

Analysis of Observational Health Care Data Using SAS®

Edited by
Douglas E. Faries
Andrew C. Leon
Josep Maria Haro
Robert L. Obenchain



From *Analysis of Observational Health Care Data Using SAS*[®]. Full book available for purchase [here](#).

Contents

Preface ix

Part 1 Introduction

Chapter 1 Introduction to Observational Studies	3
1.1 Observational vs. Experimental Studies	3
1.2 Issues in Observational Studies	5
1.3 Study Design.....	9
1.4 Methods	10
1.5 Some Guidelines for Reporting	13
Acknowledgments.....	14
References	15

Part 2 Cross-Sectional Selection Bias Adjustment

Chapter 2 Propensity Score Stratification and Regression	23
Abstract	23
2.1 Introduction.....	23
2.2 Propensity Score: Definition and Rationale	24
2.3 Estimation of Propensity Scores.....	24
2.4 Using Propensity Scores to Estimate Treatment Effects: Stratification and Regression Adjustment	26
2.5 Evaluation of Propensity Scores	27
2.6 Limitations and Advantages of Propensity Scores	29
2.7 Example	30
2.8 Summary	46
Acknowledgments.....	46
References	46
Chapter 3 Propensity Score Matching for Estimating Treatment Effects.....	51
Abstract	51
3.1 Introduction	52
3.2 Estimating the Propensity Score.....	52
3.3 Forming Propensity Score Matched Sets.....	53
3.4 Assessing Balance in Baseline Characteristics.....	55

3.5	Estimating the Treatment Effect.....	58
3.6	Sensitivity Analyses for Propensity Score Matching	61
3.7	Propensity Score Matching Compared with Other Propensity Score Methods.....	62
3.8	Case Study.....	62
3.9	Summary.....	81
	Acknowledgments.....	81
	References	82

Chapter 4 Doubly Robust Estimation of Treatment Effects..... 85

	Abstract	85
4.1	Introduction.....	85
4.2	Implementation with the DR Macro	88
4.3	Sample Analysis	95
4.4	Summary	101
4.5	Conclusion	102
	References	103

Chapter 5 Propensity Scoring with Missing Values..... 105

	Abstract	105
5.1	Introduction.....	105
5.2	Data Example	107
5.3	Using SAS for IPW Estimation with Missing Values	109
5.4	Sensitivity Analyses.....	124
5.5	Discussion	128
	References	128

Chapter 6 Instrumental Variable Method for Addressing Selection Bias..... 131

	Abstract	131
6.1	Introduction.....	131
6.2	Overview of Instrumental Variable Method to Control for Selection Bias	132
6.3	Description of Case Study	134
6.4	Traditional Ordinary Least Squares Regression Method Applied to Case Study.....	138
6.5	Instrumental Variable Method Applied to Case Study	139
6.6	Using PROC QLIM to Conduct IV Analysis	143
6.7	Comparison to Traditional Regression Adjustment Method.....	146

6.8 Discussion 147
 6.9 Conclusion 148
 Acknowledgments 149
 References 149

Chapter 7 Local Control Approach Using JMP 151

Abstract 151
 7.1 Introduction 152
 7.2 Some Traditional Analyses of Hypothetical Patient Registry
 Data 156
 7.3 The Four Phases of a Local Control Analysis 163
 7.4 Conclusion 182
 Acknowledgments 184
 Appendix: Propensity Scores and Blocking/Balancing Scores 185
 References 189

Part 3 Longitudinal Bias Adjustment

Chapter 8 A Two-Stage Longitudinal Propensity Adjustment for Analysis of Observational Data 195

Abstract 195
 8.1 Introduction 195
 8.2 Longitudinal Model of Propensity for Treatment 196
 8.3 Longitudinal Propensity-Adjusted Treatment Effectiveness
 Analyses 197
 8.4 Application 198
 8.5 Summary 207
 Acknowledgments 208
 References 208

Chapter 9 Analysis of Longitudinal Observational Data Using Marginal Structural Models... 211

Abstract 211
 9.1 Introduction 211
 9.2 MSM Methodology 213
 9.3 Example: MSM Analysis of a Simulated Schizophrenia Trial 214
 9.4 Discussion 227
 References 228

Chapter 10 Structural Nested Models.....	231
Abstract	231
10.1 Introduction.....	232
10.2 Time-Varying Causal Effect Moderation	234
10.3 Estimation	236
10.4 Empirical Example: Maximum Likelihood Data Analysis Using SAS PROC NLP.....	239
10.5 Discussion	251
Appendix 10.A.....	253
Appendix 10.B.....	256
Appendix 10.C.....	259
References	260
Chapter 11 Regression Models on Longitudinal Propensity Scores.....	263
Abstract	263
11.1 Introduction.....	263
11.2 Estimation Using Regression on Longitudinal Propensity Scores	265
11.3 Example.....	266
11.4 Summary	282
References	282
Part 4 Claims Database Research	
Chapter 12 Good Research Practices for the Conduct of Observational Database Studies	287
Abstract	287
12.1 Introduction.....	287
12.2 Checklist and Discussion.....	289
Acknowledgments.....	294
References	294
Chapter 13 Dose-Response Safety Analyses Using Large Health Care Databases.....	295
Abstract	295
13.1 Introduction.....	295
13.2 Data Structure.....	297
13.3 Treatment Model and Censoring Model Setup	300
13.4 Structural Model Implementation	304

13.5 Discussion309
 References310

Part 5 Pharmacoeconomics

Chapter 14 Costs and Cost-Effectiveness Analysis Using Propensity Score Bin Bootstrapping 315

Abstract315
 14.1 Introduction.....315
 14.2 Propensity Score Bin Bootstrapping319
 14.3 Example: Schizophrenia Effectiveness Study320
 14.4 Discussion334
 References335

Chapter 15 Incremental Net Benefit..... 339

Abstract339
 15.1 Introduction.....340
 15.2 Cost-Effectiveness Analysis341
 15.3 Parameter Estimation346
 15.4 Example.....351
 15.5 Observational Studies359
 15.6 Discussion359
 Acknowledgments.....360
 References360

Chapter 16 Cost and Cost-Effectiveness Analysis with Censored Data 363

Abstract363
 16.1 Introduction.....363
 16.2 Statistical Methods.....366
 16.3 Example.....369
 16.4 Discussion379
 Acknowledgments.....380
 References381

Part 6 Designing Observational Studies

Chapter 17 Addressing Measurement and Sponsor Biases in Observational Research 385

Abstract	385
17.1 Introduction.....	385
17.2 General Design Issues.....	386
17.3 Addressing Measurement and Sponsor Bias	388
17.4 Summary	389
References	390

Chapter 18 Sample Size Calculation for Observational Studies 391

Abstract	391
18.1 Introduction.....	391
18.2 Continuous Variables	392
18.3 Binary Variables	398
18.4 Two-Sample Log-Rank Test for Survival Data	405
18.5 Two-Sample Longitudinal Data.....	410
18.6 Discussion	423
Appendix: Asymptotic Distribution of Wilcoxon Rank Sum Test under H_a	423
References	424

Index..... 427

From *Analysis of Observational Health Care Data Using SAS®* by Douglas Faries, Andrew Leon, Josep Haro, and Robert Obenchain. Copyright © 2010, SAS Institute Inc., Cary, North Carolina, USA. ALL RIGHTS RESERVED.



From *Analysis of Observational Health Care Data Using SAS*[®]. Full book available for purchase [here](#).

Chapter 9

Analysis of Longitudinal Observational Data Using Marginal Structural Models

Douglas E. Faries
Zbigniew A. Kadziola

Abstract	211
9.1 Introduction	211
9.2 MSM Methodology	213
9.3 Example: MSM Analysis of a Simulated Schizophrenia Trial	214
9.4 Discussion	227
References	228

Abstract

Assessing treatment effectiveness in longitudinal, observational data can be complex because in observational treatment patients can change medications at any time. In addition to the need to control for selection bias at baseline due to the lack of randomization, time-varying confounders can influence treatment changes over time and, thus, affect treatment group effectiveness comparisons. One approach to producing causal treatment effect estimates—even in the presence of treatment switching, missing data, and time-varying confounders—is to use marginal structural models. To illustrate, simulated data based on an observational schizophrenia study were analyzed using a marginal structural model approach. SAS code for performing the analysis is provided, and output using data from the schizophrenia study is examined.

9.1 Introduction

Assessing the causal effect of medications in longitudinal, observational (naturalistic) data presents analytical challenges—including the need to address selection bias; missing data; and switching, stopping, and augmenting medications. Addressing the issue of selection bias is critical because treatment groups likely differ in aspects other than treatment choice, and adjustment in the analysis is necessary (Rosenbaum and Rubin, 1983; Grimes and Schulz, 2002; Haro et al., 2006). In addressing data that are both longitudinal and observational, the issue of selection bias also extends to treatment switching over time (Robins et al., 2000, Hernán et al., 2000) as patients may switch, stop, augment, or otherwise not comply for a variety of reasons. In addition, such patient/physician choices are typically based upon stochastic and/or time-varying factors that may well differ among treatments. Because of such issues, statistical methods commonly used for longitudinal analyses of randomized clinical trial data, such as intent-to-treat

(ITT) last observation carried forward (LOCF) or repeated measures models, may not be appropriate.

ITT analyses group patients based only on their initial treatment assignment and ignore all information on other medications prescribed or taken. Patient dropout in such studies is often addressed by utilizing a LOCF approach. Clearly, such a technique does not directly address treatment effectiveness when there has been a substantial amount of switching among treatments. ITT analyses certainly have their place in longitudinal, observational research, such as in studies to compare policies or treatment strategies where one is not primarily interested in the effects of individual medications (Tunis et al., 2006).

While utilizing repeated measures models with treatment as a time-dependent variable may seem to provide a simple solution, Hernán and colleagues (2004, 2005) explain that such an approach does not provide estimates with a causal interpretation (see the following) in the presence of time-dependent confounders (a predictor of subsequent outcome and subsequent treatment) that are also affected by prior treatment. For instance, any longitudinal measure of disease severity would likely be problematic because it could be associated with the outcome measure, it could predict subsequent treatment, and it could have been affected by prior treatment. Thus, even if treatment is randomized at the beginning of a study, the result of usual-care treatment over time will ultimately result in imbalance in key patient characteristics among treatment groups. To address the switching of treatments, one could ignore the data after the medication switch and use standard repeated measures mixed models that have proven very useful in longitudinal data analyses (Verbeke and Molenberghs, 2000; Mallinckrodt et al., 2003). Such an approach treats the data after the switch as missing data but clearly does not make use of the information gathered after the medication switch.

In this chapter, we examine the use of marginal structural models (MSMs) for longitudinal, observational data. To explain the potential benefits of the MSM approach, we first must briefly review the notions of counterfactual outcomes and causal effect. We will follow the notation provided by Hernán and colleagues (2002). Let \bar{a} denote the treatment history for a patient over a period of time (for example, $\bar{a} = [a(1), a(2), \dots, a(t)]$, where $a(1)$ denotes the treatment used at time 1). A counterfactual outcome for patient i on treatment sequence \bar{a} denotes that patient's outcome if, possibly contrary to fact, the patient received treatment \bar{a} . It is denoted by $Y_{\bar{a},i}^-(t)$.

Each patient has an unknown counterfactual outcome for each treatment he did not receive, plus an observed outcome for the treatment actually received. On an individual basis, treatment is said to have a causal effect on a patient's outcome if $Y_{\bar{a},i}^-(t) \neq Y_{\bar{a}',i}^-(t)$ for some time point t and treatment patterns \bar{a} and \bar{a}' . That is, the outcome for the patient differs based on the treatment taken. On a population basis, treatment is said to have a causal effect on outcome if the mean outcome had all patients followed a particular treatment pattern (\bar{a} for example) differs from the mean outcome had all patients followed a different treatment pattern (\bar{a}' for example) (that is, $E[Y_{\bar{a}}^-(t)] \neq E[Y_{\bar{a}'}^-(t)]$ for some time point and treatment pattern).

Robins and colleagues (1999) demonstrated that MSMs, under a set of assumptions discussed here, produce consistent estimates of the average causal treatment effects—even in the presence of treatment changes, time-dependent confounders, and missing at random study dropout. In this chapter, we first describe the MSM approach (Section 9.2) and then present the MSM analysis of a longitudinal schizophrenia study (Section 9.3). SAS code is provided, and SAS output is discussed to allow readers to understand the implementation of the analysis and to modify the

code for their own use. Faries and colleagues (2007) also summarize data from this study using a variety of methods, including MSMs. Some other applications of MSMs in the literature include Hernán and colleagues (2000, 2002), Ko and colleagues (2003), Brumback and colleagues (2004), and Cole and colleagues (2005), who examined time to event outcomes for HIV patients; and Bodnar and colleagues (2004), Yamaguchi and Ohashi (2004a and 2004b), Mortimer and colleagues (2005), Suarez and colleagues (2006), Peterson and colleagues (2007), and Vansteelandt and colleagues (2009), who assessed other applications.

9.2 MSM Methodology

An MSM analysis is basically a weighted repeated measures approach – using treatment as a time-varying covariate. Weights, based on inverse probability of treatment weighting, control for time-dependent confounders and essentially produce a pseudo-population with balance in both time-invariant and time-varying covariates allowing for causal treatment comparisons using standard repeated measure models. The weighting can also be adjusted to incorporate adjustments for missing data—providing validity under missing at random (MAR; missing data may depend upon observed but not unobserved measures) and missing completely at random (MCAR; missing data does not depend upon observed or unobserved measures) data.

Conducting an MSM analysis is a two-step process. First, one estimates two weights for each observation (patient visit): one adjusting for treatment selection and one adjusting for study discontinuation. Computation of these estimated weights can incorporate time-independent and time-dependent factors. The stabilized weight is recommended by Hernán (2002), and we use the notation from that manuscript (here for the treatment selection weight),

$$SW = \prod_{k=0}^t \frac{f[A(k) | \bar{A}(k-1), V]}{f[A(k) | \bar{A}(k-1), \bar{L}(k)]}$$

where $A(k)$ represents the treatment at time k and $\bar{A}(k-1)$ represents the treatment history prior to time k , V represents a vector of time-independent variables (baseline covariates), and $\bar{L}(k)$ represents a vector of time-varying covariates through time k —which includes baseline variables V . The numerator of the weight is the probability a patient is on the observed treatment at time k , given the prior treatment history and baseline covariates. The denominator is basically the same factor, except it incorporates time-varying covariates as predictors. Thus, one can see that observations where the time-varying factors are strong predictors of the current treatment selection are down-weighted in the analyses (because such observations are over-represented in the observed data).

To incorporate adjustment for early patient dropout, the same stabilized weight approach is used—except the outcome is not treatment selection but a flag variable denoting whether the patient remained in the study. The final weight for each patient’s observation is obtained by multiplying the treatment selection weights and the censoring weights.

For the second step of the MSM analysis, one simply conducts a weighted repeated measures model analysis using generalized estimating equations. In this second stage, time-dependent confounders are not included in the repeated measures model—as their effects have been incorporated into the weights. Treatment is included as a time-dependent factor, and time-invariant covariates may also be included as appropriate (just as in a cross-sectional analysis where variables may be included in a propensity model and in the analysis model).

As mentioned previously, MSMs can produce consistent estimates of the average causal treatment effects—even in the presence of treatment changes, time-dependent confounders, and missing at random study dropout. The assumptions necessary for causal inference from an MSM correspond to the same assumptions necessary for common cross-sectional bias control methods such as propensity scoring:

1. no unmeasured confounders—that is, all variables that relate to treatment assignment and outcome were collected and utilized in the analysis; called *conditional exchangeability*; formally: $Y_a(t+1) \perp\!\!\!\perp A(k) \mid \bar{A}(k-1), \bar{L}(k)$, for all \bar{a} and $t \geq k$.
2. positivity—there is a positive probability of each treatment for each set of covariates (no perfect confounding); formally $f[\bar{a}(k-1), \bar{l}(k)] > 0 \Rightarrow f[a(k) \mid \bar{a}(k-1), \bar{l}(k)] > 0$.
3. use of the correct models (weighting and analysis models).

As strong assumptions are necessary, assessing the appropriateness of the assumptions and performing sensitivity analyses are critical to a quality analysis. The no unmeasured confounders and correct models assumptions can never be fully proven and are discussed here. The positivity assumption basically says that all treatment options are possible given any combination of covariate values. Mortimer and colleagues (2005) recommended assessing this by computing predicted probabilities of treatment selection using the covariates from the models across the entire study (looking for 0 or 1 predicted probabilities). Mortimer and colleagues (2005) also provide an example of assessing the correctness of the models used in an MSM analysis using test and training data sets.

To limit the possibility of unmeasured confounding, every effort should be made to identify and collect data on potential confounders by searching the literature, examining relevant data, having discussions with experts, and utilizing potential confounding variables in the analysis. Such diligence will still never allow one to completely conclude no bias is unaccounted for, but an analyst first needs to make sure that all known confounders are addressed. Robins and colleagues (1999) and Brumback and colleagues (2004) have also provided a more formal method to study the sensitivity of an MSM analysis to unmeasured confounding. They quantify such confounding through a sensitivity parameter (α) and confounding function and assess the amount of unmeasured confounding that can be present before inferences would change. The confounding function (or α itself when a simple constant function is used) represents the difference in potential outcomes between patients in the different treatment groups. We use a simple constant function in our analysis, referring the reader to Brumback and colleagues (2004) for more options.

9.3 Example: MSM Analysis of a Simulated Schizophrenia Trial

9.3.1 Study Description

To illustrate the implementation of an MSM analysis, we simulated data based on a study of the effectiveness of medications for patients with schizophrenia in usual-care settings. A brief description of the design for the actual study follows, though the reader is referred to Tunis and colleagues (2006) for details. This was a one-year study of patients with schizophrenia, schizoaffective disorder, or schizophreniform disorder who were randomized to one of three

different treatment regimens. After randomization, the remainder of the study was observational in the sense that physicians/patients were allowed to stop or switch medications as deemed necessary in usual practice. Data on a variety of domains were captured at five post-baseline visits (approximately 2 weeks and 2, 5, 8, and 12 months post-baseline). The outcome measure of interest for this analysis was the Brief Psychiatric Rating Scale (BPRS) total score, a measure of schizophrenia symptom severity where lower scores indicate lesser symptom severity. For demonstrative purposes, we simulated data for this analysis (see Tunis et al., 2006, for actual data results) maintaining the design and data structure and focused the final comparison between two groups rather than three. These groups are referred to as *treatment* and *control* during this example analysis. However, each treatment is considered as a separate treatment in the analytical steps until the final model to demonstrate how one can handle more than two groups with the MSM approach.

9.3.2 Data Analysis

9.3.2.1 Data Overview

Before conducting the MSM analysis, we provide a brief summary of the simulated data pertaining to medication changes and the steps taken to prepare the data set for analysis. The treatment groups were balanced with respect to demographics and baseline patient characteristics due to the randomization. After randomization, study discontinuation was similar across the treatment groups though rates of switching medication differed, with almost half (47.7%) of the control group (treatment C) switching medications during the study with a lower rate for the treatment group (treatment groups A and B pooled; 20.4%). At each visit, patients were considered to be on a particular medication if they had been treated with that medication for at least the previous 14 days. Given this definition, on 41.6% of the 2,548 patient visits during the study, patients were taking treatment A, 25.1% were taking treatment B, 23.8% were taking treatment C, 4.8% were taking both A and C, and 4.8% were not taking any antipsychotic medication. A total of 17 patient visits were excluded from the analysis due to small sample sizes for certain treatment combinations (A and B $n=16$; A and B and C $n=1$). In addition, approximately 4% of the patient visits had missing covariate information (see the list of covariates later) that was imputed using a LOCF approach. Outcome data were not imputed, only covariate data and only when the outcome measure was available. The analysis data set, INPDS, used a one observation per patient per visit format. A description of the key variables follows.

Table 9.1 Description of Key Variables in MSM Analysis

Variable	Description	Variable	Description
INVSC	Investigator number	BPRS	BPRS at this visit
PATSC	Patient number	GAFC	GAF at this visit
AGEYRS	Age in years	EVNT	Events during visit
GENDER	Gender	HOSP	Hosp. during visit
ORIGIN2	Race	PR1TRTA	Trt A previous visit
THERAPY	Randomized trtmnt	PR1TRTB	Trt B previous visit
VIS	Visit number	PR1TRTC	Trt C previous visit
BAVAR	Baseline BPRS total	PR1BPRS	Previous vis BPRS
BGAF	Baseline GAF score	PR1GAF	Previous vis GAF
BEVNT	Baseline adv events	PR1EVNT	Previous vis event
BHOSP	Baseline hospitaliz.	PR1HOSP	Previous vis hosp

(continued)

Table 9.1 (continued)

Variable	Description	Variable	Description
TRTA	On Trt A during vis		
TRTB	On Trt B during vis		
TRTC	On Trt C during vis		
TRT	Treatment during vis		

9.3.2.2 Computation of Inverse Weights

Step one in conducting the MSM analysis requires estimation of the treatment selection and censoring weights (using the formula for stabilized weights [SW] from Section 9.2). For estimating the treatment selection weights, separate multinomial models for the numerator and denominator were implemented using PROC LOGISTIC with the LINK= GLOGIT option (see Program 9.1). Treatment was the dependent variable for both models and the GLOGIT option was used because the treatment choice at each visit had five potential outcomes (three individual treatments, no treatment, and one combination). Output from the LOGISTIC procedure contains the predicted probabilities of treatment selection (data sets PREDTRT0 and PREDTRT1). In addition to previous treatment, the following time-varying covariates were included in this model:

- BPRS total
- global assessment of functioning (GAF)
- events (the presence or absence of at least one moderate or severe adverse event)
- hospitalization

These variables were chosen a priori to cover the domains of symptom severity, functioning, tolerability, and resource utilization. Time-independent variables included in this initial model included age, gender, ethnicity, and baseline value for each of the time-dependent variables. Macro variables could be utilized to input all model parameters at the beginning of the code; however, we chose to simplify the understanding of the process by simply showing the models directly in the LOGISTIC statements. In addition, output from the weight models is suppressed for simplification of the output. However, one can easily remove the restriction to evaluate the weight model in more detail.

PROC GENMOD was used to compute the estimated stabilized weights to adjust for censoring using a logistic regression model (see Program 9.1). The dependent variable for the censoring weight model was a binary flag for remaining in the study. The independent variables for this model were the same as for the treatment selection weight model—though the time-varying covariates were offset by one visit due to the structure of the data as censoring looked forward (did the patient return for a following visit?) relative to treatment (what was the treatment assigned in the previous time period?). The Logit LINK function was used here because the outcome measure was binomial. The GENMOD approach could be used for both weight calculations in studies where only two treatment groups are assessed.

Partial output from the multinomial and logistic regression denominator weights models is provided in the Output from Program 9.1. Previous treatment was the strongest predictor of present treatment. None of the time-varying covariates were strong predictors of treatment changes—suggesting bias in treatment selection over time may not be particularly strong in these data. Along with previous treatment, higher (time-varying) symptom severity was found to be a predictor of censoring. Patients with more severe symptoms were more likely to discontinue from the study.

Program 9.1 Computing Treatment Selection and Censoring Weights

```

/* This section of code computes the treatment selection and censoring
weights. This is accomplished in 4 steps:
1) multinomial model to compute numerator of treatment selection weights;
2) multinomial model to compute denominator of treatment selection weights;
3) binomial model to compute numerator of censoring adjustment weights;
4) binomial model to compute denominator of censoring adjustment weights.*/

/* treatment selection weights: numerator calculation
(probability of treatment using only baseline covariates) */
PROC LOGISTIC DATA = INPDS;
  CLASS VIS THERAPY PR1TRTA PR1TRTB PR1TRTC GENDER ORIGIN2 BHOSP BEVNT;
  MODEL TRT = VIS THERAPY PR1TRTA PR1TRTB PR1TRTC GENDER ORIGIN2 BHOSP
  BEVNT AGEYRS BGAF BBPRS
  /LINK=GLOGIT;
  OUTPUT OUT=PREDTRT0 (WHERE=(TRT=_LEVEL_)) PRED=PREDTRT0;
run;

/* treatment selection weights: denominator calculation
(probability of treatment with baseline covariates and time-dependent
covariates) */
PROC LOGISTIC DATA = INPDS;
  CLASS VIS THERAPY PR1TRTA PR1TRTB PR1TRTC GENDER ORIGIN2 BHOSP BEVNT
  PR1EVNT PR1HOSP;
  MODEL TRT = VIS THERAPY PR1TRTA PR1TRTB PR1TRTC GENDER ORIGIN2 BHOSP
  BEVNT AGEYRS BGAF BBPRS PR1EVNT PR1HOSP PR1BPRS PR1GAFC
  /LINK=GLOGIT;
  OUTPUT OUT=PREDTRT1 (WHERE=(TRT=_LEVEL_)) PRED=PREDTRT1;
run;

/* censoring adjustment weights: numerator calculation
(probability of censoring using only baseline covariates) */
ODS LISTING EXCLUDE OBSTATS;
PROC GENMOD DATA = INPDS;
  CLASS PATSC VIS THERAPY TRTA TRTB TRTC GENDER ORIGIN2 BHOSP BEVNT;
  MODEL CFLAG = VIS THERAPY TRTA TRTB TRTC GENDER ORIGIN2 BHOSP BEVNT
  AGEYRS BGAF BBPRS
  /DIST = BIN LINK = LOGIT TYPE3 OBSTATS;
  REPEATED SUBJECT = PATSC / TYPE = EXCH;
  ODS OUTPUT OBSTATS = PREDCEN0 (RENAME=(PRED=PREDCEN0));
run;
ODS LISTING SELECT ALL;

/* censoring adjustment weights: denominator calculation
(probability of censoring using baseline covariates and time-dependent
covariates) */
ODS LISTING EXCLUDE OBSTATS;
PROC GENMOD DATA = INPDS;
  CLASS PATSC VIS THERAPY TRTA TRTB TRTC GENDER ORIGIN2 BHOSP BEVNT
  EVNT HOSP;
  MODEL CFLAG = VIS THERAPY TRTA TRTB TRTC GENDER ORIGIN2 BHOSP BEVNT
  EVNT HOSP AGEYRS BGAF BBPRS BPRS GAFC
  /DIST = BIN LINK = LOGIT TYPE3 OBSTATS;
  REPEATED SUBJECT = PATSC / TYPE = EXCH;
  ODS OUTPUT OBSTATS = PREDCEN1 (RENAME=(PRED=PREDCEN1));
run;
ODS LISTING SELECT ALL;

```

Output from Program 9.1

```

weight denominator model: treatment weights

Type 3 Analysis of Effects

Effect          DF      Wald
                Chi-Square    Pr > ChiSq
VIS             16      124.2743    <.0001
THERAPY         8       11.8349     0.1587
PR1TRTA         4      229.7185    <.0001
PR1TRTB         4     100.6526    <.0001
PR1TRTC         4     269.8939    <.0001
GENDER          4       3.3971     0.4937
ORIGIN2         8     10.4911     0.2322
BHOSP           4       0.4611     0.9772
BEVNT           4       3.0900     0.5429
AGEYRS          4       1.4473     0.8359
BGAF            4       1.9736     0.7406
BBPRS           4       2.2390     0.6919
PR1EVNT         4       7.5905     0.1078
PR1HOSP         4       4.6487     0.3253
PR1BPRS         4       2.5962     0.6275
PR1GAFC         4       2.5983     0.6271

weight denominator model: censoring weights

Score Statistics For Type 3 GEE Analysis

Source          DF      Chi-
                Square    Pr > ChiSq
VIS             4       13.21     0.0103
THERAPY         2        9.04     0.0109
TRTA            1        3.95     0.0468
TRTB            1        3.74     0.0532
TRTC            1        0.39     0.5346
GENDER          1        0.15     0.6964
ORIGIN2         2        0.99     0.6087
BHOSP           1        0.88     0.3473
BEVNT           1        0.93     0.3338
EVNT            1        0.61     0.4338
HOSP            1        0.05     0.8163
AGEYRS          1        4.89     0.0270
BGAF            1        0.04     0.8431
BBPRS           1        0.67     0.4115
BPRS            1       12.21     0.0005
GAFC            1        0.32     0.5726

```

To produce the overall weights for each observation (patient visit) in this analysis, the inverse probability weights for treatment selection and censoring computed here were multiplied together cumulatively in a DATA step based on the formula for SW in Section 9.2 (see Program 9.2). In the SAS code, the SQL procedure gathers data from the four output data sets from the treatment selection and weight models (PREDTRT0, PREDTRT1, PREDCENS0, and PREDCENS1) and produces the stabilized weights. Variable STABWT is the final estimate of SW. Output from Program 9.2 displays the distribution of the final weights across all patient visits in box plot form. The mean is near 1 (mean = 1.002, SD = 0.1555), as would be expected from the average of weights, and no major outliers were noted in the box plot.

Program 9.2 Merging Output from Program 9.1

/* This section of code performs the steps necessary to merge the output from the weight models (Program 9.1) to allow for computation of a single adjustment for each observation in the analysis data set (stabilized weight). This is followed by code to produce summaries of the final weights. */

```

PROC SQL;
  /*ratio of probabilities for treatment*/
  CREATE TABLE PREDTRT AS
    SELECT *, PREDTRT0/PREDTRT1 AS PREDTRT
    FROM PREDTRT1(KEEP=PATSC VIS PREDTRT1)
      NATURAL FULL JOIN
      PREDTRT0(KEEP=PATSC VIS PREDTRT0)
    ORDER PATSC, VIS
  ;
  /*ratio of probabilities for censoring*/
  CREATE TABLE PREDCEN AS
    SELECT *, PREDCEN0/PREDCEN1 AS PREDCEN
    FROM (SELECT INPUT(PATSC,BEST.) AS PATSC, INPUT(VIS,BEST.) AS VIS,
    PREDCEN0 FROM PREDCEN0)
      NATURAL FULL JOIN
      (SELECT INPUT(PATSC,BEST.) AS PATSC, INPUT(VIS,BEST.) AS VIS,
    PREDCEN1 FROM PREDCEN1)
    ORDER PATSC, VIS;
QUIT;

  /*calculate stabilized weight*/
PROC SORT DATA=INPDS OUT=WEIGHTS;
  BY PATSC VIS;
RUN;

DATA WEIGHTS;
  MERGE WEIGHTS PREDTRT PREDCEN;
  BY PATSC VIS;
  VWT=PREDTRT*PREDCEN;
  IF FIRST.PATSC THEN STABWT=VWT;
  ELSE STABWT=VWT*DUM;

  RETAIN DUM;
  DROP DUM;
  DUM=STABWT;
RUN;

  /*diagnostic plot for weights*/
DATA GRPH;
  SET WEIGHTS; /*assignment of months to visits is study-specific*/
  IF VIS = 3 THEN MONTH = 0.5;
  IF VIS = 4 THEN MONTH = 2;
  IF VIS = 5 THEN MONTH = 5;
  IF VIS = 6 THEN MONTH = 8;
  IF VIS = 7 THEN MONTH = 12;

  IF MONTH = 0.5 THEN DELETE; /*simplify plot by focusing on months
  with greater switching and thus greatest variability in weights */
RUN;

PROC SORT DATA = GRPH;
  BY MONTH;
RUN;

```


220 Analysis of Observational Health Care Data Using SAS

```
ODS RTF FILE="%SYSFUNC(PATHNAME(WORK))\FIG1.RTF";

FILENAME FIGURE "%SYSFUNC(PATHNAME(WORK))\SASGRAPH.EMF";

GOPTIONS RESET=ALL TARGET=SASEMF DEVICE=SASEMF FTEXT=DUPLEX HTEXT=.75
          CBACK=WHITE XMAX=6IN XPIXELS=1200 YMAX=5IN YPIXELS=1000
          GSFNAME=FIGURE GSFMODE=REPLACE;

SYMBOL1 COLOR=BLACK INTERPOL=JOIN
          WIDTH=2 VALUE=SQUARE
          HEIGHT=1;

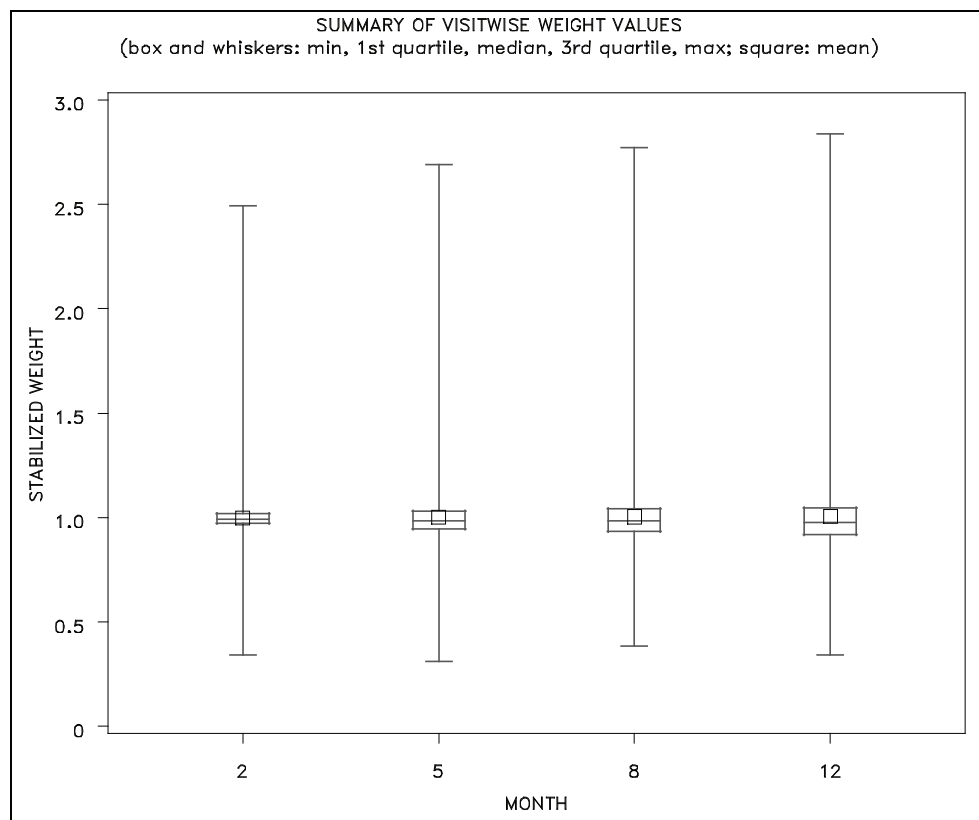
AXIS1 MINOR = NONE COLOR = BLACK LABEL=("STABILIZED WEIGHT" ANGLE=90
ROTATE=0);

PROC BOXPLOT DATA=GRPH;
  PLOT STABWT*MONTH / CFRAME = WHITE
                CBOXES = DAGR
                CBOXFILL = WHITE
                VAXIS = AXIS1;
  TITLE "SUMMARY OF VISITWISE WEIGHT VALUES";
  TITLE2 "(box and whiskers: min, 1st quartile, median, 3rd quartile, max;
square: mean)";
RUN;

GOPTIONS RESET=ALL;

ODS RTF CLOSE;
```

Output from Program 9.2



9.3.2.3 Treatment Effectiveness Analysis Model

The final step of the MSM analysis is to estimate causal treatment effects using a weighted repeated measures model with generalized estimating equations (PROC GENMOD—see Program 9.3) and an exchangeable correlation matrix. Change from baseline BPRS total score was the dependent measure in the analysis model. Independent variables for the analysis model were investigational site, age, gender, race, baseline BPRS, visit, time-varying treatment, and the treatment-by-visit interaction. The WEIGHT statement in PROC GENMOD incorporates the inverse probability weighting, which allows for the causal treatment effect estimates. The ESTIMATE statement utilized in PROC GENMOD pooled individual treatments together and produced estimated mean group differences for pooled groups (as opposed to comparing individual treatment groups). This portion of the code is not necessary for many applications. Output from Program 9.3 displays a summary of the final model results and a figure displaying the least squares means at each visit. Results showed a statistically significant treatment difference favoring the treatment group (pooled treatments A and B) relative to the control group, with an estimated average treatment difference in BPRS changes of 1.8 [0.4, 3.2], $p=.015$ across the 1-year period. Though treatment differences grew numerically over time, the treatment-by-visit interaction term was not statistically significant ($p=.158$).

Program 9.3 Running Final Analysis Model

```
/* This section of code runs the final analysis model using a weighted
repeated measures approach. The results are presented graphically using
PROC GPLOT.*/
```

```
/*final analysis model*/
PROC GENMOD DATA = WEIGHTS;
  CLASS VIS PATSC GENDER ORIGIN2 INVSC TRT;
  WEIGHT STABWT;
  MODEL CAVAR = INVSC BAVAR VIS AGEYRS GENDER ORIGIN2 TRT VIS*TRT
    / DIST=NORMAL LINK=ID TYPE3;
  REPEATED SUBJECT = PATSC / TYPE=EXCH;
  LSMEANS TRT VIS*TRT / PDIFF;
  TITLE 'MSM FINAL ANALYSIS MODEL';
  ESTIMATE 'A+B VS C'
    TRT 0 .5 .5 -1 0;
  ESTIMATE 'A+B VS C AT VIS 3'
    TRT 0 .5 .5 -1 0
    VIS*TRT 0 .5 .5 -1 0      0 0 0 0 0      0 0 0 0 0
            0 0 0 0 0      0 0 0 0 0;
  ESTIMATE 'A+B VS C AT VIS 4'
    TRT 0 .5 .5 -1 0
    VIS*TRT 0 0 0 0 0      0 .5 .5 -1 0      0 0 0 0 0
            0 0 0 0 0      0 0 0 0 0;
  ESTIMATE 'A+B VS C AT VIS 5'
    TRT 0 .5 .5 -1 0
    VIS*TRT 0 0 0 0 0      0 0 0 0 0      0 .5 .5 -1 0
            0 0 0 0 0      0 0 0 0 0;
  ESTIMATE 'A+B VS C AT VIS 6'
    TRT 0 .5 .5 -1 0
    VIS*TRT 0 0 0 0 0      0 0 0 0 0      0 0 0 0 0
            0 .5 .5 -1 0      0 0 0 0 0;
  ESTIMATE 'A+B VS C AT VIS 7'
    TRT 0 .5 .5 -1 0
    VIS*TRT 0 0 0 0 0      0 0 0 0 0      0 0 0 0 0
            0 0 0 0 0      0 .5 .5 -1 0;
  ODS OUTPUT LSMEANS=LSMEANS;
RUN;
```

222 Analysis of Observational Health Care Data Using SAS

```
/*LS means plot for the final model*/
DATA LSMEANS2;
  SET LSMEANS;
  WHERE TRT IN ('A ', ' B ', ' C');
  IF TRT IN ('A ', ' B ') THEN TRT2='A+B';
  ELSE TRT2='C  ';
RUN;

PROC SQL;
  CREATE TABLE LSMEANS3 AS
  SELECT TRT2 AS TRT, VIS, MEAN(ESTIMATE) AS ESTIMATE
  FROM LSMEANS2
  GROUP TRT2, VIS;
QUIT;

ODS RTF FILE="%SYSFUNC(PATHNAME(WORK))\FIG2.RTF";

FILENAME FIGURE "%SYSFUNC(PATHNAME(WORK))\SASGRAPH.EMF";

GOPTIONS RESET=ALL TARGET=SASEMF DEVICE=SASEMF FTEXT=DUPLEX HTEXT=.75
CBACK=WHITE
        XMAX=6IN XPIXELS=1200 YMAX=5IN YPIXELS=1000 GSFNAME=FIGURE
GSFMODE=REPLACE;

AXIS1 MINOR = NONE COLOR = BLACK LABEL=(ANGLE=90 ROTATE=0 "CHANGE IN BPRS
TOTAL SCORE");

SYMBOL1 I=JOIN W=2 L=1 C=RED V=SQUARE;
SYMBOL2 I=JOIN W=2 L=2 C=BLACK V=CIRCLE;

PROC GPLOT DATA=LSMEANS3;
  PLOT ESTIMATE*VIS=TRT/VAXIS=AXIS1;
  TITLE "MSM ESTIMATED MEAN CHANGE FROM BASELINE BPRS SCORES";
  LABEL VIS="VISIT";
  LABEL TRT="TREATMENT";
RUN;

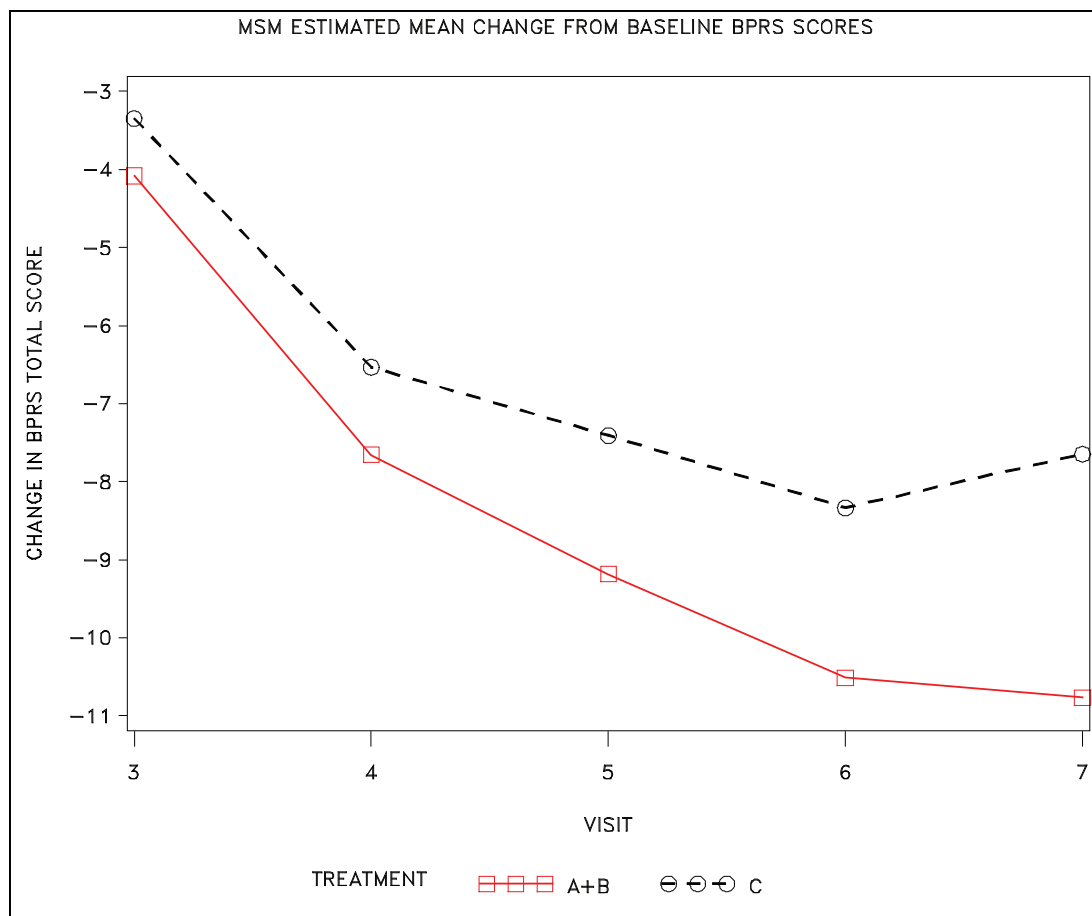
GOPTIONS RESET=ALL;

ODS RTF CLOSE;
```

Output from Program 9.3

MSM FINAL ANALYSIS MODEL							
Score Statistics For Type 3 GEE Analysis							
Source	DF	Chi-Square	Pr > ChiSq				
INVSC	19	84.25	<.0001				
BAVAR	1	71.12	<.0001				
VIS	4	37.69	<.0001				
AGEYRS	1	0.80	0.3725				
GENDER	1	0.09	0.7624				
ORIGIN2	2	0.86	0.6500				
TRT	4	11.70	0.0197				
VIS*TRT	16	21.57	0.1576				

Contrast Estimate Results							
Label	Estimate	Standard Error	Alpha	Confidence Limits		Chi-Square	Pr > ChiSq
A+B VS C	-1.7852	0.7321	0.05	-3.2201	-0.3504	5.95	0.0147
A+B VS C AT VIS 3	-0.7336	0.9615	0.05	-2.6181	1.1509	0.58	0.4455
A+B VS C AT VIS 4	-1.1225	0.9417	0.05	-2.9682	0.7232	1.42	0.2333
A+B VS C AT VIS 5	-1.7746	1.2532	0.05	-4.2307	0.6815	2.01	0.1567
A+B VS C AT VIS 6	-2.1785	1.2752	0.05	-4.6778	0.3208	2.92	0.0876
A+B VS C AT VIS 7	-3.1169	1.3185	0.05	-5.7012	-0.5327	5.59	0.0181



9.3.2.4 Sensitivity Analysis

To assess the robustness of the findings, one should evaluate the plausibility of the three main assumptions for causality (no unmeasured confounding, positivity, correct model) as well as implementing other statistical methods that are based on different assumptions. Regarding the unmeasured confounding assumption, effort was made to collect and incorporate information from experts and the literature on potential confounders prior to the analysis. The variables included in the model were selected in order to cover the domains of symptom severity, functioning, tolerability, and resource use burden. To assess the potential impact of unmeasured confounding on the results, we implemented a simple, unmeasured confounding function (we chose a constant function α) based on Brumback and colleagues (2004). Using this function, a missing confounder resulting in a shift of 0.45 (in BPRS total score) in potential outcomes on the BPRS scale favoring patients with high probability of being in the treatment group would result in a loss of the statistically significant finding. In other words, our finding of a significant difference between groups depends upon the assumption that such a confounder (or a group of confounders combining to have the same effect) does not exist. Multiple constants were evaluated using the ALPHA=c statement in the SAS code (see Program 9.4). The value of 0.45 was retained because this was the smallest value producing a p-value of approximately 0.05. While it is challenging to interpret the implications of a particular α , such an effect (<1 point on BPRS) does not appear to be extremely large, and the existence of such a confounder is certainly possible in an actual application. However, it is greater than the observed confounding (the difference between unweighted and weighted analyses) of 0.29 in the opposite direction observed in this study. Regardless of the sensitivity analysis results, one can simply not dismiss the possibility of unaccounted for factors that would result in this analysis failing to produce an estimate of a causal treatment effect. Thus, one must view the results with some caution. Different sensitivity functions can easily be tested using the provided SAS code by altering the calculation of the variables CAVAR_SENS and SAPT.

Program 9.4 Finding Level of Confounding That Would Eliminate Treatment Difference

```

/* This section of code examines the sensitivity of the results to
unmeasured confounding by finding the level of confounding that would
eliminate the observed treatment difference.*/

/*sensitivity analysis per Brumback et al. (2004) - constant
function alpha is used*/
DATA WEIGHTSS;
  SET WEIGHTS;
  BY PATSC VIS;
  IF TRT IN ('A__','_B_') THEN SNSTRT = 1;
                                ELSE SNSTRT = -1;
  SAVT = -SNSTRT*(1 - PREDTRT1);
  IF FIRST.PATSC THEN SAPT = SAVT;
                                ELSE SAPT = SAVT + DUM;
  RETAIN DUM;
  DROP DUM;
  DUM=SAPT;
  ALPHA = 0.45;
  CAVAR_SENS = CAVAR - ALPHA*SAPT;
RUN;

PROC GENMOD DATA = WEIGHTSS;
  CLASS VIS PATSC GENDER ORIGIN2 INVSC TRT;
  WEIGHT STABWT;
  MODEL CAVAR SENS = INVSC BAVAR VIS AGEYRS GENDER ORIGIN2 TRT VIS*TRT
          / DIST=NORMAL LINK=ID TYPE3;
  REPEATED SUBJECT = PATSC / TYPE=EXCH;
  LSMEANS TRT VIS*TRT / PDIFF;
  TITLE 'MSM SENSITIVITY ANALYSIS: ALPHA = 0.45';

```

```

ESTIMATE 'A+B VS C'
  TRT 0 .5 .5 -1 0;
ESTIMATE 'A+B VS C AT VIS 3'
  TRT 0 .5 .5 -1 0
  VIS*TRT 0 .5 .5 -1 0      0 0 0 0 0      0 0 0 0 0
           0 0 0 0 0      0 0 0 0 0;
ESTIMATE 'A+B VS C AT VIS 4'
  TRT 0 .5 .5 -1 0
  VIS*TRT 0 0 0 0 0      0 .5 .5 -1 0      0 0 0 0 0
           0 0 0 0 0      0 0 0 0 0;
ESTIMATE 'A+B VS C AT VIS 5'
  TRT 0 .5 .5 -1 0
  VIS*TRT 0 0 0 0 0      0 0 0 0 0      0 .5 .5 -1 0
           0 0 0 0 0      0 0 0 0 0;
ESTIMATE 'A+B VS C AT VIS 6'
  TRT 0 .5 .5 -1 0
  VIS*TRT 0 0 0 0 0      0 0 0 0 0      0 0 0 0 0
           0 .5 .5 -1 0      0 0 0 0 0;
ESTIMATE 'A+B VS C AT VIS 7'
  TRT 0 .5 .5 -1 0
  VIS*TRT 0 0 0 0 0      0 0 0 0 0      0 0 0 0 0
           0 0 0 0 0      0 .5 .5 -1 0;

RUN;

```

Output from Program 9.4

MSM SENSITIVITY ANALYSIS: ALPHA = 0.45							
Contrast Estimate Results							
Label	Estimate	Standard Error	Alpha	Confidence Limits		Chi-Square	Pr > ChiSq
A+B VS C	-1.3911	0.7303	0.05	-2.8224	0.0402	3.63	0.0568
A+B VS C AT VIS 3	-0.6203	0.9615	0.05	-2.5048	1.2642	0.42	0.5188
A+B VS C AT VIS 4	-0.8879	0.9417	0.05	-2.7335	0.9577	0.89	0.3457
A+B VS C AT VIS 5	-1.3184	1.2534	0.05	-3.7750	1.1383	1.11	0.292
A+B VS C AT VIS 6	-1.6479	1.2726	0.05	-4.1422	0.8464	1.68	0.1954
A+B VS C AT VIS 7	-2.4810	1.3143	0.05	-5.0569	0.0950	3.56	0.0591

To assess the positivity assumption, we followed the ideas of Mortimer and colleagues (2005) and estimated the probability of selection of each of the five treatment possibilities using all possible covariates across all study visits. The CATMOD procedure is used here in order to generate predicted probabilities for all possible treatment options. While theoretically there were no issues with the positivity assumption in this study (all patients had the opportunity of being switched from and to any combination of treatments at any time), the goal was to assess whether any observed set of covariates produced a predicted probability of 0 or 1 for this set of data. This was done as shown with Program 9.5 where we computed the predicted probabilities of treatment for each observed set of covariates and then summarized the predicted probabilities (summary statistics using PROC MEANS by treatment). We observed that the smallest (nonzero but $<.00015$) probabilities were for the theoretical switch from no treatment to the combination of treatments A and C at earlier study visits. In general, the smaller probabilities were associated with the no treatment and combination treatment groups—as would be expected given their observed frequencies. Sensitivity analyses were then performed without the more extreme records and then with combination patients re-assigned to the initial randomized single treatment (for example, a patient randomized to treatment A but treated with A and C would be counted as being treated only with treatment A). No major changes in the inferences were observed from these sensitivity analyses and these results are not shown.

Program 9.5 Examining Positivity Assumption for Causal Inference

```

/* This section of code examines the positivity assumption for causal
inference by examining the predicted values of all possible treatment
changes. Summary statistics are presented as well as a listing to allow
examination of outliers.*/

/*positivity check Mortimer (2005) - generate predicted probabilities*/
ODS LISTING CLOSE;
PROC CATMOD DATA = INPDS;
  DIRECT AGEYRS BGAF BBPRS PR1BPRS PR1G AFC;
  MODEL TRT = VIS THERAPY PR1TRTA PR1TRTB PR1TRTC GENDER ORIGIN2 BHOSP
    BEVNT AGEYRS BGAF BBPRS PR1EVNT PR1HOSP PR1BPRS PR1G AFC
    /PRED=PROB;
  ODS OUTPUT PREDICTEDPROBS=PREDTRT1POS (KEEP=VIS TRT THERAPY PR1TRTA
    PR1TRTB PR1TRTC GENDER ORIGIN2 BHOSP BEVNT AGEYRS BGAF
    BBPRS PR1EVNT PR1HOSP PR1BPRS PR1G AFC
    SAMPLE OBSFUNCTION PREDFUNCTION
    RENAME=(PREDFUNCTION=PREDTRT1));
RUN;
ODS LISTING;

ODS RTF FILE="%SYSFUNC(PATHNAME(WORK))\FIG3.RTF";

FILENAME FIGURE "%SYSFUNC(PATHNAME(WORK))\SASGRAPH.EMF";

GOPTIONS RESET=ALL TARGET=SASEMF DEVICE=SASEMF FTEXT=DUPLEX HTEXT=.75
  CBACK=WHITE XMAX=6IN XPIXELS=1200 YMAX=5IN YPIXELS=1000
  GSFNAME=FIGURE GSFMODE=REPLACE;

PROC UNIVARIATE DATA = PREDTRT1POS;
  VAR PREDTRT1;
  HISTOGRAM;
  PROBLOT;
  TITLE 'SUMMARY STATS ON PREDICTED VALUES';
RUN;

GOPTIONS RESET=ALL;

ODS RTF CLOSE;

PROC MEANS DATA = PREDTRT1POS;
  CLASS TRT;
  VAR PREDTRT1;
  TITLE 'Positivity Check: Summary Stats on Predicted Values by Observed
  Treatment';
RUN;

PROC SORT DATA = PREDTRT1POS;
  BY PREDTRT1;
RUN;

PROC PRINT DATA = PREDTRT1POS;
  TITLE 'Positivity Check: Sorted Listing of Predicted Values';
RUN;

  * (PROC PRINT output not shown but is used to identify individual
  patients with combinations of covariates with extreme values *;

```

Output from Program 9.5

Positivity Check: Summary Stats on Predicted Values by Observed Treatment						
The MEANS Procedure						
Analysis Variable : PREDTRT1 Predicted: Probability						
TRT	N Obs	N	Mean	Std Dev	Minimum	Maximum
A_C	2547	2547	0.0475067	0.1254780	0	0.9758886
A__	2547	2547	0.4157827	0.4165088	4.0755123E-7	0.9999826
B	2547	2547	0.2508929	0.3999598	1.8471036E-8	0.9999993
__C	2547	2547	0.2383188	0.3699969	3.4082211E-7	0.9998542
___	2547	2547	0.0474988	0.1179316	0	0.9817690

In addition to the assessment of the no unmeasured confounders and positivity assumptions, the appropriateness of the models was assessed by adding interactions, quadratic terms, and other potential confounder variables to the models. No significant changes to the outcome were noted. The model assessment was done by adjusting the appropriate models in the SAS code—which could be automated by putting the model variables as macro variables at the top of the code. In addition, we analyzed the data using other methods: an ITT LOCF analysis (using all data), a repeated measures model (excluding data after a medication switch) and an epoch analysis (see Chapter 8). Analysis of these simulated data using a simple ITT LOCF analysis failed to show a significant difference between treated and control groups (estimated difference of 1.16, $p=.334$). The repeated measures mixed model (discarding the data after switching) and epoch analyses results were fairly similar to the MSM approach. The disagreement between the ITT analysis and other analyses appears to be a result of ignoring treatment information for the data after a treatment switch, as discussed by Faries and colleagues (2007). Switchers from treatment C performed well after the switch to treatments A and B, information not attributed to treatments A and B in the ITT analysis.

9.4 Discussion

This chapter has presented the issue of assessing the effects of treatment in longitudinal, observational data—with a focus on addressing treatment switching using MSM. We were interested in the performance of MSM because this approach utilizes all of the study data and produces consistent estimates of the causal effect of treatments, even when there are treatment switching, missing data, and time-varying confounders. Validity of the MSM analysis rests on three key assumptions:

1. no unmeasured confounding
2. positivity
3. correct models

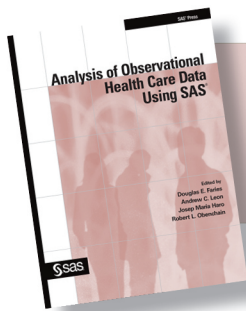
Also, the missing data are assumed to follow a missing completely at random (MCAR) or missing at random (MAR) pattern. Thus, well-planned sensitivity analyses are important. This should include an assessment of the assumptions supporting the causal effect estimation by the MSM as well as use of other methodology supported by differing assumptions. In addition, presentation of

results should make it clear to the reader that causal interpretation of the results rests on unverifiable assumptions as well as being transparent about which steps were taken to assess the robustness of the results.

In summary, marginal structural models are a promising approach for estimating the causal effects of treatment in longitudinal, observational data. Other chapters in this book provide details on alternative methodologies.

References

- Bodnar, L. M., M. Davidian, A. M. Siega-Riz, A. A. Tsiatis. 2004. "Marginal structural models for analyzing causal effects of time-dependent treatments: an application in perinatal epidemiology." *American Journal of Epidemiology* 159: 926–934.
- Brumback, B. A., M. A. Hernán, S. J. Haneuse, and J. M. Robins. 2004. "Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures." *Statistics in Medicine* 23(5): 749–767.
- Cole, S. R., M. A. Hernán, J. B. Margolick, M. H. Cohen, and J. M. Robins. 2005. "Marginal structural models for estimating the effect of highly active antiretroviral therapy initiation on CD4 cell count." *American Journal of Epidemiology* 162: 471–478. Epub 2005 Aug 2.
- Faries, D., H. Ascher-Svanum, and M. Belger. 2007. "Analysis of treatment effectiveness in longitudinal observational data." *Journal of Biopharmaceutical Statistics* 17(5): 809–826.
- Grimes, D. A., and K. F. Schulz. 2002. "Bias and causal associations in observational research." *The Lancet* 359: 248–252.
- Haro, J. M., S. Kontodimas, M. A. Negrin, M. Ratcliffe, D. Suarez, and F. Windmeijer. 2006. "Methodological aspects in the assessment of treatment effects in observational health outcomes studies." *Applied Health Economics & Health Policy* 5(1): 11–25.
- Hernán, M. A., B. Brumback, and J. M. Robins. 2000. "Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men." *Epidemiology* 11: 561–570.
- Hernán, M. A., B. Brumback, and J. M. Robins. 2002. "Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures." *Statistics in Medicine* 21: 1689–1709.
- Hernán, M. A., J. M. Robins, and L. A. Garcia Rodrigues. 2005. Comment on: Prentice, R. L., M. Pettinger, and G. L. Anderson. Discussion on "Statistical issues arising in the women's health initiative." *Biometrics* 61: 922–941.
- Hernán, M. A., S. Hernández-Díaz, and J. M. Robins. 2004. "A structural approach to selection bias." *Epidemiology* 15(5): 615–625.
- Ko, H., J. W. Hogan, and K. H. Mayer. 2003. "Estimating causal treatment effects from longitudinal HIV natural history studies using marginal structural models." *Biometrics* 59: 152–162.
- Mallinckrodt, C. H., T. M. Sanger, S. Dubé, D. J. DeBrotta, G. Molenberghs, R. J. Carroll, W. Z. Potter, and G. D. Tollefson. 2003. "Assessing and interpreting treatment effects in longitudinal clinical trials with missing data." *Biological Psychiatry* 53(8): 754–760.
- Mortimer, K. M., R. Neugebauer, M. van der Laan, and I. B. Tager. 2005. "An application of model-fitting procedures for marginal structural models." *American Journal of Epidemiology* 162(4): 382–388. Epub 2005 Jul 13.



From *Analysis of Observational Health Care Data Using SAS*[®]. Full book available for purchase [here](#).

Index

A

agglomerative clustering methods 179
Akaike Information Criterion (AIC) 265, 280
ALPHA= statement, DATA step 224
ANOVA model
 evaluating propensity scores 27–28
 evaluating treatment differences 107
 imputation strategies and 126
 LC analysis and 155
 PSBB and 317
association
 causation vs. 5–7
 defined 5
assumption of exchangeability 88

B

B/B (blocking/balancing) score
 cluster membership and 187–189
 defined 187
balancing score
 See also propensity score (PS)
 assessing balance in baseline characteristics 55–58
 defined 12, 55, 187
 evaluating across treatment groups 205–207
Bayesian simulation method 364
bias
 See also selection bias
 hidden 7
 in observational research 385–390
 overt 7
 sample size and 9, 343
binary outcomes 398–400, 411–414, 419–422
blind assessments 388
blocking, key roles played 152
%BOOT macro 122–123
%BOOTCI macro 122–123
BOUNDS statement, NLP procedure 245–246
BOXPLOT procedure 36
BPRS (Brief Psychiatric Rating Scale) 266–267, 278
BY statement 102

C

CA (covariate adjustment) method
 defined 183
 LC analysis and 165
 mortality rates analyses 157–158, 160
 of propensity scores 62
 RCTs and 347, 356
Cardiac Care Network (CCN) 62–65
case-control design 9
CATMOD procedure 225
causation, association vs. 5–7
CC (complete covariate) method
 bootstrap confidence intervals 123
 defined 106
 IPW estimation with missing values 109–110
CCN (Cardiac Care Network) 62–65
CD Trial 351–358
CDF (cumulative distribution function) 153–154
CEA
 See cost-effectiveness analysis
CEAC (cost-effectiveness acceptability curve) 344–346, 355–356
censoring
 cost-effectiveness analysis and 363–382
 induced informative 364
 parameter estimation and 347–351
CLASS statement
 DR considerations 102
 INB example 359
 UNIVARIATE procedure 136
Clinical Global Impression scale 389
clustering
 alternatives to try 178
 defined 154
 JMP considerations 168–170
 review of concepts 178–179
 sensitivity analysis and 178–179
 treatment effects 155
cohort studies 9
COMMON_SUPPORT option, WTMODEL statement 90, 92, 97–98
complete covariate method
 See CC (complete covariate) method

confidence intervals
 bootstrap method 122–123
 dose-response analyses 307
 ICER and 368
 nonparametric bootstrapping and 318
 PSBB and 323

confounding
 defined 7
 examples of 7–8
 in observational studies 7–8
 research checklist 289, 292–293
 unmeasured 29, 214

continuous outcomes
 calculating sample size 413, 416–419
 DR macro and 92
 longitudinal data 392–398
 propensity score and 59

correlation 5–7, 415

cost-effectiveness acceptability curve (CEAC)
 344–346, 355–356

cost-effectiveness analysis
 about 363–365
 incremental net benefit 339–362
 propensity score bin bootstrapping
 315–337
 with censored data 363–382

counterfactual causal effect 6

COVARIANCE= option, NLP procedure 245,
 248

covariate adjustment
See CA (covariate adjustment) method

Cox proportional hazards model
 dose-response analyses 297, 380
 propensity score matching and 61, 78–80

cross-sectional studies
 defined 9
 general design issues 386

cumulative distribution function (CDF)
 153–154

D

DATA= option, NLP procedure 245

DATA step
 ALPHA= statement 224
 creating residuals 255

databases
 dose-response analyses 295–311
 good research practices 289–294
 retrospective 287–288

defibrillator study 365, 369–380

dependent variables
 confounding and 7
 in logistic regression 25
 missing patterns 414–422

DES/BMS safety and efficacy study 62–80

dichotomous outcomes 59–60, 93, 99–101

DISCRETE option, MODEL statement 144

DIST= option, MODEL statement 90–91

divisive clustering methods 179

dose-response analyses 295–311

doubly robust (DR) estimation
 assumptions 87–88
 conceptual overview 86
 defined 85
 implementing DR macro 88–94
 limitations 101–102
 practical considerations 102
 sample analysis 95–101
 statistical expression 87–88

DR macro
 output from 89–94
 specifying outcome regression models 89
 specifying weight model 88–89

E

eCDF (empirical cumulate distribution function)
 174, 177

EM/ECM algorithm 106

empirical cumulate distribution function (eCDF)
 174, 177

end-stage renal disease (ESRD) study 296–310

endogeneity
See selection bias

erythropoiesis-stimulating agent (ESA)
 296–297

ESRD (end-stage renal disease) study 296–310

ESTIMATE statement, GENMOD procedure
 221

estimating
 ICER 368
 in structured nested models 236–239
 mean cost 366–368
 parameters 346–351
 propensity scores 24–25, 52–53, 64–65
 treatment effect 26–27, 58–61
 with RMLPS 265

Euro-QOL %D scale 389

evaluating propensity scores 27–28

exchangeability, assumption of 88

experimental studies
 defined 3

observational studies vs. 3–5
 exposure
See also independent variables
 DR analysis of probability 97
 effect on outcomes 74–80
 treatment and 3*n*

F

FDA (Food and Drug Administration) 288
 FHS (Framingham Heart Study) 5
 Fieller's Theorem 368, 380
 FMI (fraction of missing information) 126
 Food and Drug Administration (FDA) 288
 fraction of missing information (FMI) 126
 Framingham Heart Study (FHS) 5
 FREQ procedure
 IV method for addressing selection bias
 134, 139
 propensity score example 33
 PSBB and 323

G

G-computation formula (Robins) 264
 gamma distribution 333, 351, 354–356
 GEE (generalized estimating equation) 60,
 304–305, 410
 generalized models 25, 316
 GENISOS (Genetics vs. Environment In
 Scleroderma Outcome Study)
 419–422
 GENMOD procedure
 dose-response analyses 300, 304, 307
 doubly robust estimation 100
 ESTIMATE statement 221
 INB example 359
 IV method for addressing selection bias
 148
 LINK=LOGIT option 216
 propensity score example 36–37
 PSBB and 318, 320, 323
 schizophrenia study 275
 WEIGHT statement 221
 GLM procedure
 calculating least squares means 138
 creating residuals 254–255
 IV method for addressing selection bias
 147–148
 PSBB and 318, 320
 SNM example 253
 standardized weights and 109

GLOGIT option, MODEL statement 128
 %GMATCH macro 65
 goodness-of-fit testing 25
 greedy matching method 53–54
 GROUPS= option, RANK procedure 32

H

hazard ratio (HR) 296, 304, 307
 Health Collaborative Depression Study (NIH)
 198–207
 health maintenance organization (HMO) 134,
 136
 health-related quality of life (HRQOL) 232,
 239
Health Services Research journal 132
 hemodialysis study 296–310
 heteroskedasticity 168
 hidden biases 7
 HIPAA 288
 HMO (health maintenance organization) 134,
 136
 homoskedasticity 168
 HR (hazard ratio) 296, 304, 307
 HRQOL (health-related quality of life) 232,
 239

I

ICD-9 code system 298
 ICER (incremental cost-effectiveness ratio)
 316–323, 340–345, 368–379
 IML procedure 259–260
 INB (incremental net benefit)
 about 339–340
 CD Trial 351–358
 cost-effectiveness analysis 341–346, 365
 defined 316
 observational studies 359
 parameter estimation 346–351
 incremental cost-effectiveness ratio (ICER)
 316–323, 340–345, 368–379
 incremental net benefit
 See INB (incremental net benefit)
 IND (indicator variable) method
 defined 106
 IPW estimation with missing values
 110–111
 independent variables
 confounding and 7
 in logistic regression 25
 missing patterns 414–422

indicator variable (IND) method
 defined 106
 IPW estimation with missing values
 110–111

induced informative censoring 364

instrumental variable method
See IV (instrumental variable) method

International Society for Pharmacoeconomics
 and Outcomes Research (ISPOR)
 288, 316–317

inverse probability weight approach
See IPW (inverse probability weight)
 approach

IPW (inverse probability weight) approach
 CA method and 158, 160
 defined 183
 dose-response analyses 300–304
 doubly robust estimation and 86
 estimating treatment differences 107
 handling extreme weights 127–128
 mortality rates analyses 162–163
 sensitivity analysis 124
 with missing values 109–123

ISPOR (International Society for
 Pharmacoeconomics and Outcomes
 Research) 288, 316–317

ITT LOCF analyses 211–212, 227

IV (instrumental variable) method
 addressing selection bias 131–150
 applying to case study 139–143
 case study description 134–137
 challenges 147–148
 clustering support 155
 defined 12, 183
 least squares regression method 138
 overview 131–133
 QLIM procedure and 143–146
 regression adjustment method comparison
 146–147
 sensitivity analysis and 13

J**JMP**

See also LC (local control) analysis

Analyze menu 166, 180

clustering considerations 168–170

launching files 166

LTD distribution considerations 167–182

Next Number of Clusters dialog box 169

Open Data File dialog box 167

Select Columns dialog box 167, 175

K

Kaplan-Meier estimator 366, 370, 380

Kaplan-Meier survival curves 60–61, 76–77

L

LATEs (Local Average Treatment Effects) 155

LC (local control) analysis
 defined 183
 determining distribution saliency 164,
 174–177
 fundamental concepts 153–154
 identifying baseline characteristics 165,
 179–182
 patient registry data analysis 156–163
 performing sensitivity analysis 165,
 178–179
 problems with randomization 152–153
 revealing bias 163–174
 statistical methods for 154–155
 tactical phases 163–182

LCLF Study 239–252

least squares regression method 138

LIFETEST procedure 370

likelihood ratio test 198, 203

LINK=GLOGIT option, LOGISTIC procedure
 216

LINK=LOGIT option, GENMOD procedure
 216

Local Average Treatment Effects (LATEs) 155

local control analysis
See LC (local control) analysis

local treatment differences
See LTD distributions

log-likelihood 239

log-rank test 405–409

LOGISTIC procedure
 creating residuals 254–255
 IV method for addressing selection bias
 148
 LINK=GLOGIT option 216
 MODEL statement 128
 propensity score example 32–33
 schizophrenia study 267, 271

logistic regression
 DR estimation and 89–90, 95–96
 estimating propensity scores 25, 27–28
 mixed-effects model 196, 199–202

longitudinal observational study
 between-group comparison 410–422
 continuous outcomes 392–398

- defined 386
- model of propensity for treatment 195–209
- NIH Health Collaborative Depression Study 198–207
- regression models 263–283
- sensitivity analysis in 224–227
- treatment effectiveness analyses 197–198, 203–205, 211
- two-stage propensity adjustment 195–209
- LTD distributions
 - defined 152–154
 - determining saliency 174–177
 - identifying baseline characteristics 179–182
 - in LC tactical phases 164–165
 - JMP considerations 167–182
 - mortality rates analyses 161
 - systemic sensitivity analyses 178–179
- M**
- MADIT (Multicenter Automatic Defibrillator Implantation) 365, 369–380
- Mantel-Haenszel procedure 198
- Mantel-Haenszel test 398, 400–405
- many-to-one matching 55, 57
- marginal structural model
 - See* MSM (marginal structural model)
- %MATCH macro 65
- matched sets
 - defined 53
 - forming for propensity scores 53–55
- maximum likelihood data analysis 238–252
- Mayo Clinic case study 65–68
- McNemar's test 59–61, 75
- MEANS procedure
 - propensity score example 37
 - standardized weights and 109
 - summary statistics and 225
- measurement bias 385–390
- medical claims databases 287–294
- medication effectiveness study 30–46
- MEDLINE database 317
- Meta-analysis Of Observational Studies in Epidemiology (MOOSE) 14
- methods
 - See* statistical methods
- MI (multiple imputation) method
 - bootstrap confidence intervals 123
 - CEA with censored data and 364
 - defined 106
 - IPW estimation with missing values 111–115, 120
- MI procedure 112–113
- MIANALYZE procedure
 - calculating FMI 126
 - IPW estimation with missing values 113–115, 121
- Microsoft EXCEL 356
- MIMP (multiple imputation missingness pattern) method
 - bootstrap confidence intervals 123
 - defined 107
 - IPW estimation with missing values 120–122
 - sensitivity analysis 124
- missing patterns 414–422
- missing values
 - CC analysis 109–110
 - CEA with censored data 364
 - censoring and 364
 - classifying 364
 - data quality and 288
 - examples handling 107–108
 - IND analysis 110–111
 - IPW estimation with 109–123
 - MI analysis 111–115
 - MIMP analysis 120–122
 - MP analysis 115–119
 - MSM methodology and 213, 227
 - propensity scoring with 105–130
 - research checklist 289–291
- missingness pattern (MP) method
 - See* MP (missingness pattern) method
- MIXED procedure
 - IPW estimation with missing values 113
 - PSBB and 318, 320
 - standardized weights and 108–109
- MODEL statement
 - DISCRETE option 144
 - DIST= option 90–91
 - DR considerations 102
 - GLOGIT option 128
 - schizophrenia study 280
 - SELECT option 146
 - specifying outcome regression models 89
 - specifying weigh model 88–89
- Monte Carlo simulations 62
- MOOSE (Meta-analysis Of Observational Studies in Epidemiology) 14
- MORE (Multiple Outcomes of Raloxifene Evaluation) 107

mortality rates analyses
 CA method for 157–158
 estimated propensity score and 158–161
 IPW model 162–163

MP (missingness pattern) method
 bootstrap confidence intervals 123
 defined 106
 IPW estimation with missing values
 115–119
 sensitivity analysis 124

MSM (marginal structural model)
 defined 12
 dose-response analyses 295–311
 LCLF Study and 239
 methodology overview 213–214
 notation for 212
 schizophrenia study 214–227
 sensitivity analysis and 13
 treatment effectiveness analysis model
 221–223

Multicenter Automatic Defibrillator
 Implantation (MADIT) 365, 369–380

multiple imputation method
See MI (multiple imputation) method

multiple imputation missingness pattern method
See MIMP (multiple imputation missingness
 pattern) method

Multiple Outcomes of Raloxifene Evaluation
 (MORE) 107

N

National Institutes of Health (NIH)
 Framingham Heart Study 5
 Health Collaborative Depression Study
 198–207

Nelson-Aalen estimator 406

neural networks 25

Newton-Raphson algorithm 411

NIH (National Institutes of Health)
 Framingham Heart Study 5
 Health Collaborative Depression Study
 198–207

NLP procedure
 BOUNDS statement 245–246
 COVARIANCE= option 245, 248
 DATA= option 245
 LCLF Study and 243–252
 loading starting values 254
 maximizing log-likelihood 256–259
 output considerations 248–251
 PARMS statement 245

PCOV option 245
 VARDEF= option 245

NNT (number needed to treat) 59

non-experimental studies
See observational studies

nonparametric bootstrapping 318

nonparametric density estimates 56

NPARIWAY procedure 317

number needed to treat (NNT) 59

O

observational studies
See also longitudinal observational study

addressing research bias 385–390

association vs. causation 5–7

confounding in 7–8

cost estimation in 363–364

defined 3, 232

experimental studies vs. 3–5

general approaches to data analysis
 183–184

general design issues 386–387

good research practices 287–294

incremental net benefit 359

issues in 5–9

methods 10–13

replicability of 8–9

reporting guidelines 13–14

sample size calculation 391–425

selection bias in 7–8, 391

sensitivity analysis and 13

study design 9–10

Type I error in 8–9

observer bias 386–387

odds ratio
 discouraging use of 60
 propensity score example 33–34

ODS OUTPUT statement 244–245

one-to-one matching 55

optimal matching method 54

outcome regression
 doubly robust estimation and 86
 specifying models 89

outcomes
See also continuous outcomes
See also dependent variables

binary 398–400, 411–414, 419–422

dichotomous 59–60, 93, 99–101

effect of exposures on 74–80

time-to-event 60–61

OUTPUT statement 144
overt biases 7

P

pair matching on propensity scores 55
paired *t*-test
 continuous outcomes and 59
 IV method for addressing selection bias 146
parameter estimation 346–351
parametric models
 CA methods and 165, 183
 LC analysis and 183–184
 SNMM and 237–238, 242–243
PARMS statement, NLP procedure 245
partitioned estimator 365
Patient Health Questionnaire (PHQ) 30–46
patient registry data analyses 156–163
PCOV option, NLP procedure 245
pharmacoeconomics
 CEA with censored data 363–382
 good research practices 288
 incremental net benefit 339–362
 propensity score bin bootstrapping 315–337
 reporting guidelines 14
pharmacoepidemiology
 good research practices 287
 reporting guidelines 14
PHQ (Patient Health Questionnaire) 30–46
PHREG procedure 304
PLOT procedure 320
Poisson regression model 100–101
POWER procedure 394, 407, 409
PREDICTED option, OUTPUT statement 144
probit regression
 defined 52
 estimating propensity scores 24
 IV method for addressing selection bias 144
propensity score (PS)
 See also longitudinal observational study
 See also RMLPS (Regression Models on Longitudinal Propensity Scores)
 advantages of 29–30
 bias reduction methods 105–106
 computing 32–33
 defined 12, 24, 52, 183
 doubly robust estimation and 84–104
 estimating 24–25, 52–53, 64–65
 evaluating 27–28

 evaluating balance produced 36–46
 fundamental theorem 185–186
 limitations of 29–30
 medication effectiveness study 30–46
 mortality rates analyses 158–161
 objective of 24, 27
 problems estimating 186–187
 quintile classification 197
 regression adjustment of 26–30, 60
 sensitivity analysis and 13
 stratification of 26–27, 62, 198
 with missing values 105–130
propensity score bin bootstrapping (PSBB)
 about 315–319
 schizophrenia study 320–334
propensity score matching
 assessing balance in baseline characteristics 55–58
 compared to other methods 62
 DES/BMS safety and efficacy study 62–80
 estimating propensity scores 52
 estimating treatment effect 58–61
 forming matched sets 53–55
 overview 52
 sensitivity analysis for 61–62

PS

See propensity score (PS)

PSBB (propensity score bin bootstrapping)
 about 315–319
 schizophrenia study 320–334

Q

QLIM procedure 143–146
quantile-quantile plots 56

R

randomization
 fundamental problems with 152–153
 key roles played 152
randomized controlled trials
 See RCTs (randomized controlled trials)
RANK procedure
 GROUPS= option 32
 propensity score example 32
 PSBB and 323
 schizophrenia study 267, 323
RCTs (randomized controlled trials)
 as gold standard 23
 association vs. causation 5–7
 bias in 4
 CD Trial 351–358

- RCTs (randomized controlled trials) (*continued*)
- CONSORT statement for 14
 - covariate adjustment 347, 356
 - defined 3
 - generalizability of 4
 - measurement bias 386
 - methodological considerations 10–13
 - observer bias 388
 - randomization in 58
 - replicability of 8
 - reporting guidelines 14
 - strengths of 3–4
 - study designs 9–10
- RDC (Research Diagnostic Criteria) 198
- Receiver Operating Characteristic
- See* ROC (Receiver Operating Characteristic)
- REG procedure 148
- regression adjustment method
- IV method comparison 146–147
 - of propensity scores 26–30, 60
- Regression Models on Longitudinal Propensity Scores
- See* RMLPS (Regression Models on Longitudinal Propensity Scores)
- replicability of observational studies 8–9
- reporting guidelines for observational studies 13–14
- Research Diagnostic Criteria (RDC) 198
- research practices
- addressing study bias 385–390
 - checklist for 289–294
 - retrospective databases and 287–288
- response analyses using large databases 295–311
- RMLPS (Regression Models on Longitudinal Propensity Scores)
- about 263–265
 - estimation 265
 - schizophrenia study 266–282
- Robins' G-computation formula 264
- ROC (Receiver Operating Characteristic)
- CA method 157
 - defined 58
 - IPW method 162
 - propensity score estimates 158
- S**
- salient treatment differences 153
- sample size
- bias and 9, 343
 - calculating in observational studies 391–425
 - checklist considerations 289, 291
 - confounding and 9
 - in retrospective database studies 293–294
 - nonparametric bootstrapping and 318
 - skewness and 317
- schizophrenia study
- MSM and 214–227
 - PSBB and 320–334
 - RMLPS and 266–282
- scleroderma 419–422
- SELECT option, MODEL statement 146
- selection bias
- defined 387
 - good research practices and 288
 - in observational studies 7–8, 391
 - in RCTs 4
 - IV method addressing 131–150
 - revealing in LC analysis 163–174
 - treatment 52
 - types of 7
- sensitivity analysis
- clustering and 178–179
 - different imputation strategies 126–127
 - for local control approach 165, 178–179
 - for propensity score matching 61–62, 75
 - handling extreme weights 127–128
 - importance of 13, 27
 - longitudinal observational study 224–227
 - schizophrenia study 280–282
 - varying analytic methods 124–125
- SHOWCURVES option, WTMODEL statement 89–90, 97, 100
- side-by-side box plots 56
- skewness
- in cost data 347, 351, 363–364
 - sample size and 317
- smearing method 364
- SNM (structured nested model) 12
- SNMM (Structural Nested Mean Model)
- about 231–234
 - estimation 236–239
 - getting starting values 254–256
 - LCLF Study 239–252
 - maximum likelihood data analysis 238–252
 - parametric models and 237–238
 - time-varying moderation and 235–236
- sponsor bias 385–390
- SR (stratified regression) approach 124–125

standard error 92–93
 standardized differences method 56–57, 68–74
 statistical methods
 commonly used tools 12
 for assessing balance in matched samples 56
 for CEA with censored data 364–368
 for local control analysis 154–155
 issues and considerations 10–11
 research checklist 289, 293
 rule suggestions 11
 stratification of propensity scores 26–27, 62, 198
 stratified regression (SR) approach 124–125
 STROBE (STrengthening the Reporting of Observational studies in Epidemiology) 14
 Structural Nested Mean Model
 See SNMM (Structural Nested Mean Model)
 structured nested model (SNM) 12
 study designs
 general issues 386–387
 hierarchy of 10
 in observational studies 9–10
 survival data study 405–409
 systemic sclerosis 419–422

T

t-test
 paired 59, 146
 PSBB and 316–317
 two-sample 393–395
 TABULATE procedure 134–135, 139
 temporality 7
 time-to-event outcomes 60–61
 time-varying moderation
 about 232–233
 in LCLF study 241–242
 notation for 234–235
 SNMM and 235–236
 Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) 14
 treatment
 See also independent variables
 estimating differences 107
 evaluating balance across groups 205–207
 exposure and 3*n*
 longitudinal model of propensity for 195–209
 salient differences in 153

 within clusters 155
 treatment effects
 adjusted for selection bias 133
 bias reduction in 105
 clustering and 155
 computing 33–36
 doubly robust estimation of 84–104
 estimating 26–27, 58–61, 124–128
 evaluation challenges 195
 longitudinal observational study 197–198, 203–205, 211
 MSM methodology 221–223
 sensitivity analysis 124–128
 TREND (Transparent Reporting of Evaluations with Nonrandomized Designs) 14
 TTEST procedure
 IV method for addressing selection bias 134, 136–137, 139, 146
 PSBB and 333
 Tukey-Kramer adjustment for multiple comparisons 138
 two-sample log-rank test 405–409
 two-sample *t*-test 393–395
 two-sample test on binary outcome 399–400
 two-stage regression 244, 253–256
 type I error
 in observational studies 8–9
 random effects and 348

U

UNIVARIATE procedure
 CLASS statement 136
 IV method for addressing selection bias 136
 schizophrenia study 279
 unmeasured confounding 29, 214
 USRDS (United States Renal Data System) 298

V

VARDEF= option, NLP procedure 245
 variables
 binary 398–400, 411–414, 419–422
 checklist for definitions 289, 291
 continuous 392–398, 413, 416–419
 dependent 7, 25, 414–422
 independent 7, 25, 414–422
 %VMATCH macro 65

W

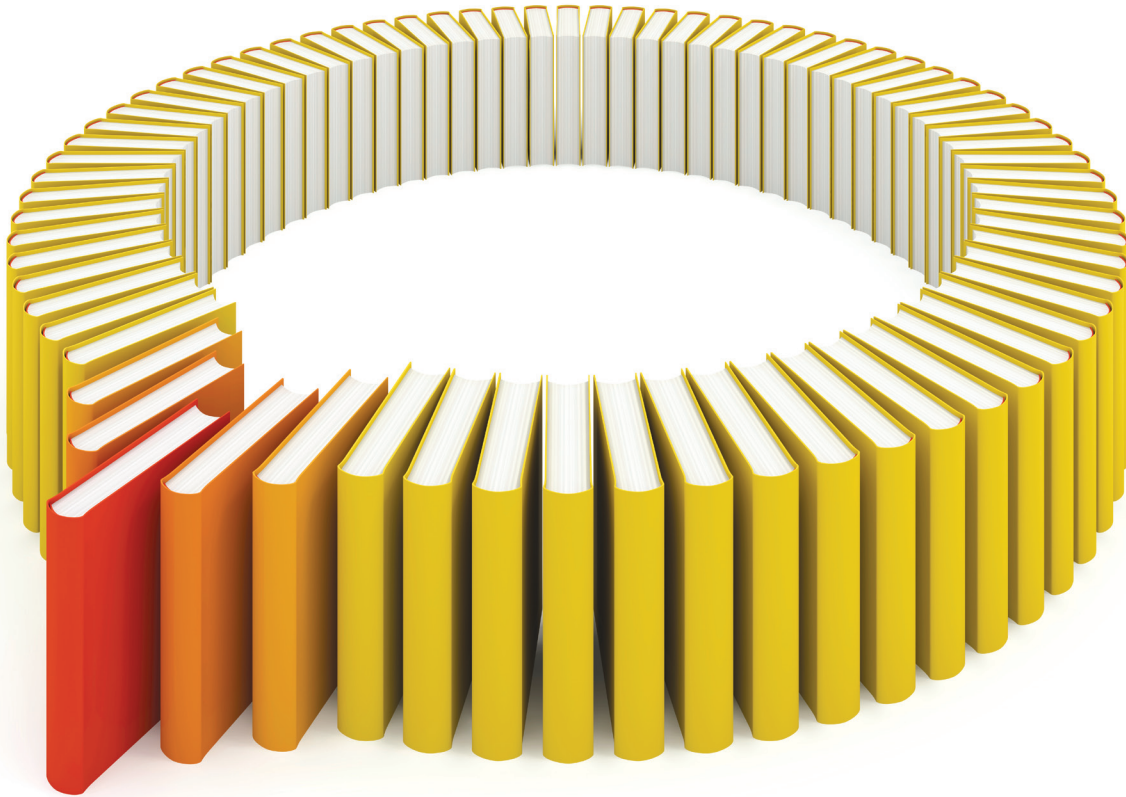
WEIGHT statement, GENMOD procedure 221
 weighted Mantel-Haenszel test 398, 400–405

WHERE statement 102
WHI (Women's Health Initiative) 5
Wilcoxon rank sum test
 about 395–398
 asymptotic distribution of 423–424
 LC approach and 176
 PBSS and 317, 333
willingness-to-pay (WTP) 341, 365
Women's Health Initiative (WHI) 5
WTMODEL statement
 COMMON_SUPPORT option 90, 92,
 97–98
 SHOWCURVES option 88–89, 97, 100
WTP (willingness-to-pay) 341, 365

Z

Z score 93

From *Analysis of Observational Health Care Data Using SAS®* by Douglas Faries, Andrew Leon, Josep Haro, and Robert Obenchain. Copyright © 2010, SAS Institute Inc., Cary, North Carolina, USA. ALL RIGHTS RESERVED.



Gain Greater Insight into Your SAS[®] Software with SAS Books.

Discover all that you need on your journey to knowledge and empowerment.

 support.sas.com/bookstore
for additional books and resources.


THE POWER TO KNOW[®]