

# Statistical Approaches to Modeling Multiple Outcomes in Psychiatric Studies

Armando Teixeira-Pinto, PhD; Juned Siddique, DrPH; Robert Gibbons, PhD; and Sharon-Lise Normand, PhD

Psychiatric Annals, Volume 39, Issue 7, July 2009

## CME EDUCATIONAL OBJECTIVES

1. Review the role of multiple outcomes in psychiatric trials.
2. Review different analytic strategies for analysis of multiple outcome measures.
3. Discuss the best approach(s) to the simultaneous analysis of multiple outcome measures.

## ABOUT THE AUTHOR

Armando Teixeira-Pinto, PhD, is with the Department of Biostatistics and Medical Informatics, Faculty of Medicine, CINTESIS, University of Porto, Portugal. Juned Siddique, DrPH, is with the Department of Preventive Medicine, Northwestern University, Chicago. Robert Gibbons, PhD, is with the Center for Health Statistics, University of Illinois at Chicago. Sharon-Lise Normand, PhD, is with the Department of Health Care Policy, Harvard Medical School; Department of Biostatistics, Harvard School of Public Health.

Address correspondence to: Armando Teixeira-Pinto, PhD, Department of Biostatistics and Medical Informatics, Faculty of Medicine, University of Porto, 4200-319 Porto, Portugal; fax: 351 22 551 3623; or e-mail [tpinto@post.harvard.edu](mailto:tpinto@post.harvard.edu).

Dr. Teixeira-Pinto and Dr. Siddique have disclosed no relevant financial relationships. Dr. Normand disclosed the following relevant financial relationship: National Institute of Mental Health (MH54693): research grant recipient. Dr. Gibbons disclosed the following relevant financial relationships: National Institutes of Mental Health: research grant recipient (R56-MH078580 and R01-MH8012201)

The WeCare data were generously provided through the efforts of Dr. Jeanne Miranda. The authors are also grateful to Dr. Hendricks Brown and Dr. Elizabeth Stuart for their valuable comments and suggestions. Dr. Miranda has disclosed no relevant financial relationships. Dr. Stuart has disclosed the following relevant financial relationships: Center for Prevention and Early Intervention, jointly funded by the National Institute of Mental Health (NIMH) and the National Institute on Drug Abuse (Grant MH066247; PI: N. Ialongo), and NIMH grant K25-MH083846: research grant recipient. Dr. Brown has disclosed the following relevant financial relationship: research grant recipient (NIMH Grant R01-MH040859).

doi: 10.3928/00485713-20090625-08

## PARTICIPANT ATTESTATION

\_\_\_ I certify that I have read the article(s) on which this activity is based, and claim credit commensurate with the extent of my participation.

## COMMERCIAL BIAS EVALUATION

Please rate the degree to which the content presented in this activity was free from commercial bias.

No bias	Significant bias			
5	4	3	2	1

Comments regarding commercial bias: \_\_\_\_\_

## INSTRUCTIONS

1. Review the stated learning objectives of the CME articles and determine if these objectives match your individual learning needs.
2. Read the articles carefully. Do not neglect the tables and other illustrative materials, as they have been selected to enhance your knowledge and understanding.
3. The following quiz questions have been designed to provide a useful link between the CME articles in the issue and your everyday practice. Read each question, choose the correct answer, and record your answer on the CME REGISTRATION FORM at the end of the quiz. Retain a copy of your answers so that they can be compared with the correct answers should you choose to request them.
4. Type your full name and address and your date of birth in the space provided on the CME REGISTRATION FORM.
5. Complete the evaluation portion of the CME REGISTRATION FORM. Forms and quizzes cannot be processed if the evaluation portion is incomplete. The evaluation portion of the CME REGISTRATION FORM will be separated from the quiz upon receipt at PSYCHIATRIC ANNALS. Your evaluation of this activity will in no way affect the scoring of your quiz.
6. Your answers will be graded, and you will be advised whether you have passed or failed. Unanswered questions will be considered incorrect. A score of at least 80% is required to pass. Your certificate will be mailed to you at the mailing address provided. Upon receiving your grade, you may request quiz answers. Contact our customer service department at (856) 994-9400.
7. Be sure to complete the CME REGISTRATION FORM on or before July 31, 2010. After that date, the quiz will close. Any CME REGISTRATION FORM received after the date listed will not be processed.
8. This activity is to be completed and submitted online only.

**Indicate the total time spent on the activity** (reading article and completing quiz). Forms and quizzes cannot be processed if this section is incomplete. All participants are required by the accreditation agency to attest to the time spent completing the activity.

## CME ACCREDITATION

This CME activity is primarily targeted to patient-caring physicians specializing in psychiatry. There are no specific background requirements for participants taking this activity. Learning objectives are found at the beginning of each CME article.

This activity has been planned and implemented in accordance with the Essential Areas and policies of the Accreditation Council for Continuing Medical Education through the joint sponsorship of Vindico Medical Education and PSYCHIATRIC ANNALS. Vindico Medical Education is accredited by the ACCME to provide continuing medical education for physicians.

Vindico Medical Education designates this educational activity for a maximum of 3 AMA PRA Category 1 Credits™. Physicians should only claim credit commensurate with the extent of their participation in the activity.

## FULL DISCLOSURE POLICY

In accordance with the Accreditation Council for Continuing Medical Education's Standards for Commercial Support, all CME providers are required to disclose to the activity audience the relevant financial relationships of the planners, teachers, and authors involved in the development of CME content. An individual has a **relevant financial relationship** if he or she has a financial relationship in any amount occurring in the last 12 months with a commercial interest whose products or services are discussed in the CME activity content over which the individual has control. Relationship information appears at the beginning of each CME-accredited article in this issue.

## UNLABELED AND INVESTIGATIONAL USAGE

The audience is advised that this continuing medical education activity may contain references to unlabeled uses of FDA-approved products or to products not approved by the FDA for use in the United States. The faculty members have been made aware of their obligation to disclose such usage.

## HOW TO OBTAIN CME CREDITS BY READING THIS ISSUE

This CME activity is primarily targeted to patient-caring physicians specializing in psychiatry. Physicians can receive *AMA PRA Category 1 Credits™* by reading the CME articles in *PSYCHIATRIC ANNALS* and successfully completing the quiz at the end of the articles. Complete instructions are given subsequently. Educational objectives are found at the beginning of each CME article.

### CME ACCREDITATION

This activity has been planned and implemented in accordance with the Essential Areas and policies of the Accreditation Council for Continuing Medical Education through the joint sponsorship of Vindico Medical Education and *PSYCHIATRIC ANNALS*. Vindico Medical Education is accredited by the ACCME to provide continuing medical education for physicians.

Vindico Medical Education designates this educational activity for a maximum of 3 *AMA PRA Category 1 Credits™*. Physicians should only claim credit commensurate with the extent of their participation in the activity.

### FULL DISCLOSURE POLICY

In accordance with the Accreditation Council for Continuing Medical Education's Standards for Commercial Support, all CME providers are required to disclose to the activity audience the **relevant financial relationships** of the planners, teachers, and authors involved in the development of CME content. An individual has a relevant financial relationship if he or she has a financial relationship in any amount occurring in the last 12 months with a commercial interest whose products or services are discussed in the CME activity content over which the individual has control. Relationship information appears at the beginning of each CME-accredited article in this issue.

### UNLABELED AND INVESTIGATIONAL USAGE

The audience is advised that this continuing medical education activity may contain references to unlabeled uses of FDA-approved products or to products not approved by the FDA for use in the United States. The faculty members have been made aware of their obligation to disclose such usage.

**Financial Disclosures:** Stanley Caroff, MD, has disclosed no relevant financial relationships. John M. Davis, MD, has disclosed no relevant financial relationships. Jan Fawcett, MD, has disclosed the following relevant financial relationships: Merck: Member of Editorial Advisory Board, Merck Manual. Paula Hensley, MD, has disclosed no relevant financial relationships. Andrew A. Nierenberg, MD, has disclosed the following relevant financial relationships: Abbott, Brain Cells Inc., Bristol Myers Squibb (BMS), Eli Lilly, GlaxoSmithKline (GSK), Innapharma, PGx, Ortho-McNeil Janssen, Novartis, Pfizer, Sepracor, Shire, and Somerset: Member of Advisory Board/Consultant: Massachusetts General Hospital (MGH) Structured

Interview Guide to the Montgomery-Åsberg Rating Scale (MADRS), and the Clinical Positive Affect Scale, licensed exclusively to the MGH Clinical Trials Network and Institute. Copyright Holder: BMS, Cederroth, Cyberonics, Eli Lilly, Forest, GSK, Ortho-McNeil Janssen; Lichtwer, NARSAD, National Institute of Mental Health; Pfizer, Stanley Foundation, Wyeth: Research Grant Recipient; and BMS, Cyberonics, Eli Lilly, Forest, GSK, Wyeth: Member of Speakers' Bureau. The staff of *Psychiatric Annals* have disclosed no relevant financial relationships.

### INSTRUCTIONS

1. Review the stated learning objectives of the CME articles and determine if these objectives match your individual learning needs.
2. Read the articles carefully. Do not neglect the tables and other illustrative materials, as they have been selected to enhance your knowledge and understanding.
3. The following quiz questions have been designed to provide a useful link between the CME articles in the issue and your everyday practice. Read each question, choose the correct answer, and record your answer on the CME REGISTRATION FORM at the end of the quiz. Retain a copy of your answers so that they can be compared with the correct answers should you choose to request them.
4. Type or print your full name and address and your date of birth in the space provided on the CME REGISTRATION FORM.
5. Complete the evaluation portion of the CME REGISTRATION FORM. Forms and quizzes cannot be processed if the evaluation portion is incomplete. The evaluation portion of the CME REGISTRATION FORM will be separated from the quiz upon receipt at *PSYCHIATRIC ANNALS*. Your evaluation of this activity will in no way affect the scoring of your quiz.
6. Send the completed form, with your \$25 payment (check, money order, or credit card information) to: VINDICO MEDICAL EDUCATION, PO Box 36, Thorofare NJ 08086. Payment should be made in US dollars drawn on a US bank.
7. Your answers will be graded, and you will be advised whether you have passed or failed. Unanswered questions will be considered incorrect. A score of at least 80% is required to pass. Upon receiving your grade, you may request quiz answers. Contact our customer service department at (856) 994-9400.
8. Be sure to mail the CME REGISTRATION FORM on or before the deadline listed. After that date, the quiz will close. Any CME REGISTRATION FORM received after the date listed will not be processed.

**Indicate the total time spent on the activity** (reading article and completing quiz). Forms and quizzes cannot be processed if this section is incomplete. All participants are required by the accreditation agency to attest to the time spent completing the activity.

## EDUCATIONAL OBJECTIVES OVERVIEW

If any readers of *Psychiatric Annals* read any psychiatric articles in any journal so that they can keep up to date and use the best-available evidence in their clinical practice, then the statistical papers in this issue are essential not only to read, but also to re-read, study, and master. What do these excellent statistical papers have to do with clinical psychiatric practice? They contain the tools that will allow you to critically appraise and interpret the psychiatric literature. They demystify the statistics used to generate the evidence and help you to become statistically literate. They give you the tools to turn data into information and knowledge.

## TABLE OF CONTENTS

711	Where Do We Go Wrong in Assessing Risk Factors, Diagnostic and Prognostic Tests? The Problems of Two-by-two Association Helen Chmura Kraemer, PhD; and Robert D. Gibbons, PhD
719	Using Nonexperimental Data to Estimate Treatment Effects Elizabeth A. Stuart, PhD; Sue M. Marcus, PhD; Marcela V. Horvitz-Lennon, MD; Robert D. Gibbons, PhD; Sharon-Lise T. Normand, PhD; and C. Hendricks Brown, PhD
729	Statistical Approaches to Modeling Multiple Outcomes in Psychiatric Studies Armando Teixeira-Pinto, PhD; Juned Siddique, DrPH; Robert Gibbons, PhD; and Sharon-Lise Normand, PhD
736	Why Does the Clinical Trial Methodology So Often Misperceive Clinical Decision Making? Focus on Moderators and Mediators of Treatment Helen Chmura Kraemer, PhD; and Robert D. Gibbons, PhD

## RESPONSIBILITY FOR STATEMENTS

All opinions expressed by authors and quoted sources are their own and do not necessarily reflect the opinions of the editors, publishers, or editorial boards of *Psychiatric Annals* or its employees, Vindico Medical Education or its employees, or the University of New Mexico. The acceptance of advertising in no way implies endorsement by the editors, publishers, or editorial boards of *Psychiatric Annals*.

The material presented at or in any *Psychiatric Annals* or Vindico Medical Education continuing education activity does not necessarily reflect the views and opinions of Vindico Medical Education or *Psychiatric Annals*. Neither *Psychiatric Annals*, Vindico Medical Education, nor the faculty endorse or recommend any techniques, commercial products, or manufacturers. The faculty/authors may discuss the use of materials and/or products that have not yet been approved by the U.S. Food and Drug Administration. Articles are intended for informational purposes only and should not be used as the basis of patient treatment. All readers and continuing education participants should verify all information before treating patients or utilizing any product.

Copyright © 2009 by SLACK Incorporated. All rights reserved. No part of this publication may be reproduced without prior written consent of the publisher.

# Statistical Approaches to Modeling Multiple Outcomes in Psychiatric Studies



## CME EDUCATIONAL OBJECTIVES

1. Review the role of multiple outcomes in psychiatric trials.
2. Review different analytic strategies for analysis of multiple outcome measures.
3. Discuss the best approach(s) to the simultaneous analysis of multiple outcome measures.

Armando Teixeira-Pinto, PhD, is with the Department of Biostatistics and Medical Informatics, Faculty of Medicine, CINTESIS, University of Porto, Portugal. Juned Siddique, DrPH, is with the Department of Preventive Medicine, Northwestern University, Chicago. Robert Gibbons, PhD, is with the Center for Health Statistics, University of Illinois at Chicago. Sharon-Lise Normand, PhD, is with the Department of Health Care Policy, Harvard Medical School; Department of Biostatistics, Harvard School of Public Health.

Address correspondence to: Armando Teixeira-Pinto, PhD, Department of Biostatistics and Medical Informatics, Faculty of Medicine, University of Porto, 4200-319 Porto, Portugal; fax: 351 22 551 3623; or e-mail [tpinto@post.harvard.edu](mailto:tpinto@post.harvard.edu).

Dr. Teixeira-Pinto and Dr. Siddique have disclosed no relevant financial relationships. Dr. Normand disclosed the following relevant financial relationship: National Institute of Mental Health (MH54693): research grant recipient. Dr. Gibbons disclosed the following relevant financial relationships: National Institutes of Mental Health: research grant recipient (R56-MH078580 and R01-MH8012201)

The WeCare data were generously provided through the efforts of Dr. Jeanne Miranda. The authors are also grateful to Dr. Hendricks Brown and Dr. Elizabeth Stuart for their valuable comments and suggestions. Dr. Miranda has disclosed no relevant financial relationships. Dr. Stuart has disclosed the following relevant financial relationships: Center for Prevention and Early Intervention, jointly funded by the National Institute of Mental Health (NIMH) and the National Institute on Drug Abuse (Grant MH066247; PI: N. Ialongo), and NIMH grant K25-MH083846: research grant recipient. Dr. Brown has disclosed the following relevant financial relationship: research grant recipient (NIMH Grant R01-MH040859).

doi: 10.3928/00485713-20090625-08

**Armando Teixeira-Pinto, PhD; Juned Siddique, DrPH; Robert Gibbons, PhD;  
and Sharon-Lise Normand, PhD**

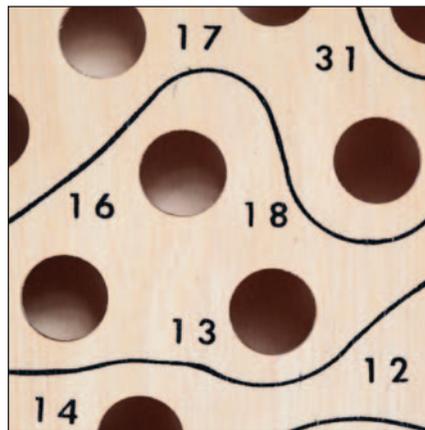
Multiple outcomes are increasingly collected both in randomized clinical trials and observational studies in order to characterize treatment or intervention effectiveness, or to investigate the association of the outcomes with other variables of interest. The decision to include more than one outcome arises for several reasons, including a lack of consensus on the most important clinical outcomes or a desire to demonstrate effectiveness on more than one outcome. The inclusion of multiple outcomes is particularly common in psychiatric studies where disease complexity is often not adequately characterized by a single outcome measure. Depression, for example, is assessed by multiple instruments.

The collection of several outcomes in a study allows different analytical strategies for analysis. The outcomes can be combined into a single composite endpoint using a variety of pooling rules or scoring algorithms. Several types of composite endpoints exist, such as taking a simple average of the outcomes or using conjunctive or compensatory rules.<sup>1</sup> Another frequently adopted option is to consider each outcome separately by analyzing each independently of the others.<sup>2</sup> However, the situation of multiple outcomes fits perfectly in the framework of several statistical methods designated as multivariate methods, and in particular, multiple informant analyses.<sup>3,4</sup>

Pooling strategies have the disadvantage of reducing the information collected and potentially attenuating important features of the data. Also, any missing observation in one outcome may reduce the sample size if a complete-case analysis is adopted, although we could complement this approach with some sort of imputation technique,<sup>5,6</sup> or may produce biased estimates when using available-case analysis even if the missingness is completely at random.<sup>7</sup> Another major drawback of pooling is that it fails when the outcomes are of different natures or are measured on different scales (ie, non-commensurate outcomes).

For example, combining a binary outcome (such as the presence of a symptom) and a continuous outcome (such as a well-being score) requires additional decisions. Often, information is wasted by dichotomizing the continuous outcome so it is “poolable” with the binary outcome.

Analyzing the outcomes separately does not require that the outcomes are commensurate (measured on the same



*Several types of composite endpoints exist, such as taking a simple average of the outcomes or using conjunctive or compensatory rules.*

scale) because each outcome is treated as if the other outcomes were not observed. Although the simplicity of such an approach is appealing, the correlation between the outcomes is effectively ignored. This could result in a loss of efficiency in the analysis leading to less power to detect treatment effects (and larger confidence intervals for the estimates). If some outcomes are missing for individuals, separate analyses may produce biased estimates of the covariate effects on the outcomes. Finally, if primary interest is in testing for an overall treatment effect, separate analyses do not provide such an estimate without further work, and individual tests for each outcome raise the issue of adjusting the  $P$

values for multiple comparisons.<sup>8</sup>

In this article, we present a multivariate method that 1) analyzes all the outcomes at the same time by taking into account their correlations and 2) allows mixtures of different types of outcomes (for example binary and continuous outcomes). Other approaches have been proposed to analyze non-commensurate outcomes in a multivariate framework but with some limitations regarding the settings where they can be applied.<sup>9,10</sup> In the section on Treating Depression in low-income women, we introduce a real data example that will be used throughout the article. In the section on statistical methods, we contrast the finding using individual analysis of the outcomes with those using multivariate methods and interpretation of the results.

#### **TREATING DEPRESSION IN LOW-INCOME WOMEN**

The WeCare Study investigated outcomes during a 12-month period in which 267 low-income, mostly minority, women in suburban Washington, D.C., were treated for major depressive disorder.<sup>11</sup> Participants were screened for depression at Women, Infant, and Children (WIC) clinics and various pediatric clinics. Subjects were randomly assigned to one of three groups: Medication, Cognitive Behavioral Therapy, and care-as-usual, the latter of which consisted of a referral to a community provider. The objective of the primary study was determination of the benefit of medication and cognitive behavioral therapy relative to community referral. Participants were interviewed by phone at baseline, every month for 6 months, and then every other month for the duration of the study. Major clinical outcomes were depression score measured using the Hamilton Depression Rating Scale (HDRS), instrumental role functioning as measured by the Social Adjustment Scale, social functioning (SF) as measured by the Short Form 36-Item Health Survey, and depression remission defined as an HDRS score of 7 or less. Smaller values of the instru-

TABLE 1.

### Social and Demographic Characteristics and Baseline Measurements of Depression Scores by Treatment Arm for Women Enrolled in the WeCare Study

Characteristic	Total (n = 267); mean (SD)	Medication (n = 88); mean (SD)	Cognitive Behavioral Therapy (n = 90); mean (SD)	Community Referral (n = 89); mean (SD)
<b>Age</b>	29.3 (7.9)	28.7 (6.6)	29.8 (7.9)	29.5 (9.1)
<b>Number of Children</b>	2.3 (1.4)	2.2 (1.2)	2.2 (1.5)	2.4 (1.6)
<b>Baseline for Social Functioning</b>	57.7 (25.3)	56.5 (24.6)	56.5 (23.9)	60.0 (27.5)
<b>Baseline for Instrumental Role Functioning</b>	3.5 (1.2)	3.6 (1.3)	3.5 (1.2)	3.3 (1.2)
<b>Baseline Hamilton Score</b>	16.9 (5.2)	17.9 (5.1)	16.3 (5.1)	16.5 (5.2)
	<b>n (%)</b>	<b>n (%)</b>	<b>n (%)</b>	<b>n (%)</b>
<b>Employment</b>				
Working or looking for work	219 (82.0)	69 (78.4)	76 (84.4)	74 (83.2)
Not working or disabled	48 (18.0)	19 (21.6)	14 (15.6)	15 (16.9)
<b>Education</b>				
Less than high school	99 (37)	37 (42)	27 (30)	35 (39)
High school	87 (33)	31 (35)	29 (32)	27 (30)
Some trade or college	63 (26)	15 (17)	26 (29)	22 (25)
College graduate	18 (7)	5 (6)	8 (9)	5 (6)
<b>Marital Status</b>				
Married or living with partner	124 (46)	43 (49)	40 (44)	41 (46)
Widowed or separated	52 (20)	17 (19)	22 (24)	13 (15)
Never married	91 (34)	28 (32)	28 (31)	35 (39)
<b>Ethnicity</b>				
Black	117 (44)	34 (39)	41 (46)	42 (47)
White	16 (6)	6 (7)	6 (7)	4 (4)
Latina	134 (50)	48 (55)	43 (48)	43 (48)
<b>Schooling</b>				
Less than high school	99 (37)	37 (42)	27 (30)	35 (39)
High school or GED	87 (33)	31 (35)	29 (32)	27 (30)
Some trade or college	63 (23)	15 (17)	26 (29)	22 (25)
College graduate	18 (7)	5 (6)	8 (9)	5 (6)

*SD = Standard deviation*

TABLE 2.

### Comparison of Instrumental Role Functioning, Remission, and Treatment Arms Stratified by Missingness of 6-month Social Functioning Score

Characteristic	Social Functioning Missing (n = 83)	Social Functioning observed (n = 184)	P value
Instrumental Role Functioning, mean (SD)	4.0 (1.5)	2.3 (1.2)	< 0.001
Depression Remission (HDRS of 7 or less), n (%)	33 (40)	74 (40)	0.944
Treatment, n (%)			
Medication	7 (8)	81 (44)	< 0.001
Cognitive Behavioral Therapy	17 (21)	73 (40)	
Community Referral	59 (71)	30 (36)	

SD = Standard deviation; HDRS = Hamilton Depression Rating Scale

mental functioning score and larger values of the social functioning score correspond to better outcomes. Baseline information included age, ethnicity, income, marital status, number of children, health insurance, education, employment, and stressful life events (see Table 1, 731).

Outcomes for the first 6 months of the study were reported in Miranda J et al.<sup>11</sup> In this article, we use depression remission, instrumental role functioning, and social functioning to demonstrate statistical approaches for assessing the treatment effect on these three outcomes. For illustration purposes, several social functioning scores at 6 months were deleted to demonstrate problems with conventional methods when data are missing. The primary research question addressed in this paper is whether the Medication and Cognitive Behavioral treatment groups had better depression and functioning outcomes at 6-months as compared to the care-as-usual group (Community Referral).

## STATISTICAL METHODS

### Separate Analyses of Each Outcome

A common approach used when analyzing multiple outcomes is to analyze each outcome separately by regressing each outcome on treatment indicators and additional covariates. In the WeCare Study,

outcomes are adjusted for baseline depression in order to correct for chance initial differences among treatment groups. The regression models depend on the type of the outcome that is being modeled. For example, for continuous outcomes, a linear regression model is typically assumed, while for binary outcomes, logistic, or probit regression models are used.

For depression remission, we use a probit regression model to estimate the treatment effect adjusted for baseline depression. A probit regression model gives very similar results to the logistic regression. The regression coefficients of the probit model are approximately 1.6 times the coefficients obtained from a logistic regression model. This does not mean that the estimated effects in the probit model are larger than the ones given by the logistic model. In fact, they are very similar but measured in a different scale. We use a probit model because it allows a direct comparison of treatment effects from a multivariate approach next discussed. For the two functioning scores, we estimate linear regressions models for each outcome with treatment indicators and the respective baseline measurements as covariates.

### Multivariate Approach Using a Latent Variable

Rather than modeling each outcome separately, consider a multivariate ap-

proach that models the three outcomes in a similar way as the separate models but that additionally takes into account the correlation between the outcomes. Why would an investigator want to adopt this analytical strategy? When the study outcomes have no missing values (or they are missing completely at random), analyzing each outcome separately will provide unbiased estimates for the treatment effects, even if the outcomes are correlated. In this case, the separate models for each outcome will give correct treatment effect estimates but will have larger standard errors than if the correlations among outcomes were taken into account. With sufficiently large sample sizes, investigators may not be concerned so that the tradeoff between simplicity of the analysis procedure and larger errors might favor the simple, one-outcome-at-a-time approach.

However, what happens in the more common case when data are not missing completely at random? In the WeCare data we deleted the social functioning scores of several participants. Women missing these observations have higher instrumental role functioning scores than women with observed SF scores (see Table 2). This suggests that women with missing SF are sicker than those with measured SF. Moreover, most women with missing SF belong to the Community Referral arm (71%), and only 8% of the missing SF arises from

TABLE 3.

**Regression Coefficients Associated with Cognitive Behavioral Therapy, Medication, and Community Referral (adjusted for the respective baselines) for the Three Scores: Remission, Instrumental Role Functioning, and Social Functioning\***

6-month Outcome	Separate Regressions			Multivariate Latent Variable Model		
	Coef.	(SE)	P value	Coef.	(SE)	P value
<b>Depression Remission (HDRS of 7 or less)</b>	N = 267 participants			N = 267 participants		
Hamilton score at baseline	-0.03	(0.02)	0.551	-0.02	(0.01)	0.146
Cognitive Behavioral therapy	-0.01	(0.19)	0.944	0.01	(0.19)	0.960
Medication	0.38	(0.19)	0.050	0.36	(0.19)	0.062
<b>Community Referral</b>	<b>Reference Group</b>					
Instrumental Role Functioning	N = 267 participants			N = 267 participants		
Instrumental Role Functioning at baseline	0.21	(0.07)	0.003	0.11	(0.06)	0.080
Cognitive Behavioral therapy	-0.11	(0.22)	0.606	-0.09	(0.22)	0.691
Medication	-0.87	(0.22)	< 0.001	-0.84	(0.22)	< 0.001
<b>Community Referral</b>	<b>Reference Group</b>					
SD of the error term	1.42	(0.06)	< 0.001	1.04	(0.06)	< 0.001
Social Functioning	N = 184 participants			N = 267 participants		
Social Functioning at baseline	0.26	(0.07)	< 0.001	0.27	(0.06)	< 0.001
Cognitive Behavioral therapy	2.74	(4.94)	0.581	7.67	(4.73)	0.107
Medication	6.99	(4.86)	0.153	13.37	(4.70)	0.005
<b>Community Referral</b>	<b>Reference Group</b>					
SD of the error term	22.48	(1.17)	< 0.001	16.94	(0.93)	< 0.001
SD of the latent variable				0.94	(0.09)	< 0.001

\*Comparison between the approach using separate regressions for each outcome and the multivariate approach using a latent variable model.

SE = Standard error; SD = Standard deviation; HDRS = Hamilton Depression Rating Scale

the Medication arm. Statistical theory tells us that a model for subjects with observed SF will produce biased estimates for the treatment effect.<sup>5,6</sup>

There are several options available to researchers to model multiple outcomes when measured on the same scale, referred to as commensurate outcomes. For example, for normally distributed outcomes, we can use a multivariate linear regression, or for multiple binary outcomes, we can use a generalized linear mixed model.<sup>10,12,13</sup> With outcomes that are not measured on

the same scale, such as non-commensurate outcomes, there is no simple multivariate distribution to use. The difficulty arises because there is no obvious way to express the multivariate distribution for the mixed type of outcomes.

What can be used instead? The trick is to include a common unobserved (or latent) variable for all three regression equations. This latent variable establishes the link between the regression equations — the outcomes are measured on the same individuals so the latent variable

induces the needed correlations among the outcomes. We assume the latent variable completely specifies the correlation among the outcomes (ie, given the latent variable the outcomes are assumed to be independent). This permits examination of the outcomes as independent of each other by accounting for the correlation through the latent variable.

The latent variable is assumed to have a normal distribution with mean 0 and some variance and is scaled (multiplied by a value that has to be estimated) to ac-

### **/\* SAS Code for Using the PROC NL MIXED to Fit the Latent Model for the WeCare Data**

#### **Abbreviations:**

dr: depression remission  
 irf: instrumental role functioning  
 sf: social functioning  
 HRDSbline: Hamilton score at baseline  
 irfbline: IRF at baseline  
 sfbline: SF at baseline  
 cbt: Cognitive Behavioral Therapy arm  
 medic: Medication arm  
 tau : latent variable

#### **PROC NL MIXED data = wecare\_data\_withmiss GCONV=1E-15;**

```
parms a_dr=4 b_dr=3 c_dr=3 f_dr=0.03
      a_irf=1.7 b_irf=.76 c_irf=.87 f_irf=.21 sigma_irf=1
      a_sf=67 b_sf=-5 c_sf=-12 f_sf=0.25 sigma_sf=12
      sigmatau=1; *initial estimates obtained from the separated regressions;
bounds sigma_irf>0, sigma_sf>0, sigmatau>0.01; *constraining the std deviations to be positive;
stdconst = sqrt(1+sigmatau**2); *constant used in the probit model to obtain the marginal effects;

*construction of the likelihood;
p = a_dr*stdconst + b_dr*stdconst*cbt + c_dr*stdconst* medic + f_dr*stdconst*HRDSbline +
tau;
mean_irf = a_irf + b_irf*cbt + c_irf*medic + f_irf*irfbline + sigma_irf*tau;
mean_sf = a_sf + b_sf*cbt + c_sf* medic + f_sf*sfbline + sigma_sf*tau;
ll_dr = dr*log(PROBNORM(p)) + (1-dr)*log(PROBNORM (-p));
ll_irf = -.5*((irf - mean_irf)/sigma_irf)**2 -log(sigma_irf);
ll_sf = -.5*((sf-mean_sf)/sigma_sf)**2 -log(sigma_sf);
ll = ll_dr+ll_irf+ll_sf; *log likelihood for individual with complete observations;
if missing (sf) then ll=ll1+ll2; *log likelihood for individual with missing SF;

*model;
model medic ~ general(ll);
random tau ~ normal(0, sigmatau**2) subject=MID; *latent variable;
run;
```

commodate the different nature of each outcome. The only restriction regarding the correlation is that the outcomes have to be positively correlated. If some of the correlations are negative, they can easily be changed to positive by inverting the outcome scale. This is accomplished by multiplying one of the negatively correlated outcomes by minus one. In the WeCare example, the outcome instrumental role functioning is negatively correlated with depression remission and social

functioning. We therefore multiply each participant's instrumental role functioning score by minus one, changing the correlation with the remaining outcomes to positive. The covariate effects for instrumental role functioning are afterwards multiplied again by minus one in order for them to be interpreted in the correct scale.

#### **Interpretation of the Regression Parameters**

The regression equations used in the

latent variable approach are conditional on the latent variable. For this reason, the regression parameters in the latent variable model also have to be interpreted somewhat differently than when fitting separate models. However, investigators are most interested on the unconditional effects, similar to the usual interpretation in regression models. For the continuous outcomes, the regression parameters can be interpreted in the same manner as those from a separate analyses approach. For the binary outcomes, this is not the case. To compute the treatment effect that is comparable to that obtained from a separate analysis, the regression coefficients are divided by the square root of (one plus the variance of the latent variable). A more detailed discussion about this can be found in Teixeira-Pinto and Normand.<sup>14</sup>

In Table 3 (see page 733), the treatment estimates for remission and instrumental role functioning and the outcomes that have no missing data, are virtually identical between the separate regression and the multivariate latent variable approaches. The Table 3 (see page 733) estimates have been transformed and are therefore directly comparable to those from the separate regression models. The coefficients from the probit model for the depression remission outcome can be approximated to odds ratios by taking the exponential of 1.6 times the coefficients. Using the coefficients in Table 3 (see page 733) from the latent variable model, the odds ratios of remission when assigned Medication and Cognitive Behavioral Therapy relative to Community Referral are  $\exp(0.36 \times 1.6) = 1.78$  (95% Confidence Interval = (1.11; 4.55)) and  $\exp(0.01 \times 1.6) = 1.02$  (95% Confidence Interval = (0.63; 2.56)), respectively.

The estimated treatment effects for social functioning are very different between the two analytical approaches. In separate analyses, the Medication arm has an average increase of 6.99 (SE = 4.86) points compared to the Community Referral but this effect is not statistically significant (*P*

= 0.153). In contrast, the multivariate regression approach using the latent variable yields a statistically significant benefit ( $P = 0.005$ ) of 13.37 SE = 4.70) points in the Medication arm compared with the Community Referral arm. Why does this occur? Because the outcomes are correlated, information is “borrowed” from the other outcomes through the correlation in order to compensate for the missing outcome. This information is passed through the latent variable and consequently, the estimation of the regression parameters for social functioning should be less biased. For the other outcomes, the estimates are identical to the separate analysis because they have complete data. In fact, the correlation between depression remission and the two other outcomes is -0.39 and 0.36 for instrumental role function and social function, respectively, and the correlation between these last two is -0.40. The effect in the Cognitive Behavioral Therapy arm is not statistically significant in both approaches but the effect estimate obtained with the latent variable model is almost three times larger than the estimate obtained with the separate regressions (7.67 (SE = 4.73) versus 2.74 (SE = 4.94) points).

Because we deleted the missing data, we have the complete data and are able to calculate the bias of both approaches by fitting the model for the complete data. The treatment estimates obtained with the complete data for the social functioning are 12.07 for the Medication arm and 5.13 for the Cognitive Behavioral Therapy arm, clearly closer to the results obtained with the latent variable model.

### CONCLUDING REMARKS

Historically, the use of multivariate methods for the general linear model required complete data and an assumption that the multiple outcomes of interest had jointly normal distributions (ie, multivariate normality). Although this restricted their application to continuous and normally distributed outcomes, the advantage of the multivariate approach was that it

provided more parsimonious hypothesis tests and interval estimates than a series of univariate tests. With the development of generalized non-linear mixed-effects regression models that can accommodate missing data under fairly general statistical assumptions, the advantages of the multivariate approach can now be extended to reduction in bias produced by missing data. As illustrated here, these methods can be applied to outcomes measured on different scales, including a mixture of discrete and continuous outcomes.

We have presented a multivariate strategy based on latent variable, model-to-model, mixed types of outcomes. This model is an alternative to analyzing each outcome separately, which disregards the potential correlation between the outcomes and to pooling information into a composite endpoint and which loses information contained in the data. The dataset used as an example illustrates the advantage of the latent variable model in a situation when one of the outcomes is missing for some subjects and when the missing mechanism is not completely at random. In such cases, modeling the outcomes separately may produce biased estimates for regression parameters.

The latent variable approach has some disadvantages. The model is implemented only in a limited number of statistical packages such as MPlus<sup>15</sup> and AMOS.<sup>16</sup> Although most statistical software does not have this procedure available, a program can be written to provide model estimates (see Appendix, page 734, for an example of an SAS program). The latent variable model makes some assumptions that are not easily verifiable such as the number of latent variables necessary to specify the correct correlation or the assumption that the latent variable arises from a normal distribution (although there has been some recent developments regarding this last point).<sup>17</sup> Because of the increasing frequency of multiple outcomes reported in psychiatric studies, however, the potential benefits in terms of increases in pre-

cision of estimation and power of testing by adopting a multivariate approach are extremely promising.

### REFERENCES

1. Neuhauser M. How to deal with multiple endpoints in clinical trials. *Fundam Clin Pharmacol*. 2006;20(6):515-523.
2. Pocock SJ, Geller NL, Tsiatis AA. The analysis of multiple endpoints in clinical trials. *Biometrics*. 1987;43(3):487-498.
3. Deskalakis C, Laird NM, Murphy JM. Regression analysis of multiple-source longitudinal outcomes: a “Stirling County” depression study. *Am J Epidemiol*. 2002;155(1):88-94.
4. Horton NJ, Fitzmaurice GM. Regression analysis of multiple source and multiple informant data from complex survey samples. *Stat Med*. 2004;23(18):2911-2933.
5. Siddique J, Brown CH, Hedeker D, et al. Missing data in longitudinal clinical trials, Part B: analytic issues. *Psychiatric Annals*. 2008;38(12):793-801.
6. Lavori PW, Brown CH, Duan N, Gibbons RD, Greenhouse J. Missing data in longitudinal clinical trials, Part A: design and conceptual issues. *Psychiatric Annals*. 2008;38(12):784-792.
7. Li X, Caffo B, Scharfstein D. On the potential for illogic with logically defined outcomes. *Biostat*. 2007;8(4):800-804.
8. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ*. 1995;310(6973):170.
9. Catalano PJ. Bivariate modelling of clustered continuous and ordered categorical outcomes. *Stat Med*. 1997;16(8):883-900.
10. Fitzmaurice GM, Laird NM. Regression models for mixed discrete and continuous responses with potentially missing values. *Biometrics*. 1997;53(1):110-122.
11. Miranda J, Chung JY, Green BL, et al. Treating depression in predominantly low-income young minority women: a randomized controlled trial. *JAMA*. 2003;290(1):57-65.
12. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73(1):9-22.
13. Hedeker D, Gibbons RD. *Longitudinal Data Analysis*. New York, NY: Wiley & Sons; 2006.
14. Teixeira-Pinto A, Normand SL. Correlated bivariate continuous and binary outcomes: issues and applications. *Stat Med*. 2009;28(13):1753-1773.
15. Muthén L, Muthén BO. *Mplus Statistical Analysis with Latent Variables User's Guide*. Version 5. Los Angeles, CA: Muthén & Muthén; 1998-2007.
16. Arbuckle JL. *Amos 7.0 User's Guide*. Volume 1. Chicago, IL: SPSS; 1995-2006.
17. Wang CP, Brown CH, Bandeen-Roche K. Residual diagnostics for growth mixture models: examining the impact of a preventive intervention on multiple trajectories of aggressive behavior. *Journal of the American Statistical Association*. 2005;100:1054-1076.