 2013;7:177.

76. Hashimoto K, Koizumi H, Nakazato M, Shimizu E, Iyo M. Role of brain-derived neurotrophic factor in eating disorders: recent findings and its pathophysiological implications. *Prog Neuropsychopharmacol Biol Psychiatry*. 2005;29(4):499–504.

77. Sullivan PF, Prescott CA, Kendler KS. The subtypes of major depression in a twin registry. *J Affect Disord*. 2002;68(2-3):273–84.

78. Sullivan PF, Kessler RC, Kendler KS. Latent class analysis of lifetime depressive symptoms in the national comorbidity survey. *Am J Psychiatry*. 1998;155(10):1398–406.

79. Garrett ES, Zeger SL. Latent class model diagnosis. *Biometrics*. 2000;56(4):1055–67.

80. Thase ME. Atypical depression: useful concept, but it's time to revise the DSM-IV criteria. *Neuropsychopharmacology*. 2009;34(13):2633–41.

81. Trivedi MH, Rush AJ, Ibrahim HM, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psych. *Psychol Med*. 2004;34(1):73–82.

82. Cameron IM, Crawford JR, Cardy AH, et al. Psychometric properties of the Quick Inventory of Depressive Symptomatology (QIDS-SR) in UK primary care. *J Psychiatr Res*. 2013;47(5):592–8.

83. Finch W, Bronk K. Conducting confirmatory latent class analysis using Mplus. *Struct Equ Model*. 2011;18:132–151.

84. Wurpts IC, Geiser C. Is adding more indicators to a latent class analysis beneficial or detrimental? Results of a Monte-Carlo study. *Front Psychol*. 2014;5:920.

85. Insel T, Cuthbert B, Garvey M, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry*. 2010;167(7):748–51.

86. Lamers F, Rhebergen D, Merikangas KR, de Jonge P, Beekman ATF, Penninx BWJH. Stability and transitions of depressive subtypes over a 2-year follow-up. *Psychol Med*. 2012;42(10):2083–93.

87. Rodgers S, Ajdacic-Gross V, Müller M, et al. The role of sex on stability and change of depression symptom subtypes over 20 years: a latent transition analysis. *Eur Arch Psychiatry Clin Neurosci*. 2014;264(7):577–88.

88. Lanza ST, Collins LM. A new SAS procedure for latent transition analysis: transitions in dating and sexual risk behavior. *Dev Psychol*. 2008;44(2):446–56.

89. Schrijvers D, Hulstijn W, Sabbe BGC. Psychomotor symptoms in depression: a diagnostic, pathophysiological and therapeutic tool. *J Affect Disord*. 2008;109(1-2):1–20.

90. Benazzi F. Psychomotor changes in melancholic and atypical depression: unipolar and bipolar-II subtypes. *Psychiatry Res*. 2002;112(3):211–20.

91. Iwanami T, Maeshima H, Baba H, et al. Psychomotor agitation in major depressive disorder is a predictive factor of mood-switching. *J Affect Disord*. 2015;170:185–9.

92. Leventhal AM, Gelernter J, Oslin D, Anton RF, Farrer LA, Kranzler HR. Agitated depression in substance dependence. *Drug Alcohol Depend*. 2011;116(1-3):163–9.

93. Calugi S, Cassano GB, Litta A, et al. Does psychomotor retardation define a clinically relevant phenotype of unipolar depression? *J Affect Disord*. 2011;129(1-3):296–300.

94. 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*. 2008;69(12):1847–55.

95. Frank E, Cassano GB, Rucci P, et al. Predictors and moderators of time to remission of major depression with interpersonal psychotherapy and SSRI pharmacotherapy. *Psychol Med*. 2011;41(1):151–62.

96. Buyukdura JS, McClintock SM, Croarkin PE. Psychomotor retardation in depression: biological underpinnings, measurement, and treatment. *Prog Neuropsychopharmacol Biol Psychiatry*. 2011;35(2):395–409.

97. Gold PW. The organization of the stress system and its dysregulation in depressive illness. *Mol Psychiatry*. 2014;20(1):32–47.

98. Romera I, Pérez V, Ciudad A, et al. Residual symptoms and functioning in depression, does the type of residual symptom matter? A post-hoc analysis. *BMC Psychiatry*. 2013;13:51.

99. Menza M, Marin H, Opper RS. Residual symptoms in depression: can treatment be symptom-specific? *J Clin Psychiatry*. 2003;64(5):516–23.

100. Fried EI, Nesse RM. Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study. *J Affect Disord*. 2015;172:96–102.

101. Ferrari AJ, Charlson FJ, Norman RE, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med*. 2013;10(11):e1001547.

102. Greer TL, Kurian BT, Trivedi MH. Defining and measuring functional recovery from depression. *CNS Drugs*. 2010;24(4):267–84.

103. Collins JJ, Baase CM, Sharda CE, et al. The assessment of chronic health conditions on work performance, absence, and total economic impact for employers. *J Occup Environ Med*. 2005;47(6):547–57.

104. Frey JJ, Osteen PJ, Berglund PA, Jinnett K, Ko J. Predicting the Impact of Chronic Health Conditions on Workplace Productivity and Accidents: Results From Two US Department of Energy National Laboratories. *J Occup Environ Med*. 2015.

105. Gilmour H, Patten SB. Depression and work impairment. *Heal reports*. 2007;18(1):9–22.

106. McKnight PE, Kashdan TB. The importance of functional impairment to mental health outcomes: a case for reassessing our goals in depression treatment research. *Clin Psychol Rev*. 2009;29(3):243–59.

107. American Psychiatric Association. *Practice Guideline for the Treatment of Patients with Major Depressive Disorder*.2010. Washington, DC: American Psychiatric Association Available at: http://www.psychiatryonline.com/.

108. Zimmerman M, McGlinchey JB, Posternak MA, Friedman M, Boerescu D, Attiullah N. Discordance between self-reported symptom severity and psychosocial functioning ratings in depressed outpatients: implications for how

remission from depression should be defined. *Psychiatry Res*. 2006;141(2):185–91.

109. Ferguson M, Dennehy EB, Marangell LB, Martinez J, Wisniewski SR. Impact of fatigue on outcome of selective serotonin reuptake inhibitor treatment: secondary analysis of STAR*D. *Curr Med Res Opin*. 2014;30(10):2109–18.

110. Van Noorden MS, Giltay EJ, den Hollander-Gijsman ME, van der Wee NJA, van Veen T, Zitman FG. Gender differences in clinical characteristics in a naturalistic sample of depressive outpatients: the Leiden Routine Outcome Monitoring Study. *J Affect Disord*. 2010;125(1-3):116–23.

111. Rapaport MH, Clary C, Fayyad R, Endicott J. Quality-of-life impairment in depressive and anxiety disorders. *Am J Psychiatry*. 2005;162(6):1171–8.

112. Rush AJ, Trivedi MH, Wisniewski SR, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry*. 2006;163(11):1905–17.

113. Marks I. *Behavioral Psychotherapy*. Bristol: John Writght; 1986.

114. Trivedi MH, Rush AJ, Wisniewski SR, et al. Factors associated with health-related quality of life among outpatients with major depressive disorder: a STAR*D report. *J Clin Psychiatry*. 2006;67(2):185–95.

115. Daly EJ, Trivedi MH, Wisniewski SR, et al. Health-related quality of life in depression: a STAR*D report. *Ann Clin Psychiatry*. 2010;22(1):43–55.

116. IsHak WW, Mirocha J, Pi S, et al. Patient-reported outcomes before and after treatment of major depressive disorder. *Dialogues Clin Neurosci*. 2014;16(2):171–83.

117. Joffe H, Chang Y, Dhaliwal S, et al. Lifetime history of depression and anxiety disorders as a predictor of quality of life in midlife women in the absence of current illness episodes. *Arch Gen Psychiatry*. 2012;69(5):484–92.

118. Lam RW, Filteau M-J, Milev R. Clinical effectiveness: the importance of psychosocial functioning outcomes. *J Affect Disord*. 2011;132 Suppl :S9–S13.

119. The Methodology Center, The Pennsylvania State University. LCA and LTA FAQ. 2014. Available at: https://methodology.psu.edu/ra/lca/faq#t18n113. Accessed March 31, 2015.

120. Insel TR, Wang PS. The STAR*D trial: revealing the need for better treatments. *Psychiatr Serv*. 2009;60(11):1466–7.

121. Hyman S. Mental health: Depression needs large human-genetics studies. *Nature*. 2014;515(7526):189–91.

122. NIH Office of Research on Women's Health. *Moving Into the Future with New Dimensions and Strategies for Women's Health Research: A Vision for 2020 for Women's Health Research*.; 2010. Available at: http://orwh.od.nih.gov/research/strategicplan/index.asp.

123. Lamers F, van Oppen P, Comijs HC, et al. Comorbidity patterns of anxiety and depressive disorders in a large cohort study: the Netherlands Study of Depression and Anxiety (NESDA). *J Clin Psychiatry*. 2011;72(3):341–8.

124. 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*. 2008;165(3):342–51.

125. Zimmerman M, Posternak MA, Chelminski I. Implications of using different cut-offs on symptom severity scales to define remission from depression. *Int Clin Psychopharmacol*. 2004;19(4):215–20.

126. Gordon E, Rush AJ, Palmer DM, Braund TA, Rekshan W. Toward an online cognitive and emotional battery to predict treatment remission in depression. *Neuropsychiatr Dis Treat*. 2015;11:517–31.

127. Gogtay N, Insel T. *Changing NIMH Clinical Trials: Efficiency, Transparency, and Reporting*.; 2014. Available at: http://www.nimh.nih.gov/funding/opportunities-announcements/clinical-trials-foas/changing-nimh-clinical-trials-efficiency-transparency-and-reporting.shtml. Accessed April 4, 2015.

128. Correll CU, Carbon M. Efficacy of pharmacologic and psychotherapeutic interventions in psychiatry: to talk or to prescribe: is that the question? *JAMA Psychiatry*. 2014;71(6):624–6.