Stockholms universitet

Finding therapy-response markers of lithium treatment of bipolar disease using causal interaction analysis

Carolina Fransson

Kandidatuppsats i matematisk statistik
Bachelor Thesis in Mathematical Statistics

Kandidatuppsats 2015:20
Matematisk statistik
Juni 2015

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Bachelor Thesis **2015:20**
http://www.math.su.se

Finding therapy-response markers of lithium treatment of bipolar disease using causal interaction analysis

Carolina Fransson*

June 2015

Abstract

In this thesis, we aim to identify genetic markers associated with lithium treatment response of patients with bipolar disorder. Identifying such genetic markers is a step towards personalized genotype-based treatment of bipolar disorder. By using causal inference we estimate bounds on the causal interaction between lithium treatment and genetic markers on treatment response. In particular, we aim to identify genetic markers that might either block the effect of lithium treatment or be a prerequisite for treatment response. For this reason we also estimate bounds on the causal interaction under the assumption of monotonic effects of lithium treatment and genetic markers on treatment response. We use a weighed logistic regression model to estimate the bounds, with Inverse Probability of Treatment Weights to control for confounding. We find that, to some extent, the genetic markers interact with lithium treatment. A small number of genetic markers potentially affect lithium treatment in only one direction.

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: fransson.carolina@gmail.com. Supervisor: Tom Britton.

Sammanfattning

Kunskap om hur genetiska faktorer påverkar resultatet av medicinering kan i förlängningen leda till skräddarsydda behandlingar baserade på patientens genetiska profil. I den här uppsatsen undersöker vi genetiska markörers inverkan på effekterna av litiumbehandling av bipolär sjukdom. Bipolär sjukdom är en psykisk sjukdom som kännetecknas av omväxlande perioder av mani och depression och en vanlig behandling av bipolär sjukdom är medicinering med litium. Vi använder kausal inferens för att skatta övre och undre gränser för hur vanligt förekommande det är att genetiska markörer interagerar kausalt med litiumbehandling. Vi genomför även analysen under antagandet att de genetiska markörerna inverkar monotont på resultatet av litiumbehandlingen. Detta eftersom genetiska markörer som, i den mån de påverkar resultatet av litiumbehandlingen, antingen blockerar effekten av behandlingen eller är en förutsättning för att behandlingen ska vara framgångsrik är av särskilt intresse. Vi skattar gränserna med en viktad logistisk regressionsmodell. Genom att använda "Inverse Probability of Treatment Weights" som vikter justerar vi för störningsfaktorer som varierar med tiden. Resultaten av analysen tyder på att de genetiska markörerna i viss mån samverkar med litiumbehandling. För ett mindre antal genetiska markörer tyder resultaten på att effekterna kan vara monotona.

Acknowledgments

I would like to express my gratitude to my supervisors Yudi Pawitan and Xia Shen for your guidance, time and patience and for giving me the opportunity to work on this interesting topic. I am also thankful to my supervisor Tom Britton for your advice on how to structure this report. I thank the Department of Medical Epidemiology and Biostatistics at the Karolinska Institutet for being so welcoming. Finally, I would like to thank Arvid Sjölander for sharing your programming code, parts of it were quite useful.

Contents

Abbreviations

SNP	Single Nucleotide Polymorphism
MSM	Marginal Structural Model
IPTW	Inverse Probability of Treatment Weight
CI	Confidence Interval
AIC	Akaike's information criterion

Chapter 1

Introduction

Bipolar disorder, or manic depression, is a mental disorder characterized by recurring episodes of mania, depression and in some cases "mixed episodes" (Mondimore, 2014). It is estimated that around 2 percent of the population suffers from some form of bipolar disorder. In the manic phase of a bipolar episode, individuals often experience inappropriately good or unusually irritable moods. Symptoms of the depressive phase of bipolar disorder may be feelings of hopelessness, loss of interest in activities once enjoyed etc. An individual with mixed episodes have symptoms characteristic of both mania and depression simultaneously, or in rapid sequence. An episode may last for a varying period of time, from a few hours to years. Many are free of symptoms between episodes. The cause is not clearly understood, but both genetic and environmental factors play a role. There is currently no cure for bipolar disorder. There are a range of medications developed to treat bipolar disease, including the mood stabilizer lithium.

In this thesis, we aim to identify genetic markers associated with lithium treatment response of patients with bipolar disorder. As genetic markers, we used Single Nucleotide Polymorphisms, SNP's (Xu, 2014). A SNP is a type of DNA sequence variation in the form of different sequence alternatives (alleles) at a single base pair position, commonly occurring within a population.

The structure of this report is as follows. In Chapter 2, we give a description of the data material, mainly focusing on the potential presence of time-dependent confounding, since this is a major problem if not controlled for properly. In the third chapter, we present the tools we will use to assess impact of genetic markers on lithium treatment response and to control for time-dependent confounding. We also describe important components of the process of analyzing the data. In the next chapter, Chapter 4, we present the results and conclusions and in Chapter 5 we give a brief discussion of the results.

Chapter 2

Material

The data analyzed in this project originates from a cohort study using clinical data from the Swedish National Quality Register for Bipolar Disorder (BipoläR). BipoläR was launched in 2004 and contains individualized data concerning basic demographic variables, presence of affective illness in the family, treatment with antidepressants and/or mood stabilizers and treatment outcome (Registercentrum Västra Götaland, 2015).

The study (Karantia et al., 2015) consists of 7354 bipolar patients registered in BipoläR between 2004 and 2011. The study subjects were followed between 2004 and 2011 with annual updates. Data were collected by treating physicians and participation was voluntary, both for bipolar patients and clinicians. Of the 7354 bipolar patients genotype data are available for 1318 bipolar patients which constitute the data material analyzed in this project. For these 1318 patients 4587 follow-up observations were made. The patients left blood samples from which DNA was extracted at the Karolinska Institutet Biobank. Genotyping were thereafter conducted at the Broad Institute of MIT and Harvard. In this project, 12168 SNP's are included. Each SNP is coded as 0,1 or 2 depending on whether the allele occur for 0,1 or 2 chromosomes. The data material has also been used by Tidemalm et al. (2014) in a cohort study aiming to investigate risk factors for attempted suicide in bipolar patients.

2.1 Description of data

We include variables measuring current and previous treatment of the disorder, occurrence of episodes, occurrence of episodes last year, age, type of bipolar disorder, date of follow up and gender. For a complete list of the variables, see Appendix A. Of the bipolar patients, 62% are women. Lithium treatment assignment is common among the patients, for 77% of the observations the patient is assigned to lithium treatment. The average age at registration is 50.3 years with standard deviation 14.2. (Figure 2.1) The prevalence of bipolar disorder Type I is 47% among the patients, the prevalence of bipolar disorder

Type II is 41%, the prevalence of bipolar type schizoaffective disorder is 1.6% and for 10% of the patients the type of bipolar disorder is not specified.

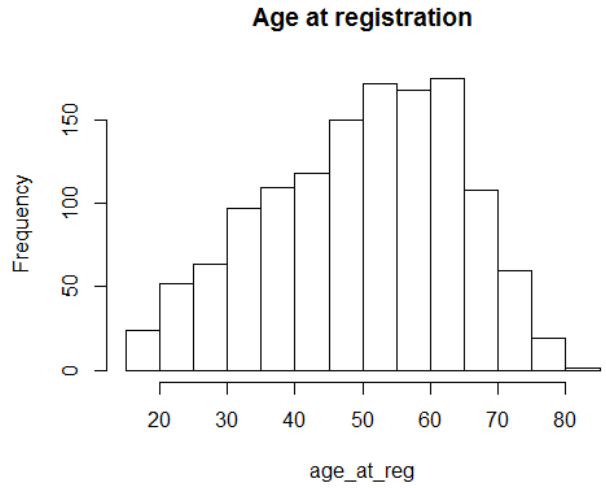

FIGURE 2.1: Age at registration in BipoläR

In our analysis, the impact of genetic markers and lithium treatment on the occurrence of bipolar episodes is central. Estimating treatment effects requires appropriate control for confounding. In observational studies in which the study subjects are followed over time, time-dependent confounding of the association between the outcome and treatments of interest is often present. One potential time-dependent confounder is previous bipolar episodes since treatment is likely to be affected by a patient's history of bipolar episodes and it is also possible that bipolar episodes tend to reoccur. Table 2.1 classifies data by current lithium treatment, by occurrence of bipolar episodes in the current year and by occurrence of episodes during the preceding year.

TABLE 2.1: Current lithium treatment, bipolar episodes in the current year and episodes during the preceding year

No current lithium treatment			Current lithium treatment		
Episodes current year	Episodes last year		Episodes current year	Episodes last year	
	No	Yes		No	Yes
No	200	202	No	1180	795
Yes	96	539	Yes	353	1181

From Table 2.1 we construct Tables 2.2-2.4, marginalizing over last year's episodes, current year's episodes and current lithium treatment respectively.

TABLE 2.2: Proportion of the participants experiencing bipolar episodes among those assigned to lithium treatment and among those not assigned to lithium treatment

	Episodes	
Lithium	No	Yes
No	0.39	0.61
Yes	0.56	0.44

TABLE 2.3: Proportion of the participants assigned to lithium among those for which bipolar episodes occurred during the preceding year and among those for which bipolar episodes did not occur

	Lithium	
Epi. last	No	Yes
No	0.16	0.84
Yes	0.27	0.73

TABLE 2.4: Proportion of the participants experiencing bipolar episodes among those for which episodes occurred the preceding year and among those for which episodes did not occur

	Episodes	
Epi. last	No	Yes
No	0.75	0.25
Yes	0.37	0.63

Table 2.2 suggests that the occurrence of episodes is more frequent among bipolar patients not assigned to lithium treatment than among those assigned to lithium treatment. Table 2.3 suggests that lithium treatment is more common among those who did not experience episodes during the preceding year. Table 2.4 suggests that bipolar episodes tend to reoccur. Thus, we have reasons to suspect that occurrence of episodes is a confounder of the association between future treatment and future bipolar episodes. The causal graph (see Chapter 3.1.6 for further information) shown in Figure 2.2 summarizes the suspected causal relations between lithium treatment and the occurrence of bipolar episodes.

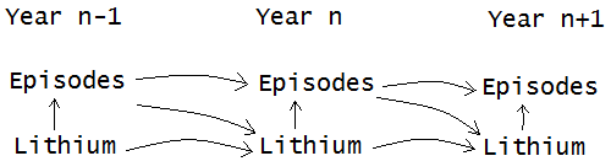

FIGURE 2.2: Causal graph summarizing the (suspected) causal relations between lithium treatment and the occurrence of bipolar episodes. Each arrow represents the presence of causal effects.

Chapter 3

Methods

3.1 Tools for causal analysis

The information in this chapter is based on Sjölander et al. (2014a), unless stated otherwise.

3.1.1 Notation

Using the same notation as Sjölander et al. (2014a) we let Y denote a response variable with binary outcome (0 or 1) for some individual in a population under study. Furthermore, we let X and Z denote two different categorical explanatory variables with levels $0, 1, 2 \ldots, K_X$ and $0, 1, 2 \ldots, K_Z$ respectively, where $K_X > 0$ and $K_Z > 0$.

3.1.2 Potential outcomes

If, for a certain individual $X = x$ and $Z = z$ then it is clearly impossible to observe the value that Y would have attained, had $X \neq x$ and/or $Z \neq z$. The assumptions (Greenland and Brumback, 2002) of the potential-outcome model are however:

1. Each individual could have been assigned to any one of the treatment levels $0, ..., K_X$ and $0, ..., K_Z$.

2. The value of Y that we would have observed at a specific time point, had the explanatory variables been intervened to take the values x and z respectively, possibly contrary to fact, exists for each individual in the population for all $x \in \{0, \ldots, K_X\}$ and $z \in \{0, \ldots, K_Z\}$. We denote this value, called the potential or counterfactual outcome, by Y_{xz}.

3.1.3 Causal interaction

To assess the impact of genetic markers on lithium treatment response we introduce the concept of causal interaction, not to be confused with statistical interaction. Absence of the latter often refers to the presence of additive effects on the outcome of interest on some scale.

If, for an individual, $Y_{xz} \neq Y_{\tilde{x}z}$ for some levels x and $\tilde{x}$ of X and z of Z then X is said to have a causal effect on Y for that individual. Similarly, if $Y_{xz} \neq Y_{x\tilde{z}}$ for some levels z and $\tilde{z}$ of Z and x of X then Z is said to have a causal effect on Y for that individual. Thus, if for a given individual, we could have prevented or caused the outcome by intervention in the level of an explanatory variable, that explanatory variable is said to have a causal effect on the outcome for that individual.

If X and Z both have causal effect on Y for a given individual then we say that there is causal interaction between X and Z for that individual. Equivalently, there is causal interaction between X and Z for a certain individual if the causal effect of both X and Z depends on the level of the other.

3.1.4 Monotonic effects

If, for all individuals, an explanatory variable always affect the outcome in a certain direction, that variable is said to have monotonic effects. We introduce the concept of monotonic effects.

We say that X has a positive monotonic effect on Y if for all individuals

$$Y_{xz} \geq Y_{\hat{x}z} \text{ if } x \geq \hat{x} \text{ for all } z \in \{0, \ldots, K_Z\}. \tag{3.1}$$

That is, under the assumption that X has positive monotonic effects, interventions to set X to a higher level can never prevent the outcome.

We say that X has a negative monotonic effect on Y if

$$Y_{xz} \leq Y_{\hat{x}z} \text{ if } x \geq \hat{x} \text{ for all } z \in \{0, \ldots, K_Z\} \tag{3.2}$$

for all individuals. Similarly, we say that Z has a positive monotonic effect on Y if for all individuals

$$Y_{xz} \geq Y_{x\hat{z}} \text{ if } z \geq \hat{z} \text{ for all } x \in \{0, \ldots, K_X\} \tag{3.3}$$

and we say that Z has a negative monotonic effect on Y if for all individuals

$$Y_{xz} \leq Y_{x\hat{z}} \text{ if } z \geq \hat{z} \text{ for all } x \in \{0, \ldots, K_X\}. \tag{3.4}$$

3.1.5 Bounds on Causal Interaction

Sjölander et al. (2014a) have derived bounds for the proportion of individuals in a population for which there is causal interaction.

We denote the proportion of individuals for which there is causal interaction between X and Z on Y by θ and we let p_{xz} denote the proportion of individuals for which $Y_{xz} = 1$. That is to say, p_{xz} is the proportion of individuals for which the outcome would occur if the levels of X and Z had been x and z respectively for all individuals, possibly contrary to fact.

The lower bound on θ is given by

$$\begin{aligned} \max_{x>x' \geq 0,\ z>z' \geq 0} (&\pm 0.5\psi_{xx'zz'}, \psi_{xx'zz'} - p_{x'z'}, -\psi_{xx'zz'} - p_{x'z}, \\ &- \psi_{xx'zz'} - p_{xz'}, \psi_{xx'zz'} - p_{xz}, -\psi_{xx'zz'} + p_{x'z'} - 1, \\ &\psi_{xx'zz'} + p_{xz'} - 1, \psi_{xx'zz'} + p_{x'z} - 1, -\psi_{xx'zz'} + p_{xz} - 1) \end{aligned} \tag{3.5}$$

where $\psi_{xx'zz'} = p_{xz} + p_{x'z'} - p_{xz'} - p_{x'z}$. The upper bound on θ is given by

$$\begin{aligned} \min_{\omega_X,\ \omega_Z} (1, &\sum_{x,z} I(x \in \omega_X)(1 - p_{xz}) + I(x \in \omega_X^c)p_{xz}, \\ &\sum_{x,z} I(z \in \omega_Z)(1 - p_{xz}) + I(z \in \omega_Z^c)p_{xz}) \end{aligned} \tag{3.6}$$

where $\omega_X = \{2^0, 2^1, \ldots, 2^{K_X}\}$, $\omega_Z = \{2^0, 2^1, \ldots, 2^{K_Z}\}$ and the indicator $I(a \in A) = 1$ if $a \in A$ and 0 otherwise. If the p_{xz}'s are known, these bounds are guaranteed to contain θ.

Sjölander et al. (2014a) also derived bounds valid under the assumption of positive monotonic effects of X and Z on Y. These bounds are at least as narrow as the bounds under no assumptions of monotonic effects. The lower bound on θ under the monotonicity assumption is given by

$$\max_{x>x'>x''\geq 0,\ z>z'>z''\geq 0} (\pm\psi_{xx'zz'}, \pm(\psi_{xx'zz'}+\psi_{xx'zz''}), \pm(\psi_{xx'zz'}+\psi_{x'x''zz'}), \tag{3.7}$$
$$\psi_{xx'zz'}-\psi_{x'x''zz'}, \psi_{xx'zz'}-\psi_{xx'z'z''}, \pm(\psi_{xx'zz'}+\psi_{x'x''z'z''}),$$
$$\pm(\psi_{xx'zz'}-\psi_{x'x''z'z''}), \psi_{x'x''zz'}-\psi_{xx'z'z''}-\psi_{x'x''z'z''},$$
$$\psi_{xx'z'z''}-\psi_{x'x''zz'}-\psi_{x'x''z'z''}, \pm(\psi_{xx'zz'}+\psi_{x'x''zz'}+\psi_{x'x''z'z''}),$$
$$\pm(\psi_{xx'zz'}+\psi_{xx'z'z''}+\psi_{x'x''z'z''}), \psi_{xx'zz'}\pm(\psi_{x'x''zz'}-\psi_{xx'z'z''}),$$
$$\psi_{xx'zz'}\pm(\psi_{x'x''zz'}\pm\psi_{xx'z'z''})-\psi_{x'x''z'z''})$$

and the upper bound is given by

$$\min_{\omega_X,\omega_Z} (\delta_{K_XK_Z00}\pm\delta_{1001}, \delta_{K_XK_Z00}+\sum_{x\in\omega_X\setminus\{2^{(K_X)}\}}\delta_{xK_Z(x+1)0}, \delta_{K_XK_Z00}+\sum_{z\in\omega_Z\setminus\{2^{(K_Z)}\}}\delta_{K_Xz0(z+1)}) \tag{3.8}$$

where $\delta_{xx'zz'}=p_{xz}-p_{x'z'}$. Let $\Delta=p_{K_XK_Z}-p_{00}$ denote the proportion of individuals for which $Y_{K_XK_Z}=1$ but $Y_{00}=0$. Under assumptions (3.1) and (3.3) of positive monotonicity, Δ is the proportion of individuals for which there exists x, $\hat{x}$,z and $\hat{z}$ such that $Y_{xz}=1$ but $Y_{\hat{x}\hat{z}}=0$, $0\leq\hat{x}\leq x\leq K_X$ and $0\leq\hat{z}\leq z\leq K_Z$. That is to say, since (3.1) and (3.3) is assumed, Δ is the proportion of individuals for which at least one of X and Z have an causal effect, so that under assumptions (3.1) and (3.3)

$$\frac{\theta}{\Delta} \tag{3.9}$$

is the proportion of individuals for which there is causal interaction among individuals for which there are causal effects.

3.1.6 Confounding

A confounder is a variable associated both with the occurrence of the outcome and the treatment of interest. We say that the covariate C is a confounder of the association between the treatment Z and the outcome Y if C is a cause of both Z and Y.

Figure 3.1 shows a causal graph (Greenland and Brumback, 2002) summarizing the fictional causal relations between the variables X, Y, C and Z in a population. An arrow from a variable X (the causal variable) to another variable Y (the affected variable) represents the presence of individuals for which X has a causal effect on Y not mediated through any of the other variables in the graph. If there is no sequence of directed arrows from a variable Y to a variable X the variable Y is assumed to have no causal effects on X. That is, no alternation of the distribution of Y would have an impact on the

distribution of X (Greenland and Brumback, 2002). If, as in Figure 3.1, there exists a sequence of directed arrows from a variable C to a treatment Z (not mediated through Y) and a sequence of directed arrows from C to the outcome Y (not mediated through Z) then C is a confounder of the association between Y and Z.

Due to confounding, estimation of the p_{xz}'s will be biased, if C is not controlled for.

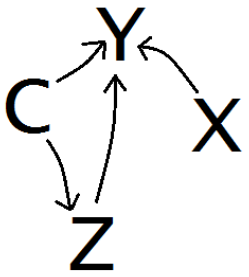

FIGURE 3.1: Example of confounding of the association between Z and Y by C

As noted, potential outcomes are unknown and impossible to observe except for the actual setting of the explanatory variables. Therefore estimation of p_{xz} is required in order to assess bounds on θ. Due to confounding, in general $P(Y = 1|X = x, Z = z) \neq P(Y_{xz} = 1) = p_{xz}$ for observational studies. That is, in the absence of appropriate control for confounding, estimation of the p_{xz} will be biased. If

$$P(Y_{xz} = 1|C) = P(Y = 1|X = x, Z = z, C) \tag{3.10}$$

for a set of measured covariates C, we say that C is sufficient for confounding control. If (3.10) hold, then

$$E(P(Y = 1|X = x, Z = z, C)) = E(P(Y_{xz} = 1|C)) =$$

$$= \sum_c P(Y_{xz} = 1|C = c)P(C = c) = p_{xz}.$$

3.2 Logistic regression model

A logistic regression model, also called a *logit* model, assume a binomial distribution of a binary response variable (Agresti, 2002).

The logistic regression model for $P(Y = 1|C = c) = E(Y|C = c)$ is

$$logit(P(Y = 1|C = c)) = \beta_0 + \sum_{i=1}^{p} \beta_i c_i \tag{3.11}$$

where $c = (c_1, \ldots, c_p)$ is the observed p-dimensional predictor variable C.

If $C = (C_1, \ldots, C_p)$ consists of factor variables the logistic regression model for $P(Y = 1|C = c) = E(Y|C = c)$ is

$$logit(P(Y = 1|C = c)) = \beta_0 + \sum_{i=1}^{p} \beta_i(c_i). \tag{3.12}$$

For the functions β_i to be identifiable, we impose constraints on β_i, $i = 1, \ldots, p$, say $\beta_i(c_i^*) = 0$ for some level c_i^* of C_i, $i = 1, \ldots, p$.

3.3 Marginal Structural Models

In observational studies where subjects typically are followed during a relatively long period of time, time-dependent confounders are often present. A covariate that satisfies the following two criteria

1. is associated with future treatment
2. is associated with future outcome

is a potential time-dependent confounder for the effect of treatment on outcome. If time-dependent confounders are also affected by previous treatment, conventional statistical methods might lead to biased estimates of treatment effect (Robins et al., 2000). Since we suspect that the occurrence of bipolar episodes last year is a confounder of the association between lithium treatment and episodes during the current year, controlling for confounding requires a careful approach. By using Marginal structural models (MSM's) we adjust for time-dependent confounding so that parameters of the model are consistently estimated, under the assumption that there are no unmeasured confounders (Robins et al., 2000).

A marginal structural model is a model for $E(Y_{xz}) = p_{xz} = P(Y_{xz} = 1)$ and can, for instance, be parameterized by

$$logit(p_{xz}) = \gamma_0 + \gamma_1(x) + \gamma_2(z) + \gamma_3(x, z) \tag{3.13}$$

where p_{xz} is uniquely determined by γ_i , $i = 0, 1, 2, 3$, if we impose constraints similar to the constraints on (3.12).

Let us now consider a logit model for $E(Y|X = x, Z = z) = P(Y = 1|X = x, Z = z)$ corresponding to (3.13) parameterized by:

$$logit(E(Y|X = x, Z = z)) = \gamma_0' + \gamma_1'(x) + \gamma_2'(z) + \gamma_3'(x, z). \tag{3.14}$$

If X and Z are randomized or unconfounded, $\gamma_i' = \gamma_i$, $i = 0, 1, 2, 3$. Fitting model (3.14) to observed data using ML will then result in asymptotically unbiased estimations of

γ_i', $i = 0, 1, 2, 3$ since maximum likelihood estimators are asymptotically unbiased under the regularity conditions (Liero and Zwanzig, 2011). Hence, estimating the parameters in (3.14) will result in asymptotically unbiased parameters of the model (3.13) when confounding is not present.

If the (possibly multivariate) covariate C is a confounder of the association between, say, Z and Y, we can use the Inverse-Probability-of-Treatment-Weights (IPTW's) to adjust for the confounding of C. A marginal structural model will then be fitted according to the following two steps:

1. The probability for each participant to have his/her own treatment history conditioned on the covariate C is estimated. These estimated probabilities are then used to construct the IPTW's. The IPTW for an individual assigned to treatment $Z = z$ is
$$\frac{1}{P(Z = z|C)}. \tag{3.15}$$
2. A weighed model for the probability of the outcome to occur conditioned on X is fitted. The IPTW's are used as weights. Under the assumption of no unmeasured confounding the estimations of $E(Y_{xz})$ made in this step will be asymptotically unbiased (Robins et al., 2000).

We will use logistic regression for the estimations made in these two steps. Thus, the marginal structural model is (3.13) with the estimated IPTW's used as weights. The IPTW's are estimated using the model

$$logit(P(Z = 1|C = c)) = \alpha_0 + \sum_{i=1}^{p} \alpha_i(c_i) \tag{3.16}$$

where $c_i,\ i = 1, \ldots, p$ constitute c.

To see why this approach results in consistent estimators, assume that there is no unmeasured confounders of the association between outcome and treatment and that the causal graph 3.1 is accurate. If we were to fit the model (3.14) to data, the following function would be maximized in order to maximize the likelihood function:

$$\sum_{i=1}^{n} Y_i(log(\pi_i(x_i, z_i))) + (1 - Y_i)log(1 - \pi_i(x_i, z_i)) \tag{3.17}$$

where Y_i is outcome for participant i, x_i and z_i is the observed values of X and Z for participant i and $\pi_i(x_i, z_i) = E(Y_i|X = x_i, Z = z_i)$. When using the IPTW's as weights the following function is maximized:

$$\sum_{i=1}^{n} w_i Y_i(log(\pi_i(x_i, z_i))) + w_i(1 - Y_i)log(1 - \pi_i(x_i, z_i)) \tag{3.18}$$

where w_i is the IPTW for for participant i.

Using a similar reasoning as Robins et al. (2000), by using the IPTW's model (3.18) is fitted to a pseudo-population consisting of w_i copies of person number i, $i = 1, 2, \ldots, N$, where N is the number of individuals in the actual population. To see why this works, assume that within stratum $C = c$ consisting of $n = n_1 + n_2$ individuals n_1 individuals were assigned to the treatment $Z = z_1$ and n_2 individuals were assigned to the treatment $Z = z_2$. If we were to construct a pseudo-population by creating $\frac{1}{n_1/n}$ copies of each individual assigned to $Z = z_1$ and $\frac{1}{n_2/n}$ copies of each individual assigned to $Z = z_2$ then the pseudo-population would consist of $\frac{n_1}{n_1/n} + \frac{n_2}{n_2/n} = 2n$ individuals of which n were assigned to $Z = z_1$ and n to $Z = z_2$. This corresponds to using a model with all possible interaction terms included to estimate the IPTW's.

By using IPTW's, we create a fictional pseudo-population in which the treatment Z is unconfounded by the covariate C. Furthermore, the p_{xz}'s are not altered in the pseudo-population which we show in Appendix C. Hence, in the pseudo-population, the p_{xz}'s can be consistently estimated by fitting a model for $E(Y|X = x, Z = z)$.

3.4 Data analysis

3.4.1 Assumptions

A person's genotype is already settled before birth and it is impossible that the association between a certain SNP and the occurrence of bipolar episodes is confounded by some of the covariates (Appendix B) included in the data material. To the extent that there is a causal relation between an individual's genotype and the covariates, the covariates are a consequence of the genotype. Therefore, we do not include the SNP's when estimating the IPTW's.

If the covariates listed in Appendix A are not sufficient for confounding control the estimations of the p_{xz}'s will not be consistent. There is no way to decide if that is the case (Robins et al., 2000). Due to the complexity of the world, it is not reasonable to assume sufficiency for confounding control. Before prescribing lithium treatment, a psychiatrist will consider factors such as the patient's history of bipolar symptoms, (that is the severity of the illness) the bipolar type and the treatment history. These factors describe much of the mechanisms behind lithium prescription. This suggests that the untestable assumption that the bias inflicted by unmeasured confounders is small is plausible.

3.4.2 Model selection based on Akaike's information criterion

In section 3.3 we described how Marginal Structural Models can be used to control for time-dependent confounding. MSM's can also be used to control for time-invariant

confounding. There are a number of covariates (see Appendix A) that are potential confounders of the association between lithium assignment and the occurrence of bipolar episodes.

We use Akaike's information criterion, AIC, to compare models for the IPTW's. For a given model M and for a given sample, the definition of AIC is (Agresti, 2002)

$$AIC = -2(L_M - d) \tag{3.19}$$

where d is the number of parameters in M and L_M is the maximized log-likelihood for M. When using AIC to compare models, models with small AIC is preferred. AIC is the sum of two functions, $-2L_M$ which rewards goodness of fit and $2d$ which penalizes models with many parameters.

We use the built in R function step (R Core Team, 2015) to select a model for the probability to be assigned to lithium treatment. The function step performs backward elimination based on AIC.

3.4.3 Estimation of the bounds

To estimate the p_{xz}'s, we use the weighed logistic regression model

$$logit(P(Y = 1|X = x, Z = z)) = \gamma_0 + \gamma_1(x) + \gamma_2(z) + \gamma_3(x, z) \tag{3.20}$$

with the IPTW's as weights, where Z denotes current lithium treatment and X denotes the SNP under consideration. The estimated bounds on θ are then calculated according to (3.5) and (3.6) for no monotonicity assumptions and according to (3.7) and (3.8) for monotonicity assumptions. Note that (3.7) and (3.8) can be used to calculate bounds on θ under assumptions of both negative and positive monotonic effects since the direction of a monotonic effect only depends on the coding. Bounds were not estimated under the assumption of lithium treatment having positive monotonic effects, since that would mean that lithium treatment could never prevent bipolar episodes.

3.4.4 Confidence intervals by Bootstrapping

We calculate bootstrapped (Alm and Britton, 2008) 95% percentile confidence intervals for each estimated bound based on 500 bootstrap replicates. Bootstrapping is made from individuals, not observations, and thus the number of observation varies between bootstrap replicates due to drop outs. For each resampled set of participants, the IPTW's are recalculated and used as weights when estimating the p_{xz}'s. When monotonicity is assumed, the number of replicates for which the monotonicity assumptions are violated is recorded.

3.4.5 Multiple testing

In this project, confidence intervals are calculated for more than 12.000 SNP's under various assumptions. If the individual confidence level is $1-\alpha$ then the overall confidence level will, if left unadjusted, be considerably smaller than $1-\alpha$. That is, the risk of false discoveries will be larger than α. Adjusting CI's for multiple testing will typically result in wider intervals. The number of CI's is quite large, and adjustments will probably lead to a situation where no conclusions can be drawn. During this discovery stage we want to maintain high power, even at the expense of potentially more false positives. For this reason, we will calculate confidence intervals with individual confidence levels 95%. It is important to bear in mind that some of these confidence intervals will probably not attain zero by pure chance.

3.4.6 Coding of genetic markers

It may be of interest to investigate the effect of having none (SNP level 0) or any (SNP level 1 or 2) SNP, in particular under monotonicity assumptions. For this reason we group level 1 and 2 of SNP's so that SNP's are coded as binary (none/any) in addition to the ternary coding of the SNP's. We perform a analysis with SNP's coded as binary analogous to the analysis with SNP's coded as ternary.

Chapter 4

Results

4.1 Model for the weights

As described is Chapter 3.4.2 we use backward elimination based on AIC in order to select a model for the probability to be assigned to lithium treatment. We start with an initial model with no interaction terms presented in Appendix B. The initial model has deviance 1142.8 on 3215 degrees of freedom. The deviance D can be used to test the null hypothesis that the initial model is accurate against the alternative of the saturated model, since D is (asymptotically) χ^2 distributed with 3215 degrees of freedom under the null. $P(\chi^2_{3215} > 1142.8) \approx 1$, indicating a good fit for the initial model. For this reason, we do not consider more complicated models. The selection process results in the following model for the probability to be assigned to lithium treatment:

$$logit(P(Z_t = 1|Z_{t-1} = z_{t-1}, Y_{t-1} = y_{t-1})) = \alpha_0 + \alpha_1 z_{t-1} + \alpha_2 y_{t-1} \tag{4.1}$$

where Z_t denotes current lithium treatment, Z_{t-1} denotes lithium treatment in the preceding year and Y_{t-1} denotes the occurrence of bipolar episodes in the preceding year. Model (4.1) is presented in Appendix B.

4.2 Bounds on causal interaction

We estimate bounds on θ, the proportion of individuals for which there are causal interaction between the genetic markers and lithium treatment. We also calculate bootstrap confidence intervals for the bounds on θ with (individual) confidence level 95%. This analysis is carried out with binary coding of the SNP's (none or any) and with ternary coding (none, single or double).

4.2.1 Bounds with no monotonicity assumption

4.2.1.1 Binary coding of genetic markers

With binary coding (Chapter 3.4.6) of the SNP's and under no assumption of monotonic effects, the estimated upper bound on θ attains one for all SNP's and the confidence bounds on the upper bound on θ collapse into one.

Figure 4.1 shows the estimated lower bound on θ, which we call θ_l, and the upper and lower confidence bound on θ_l sorted by the estimated value of θ_l.

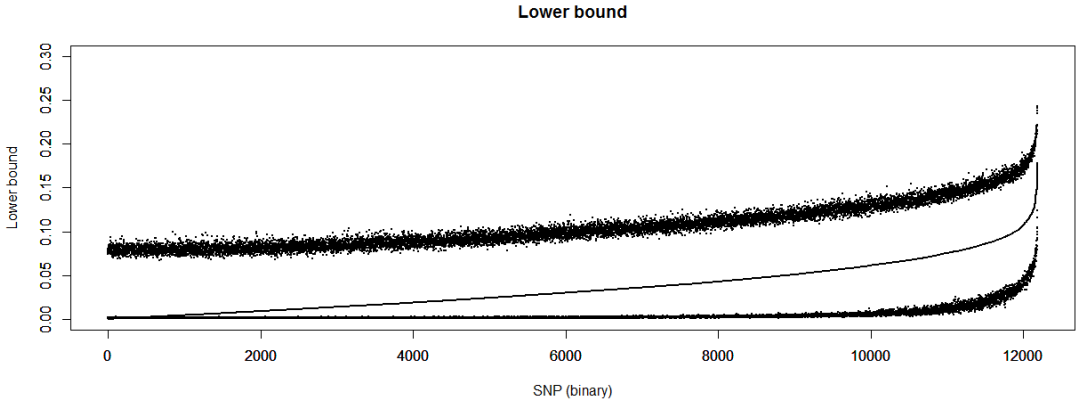

FIGURE 4.1: The estimated lower bound on θ, θ_l, and lower and upper confidence bounds on θ_l with binary coding of the SNP's

Table 4.1 contains the estimated θ_l and the confidence bounds on θ_l for the ten SNP's for which the lower confidence bound on θ_l is largest.

TABLE 4.1: Estimated θ_l and (individual) 95% confidence intervals on θ_l for the ten SNP's for which the lower confidence bound on θ_l is largest with binary coding of the genetic markers

SNP	Estimated θ_l	Confidence interval for θ_l
rs9974713	0.15	(0.090,0.20)
rs2285689	0.15	(0.090,0.21)
rs247908	0.16	(0.092,0.21)
rs11188500	0.15	(0.093,0.22)
rs3751212	0.16	(0.097,0.22)
rs1921372	0.17	(0.099,0.24)
rs10461813	0.16	(0.10,0.24)
rs12153044	0.16	(0.10,0.24)
rs4141835	0.18	(0.12,0.24)
rs6805636	0.18	(0.12,0.24)

4.2.1.2 Ternary coding of genetic markers

With ternary coding (Chapter 3.4.6) of the SNP's and under no assumption of monotonic effects, the estimated upper bound on θ attains one for all SNP's and the confidence bounds on the upper bound on θ collapse into one.

Figure 4.2 shows the estimated lower bound on θ, denoted by θ_l, and the upper and lower confidence bound on θ_l sorted by the estimated value of θ_l.

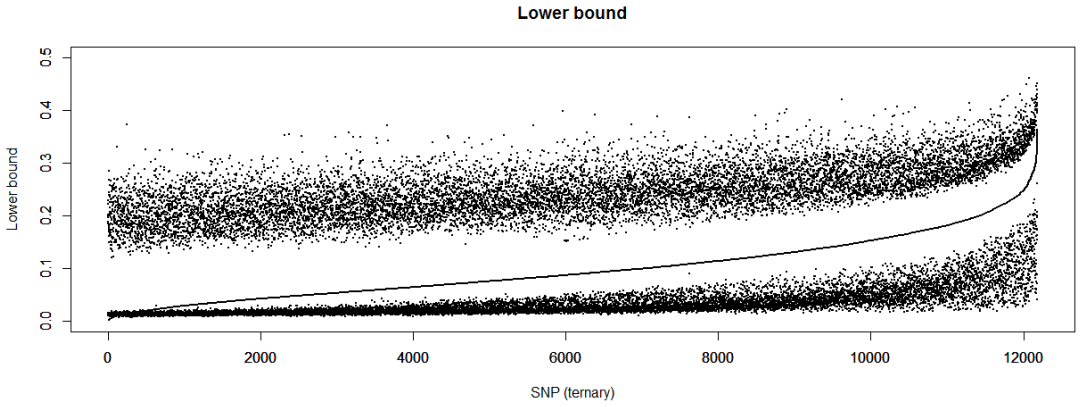

FIGURE 4.2: The estimated lower bound on θ, θ_l, and lower and upper confidence bounds on θ_l with ternary coding of the SNP's

Table 4.2 contains the estimated θ_l and the confidence bounds on θ_l for the ten SNP's for which the lower confidence bound on θ_l is largest.

TABLE 4.2: Estimated θ_l and (individual) 95% confidence intervals on θ_l for the ten SNP's for which the lower confidence bound on θ_l is largest with ternary coding of the genetic markers

SNP	Estimated θ_l	Confidence interval for θ_l
rs11762251	0.28	(0.20,0.36)
rs4817536	0.27	(0.20,0.33)
rs7495931	0.30	(0.20,0.38)
rs1457614	0.28	(0.20,0.35)
rs12466870	0.29	(0.20,0.39)
rs16842755	0.32	(0.21,0.41)
rs8000327	0.30	(0.21,0.41)
rs7713886	0.29	(0.22,0.35)
rs12972417	0.29	(0.23,0.36)
rs8139063	0.33	(0.26,0.39)

These results suggest that, on the whole, there are some causal interaction between the SNP's and lithium treatment on the occurrence of bipolar episodes. That is, there are SNP's that influence how some individuals respond to lithium treatment.

4.2.2 Bounds under assumption of positive monotonic effects

Negative monotonic effects of lithium treatment are assumed, both for the binary and for the ternary coding of the SNP's. That is, we assume that lithium treatment can not cause bipolar episodes. We make no assumptions of positive monotonic effect of lithium treatment since it is unreasonable to assume that lithium treatment, which is a well documented mood stabilizer, can never prevent bipolar episodes. Positive monotonic effects of the genetic markers are assumed. In other words we assume that the occurrence of the SNP in question can block the effect of lithium treatment on bipolar episodes but never prevent episodes.

4.2.2.1 Binary coding of genetic markers

With binary coding of the SNP's and under the assumption of positive monotonic effects of the SNP's, the lower confidence bound on the lower bound on θ attain zero for all SNP's except for four SNP's, which are tabulated in Table 4.3. We note that the confidence intervals for θ_l and θ_u overlap to a great extent for the four SNP's. The lower confidence bound on θ_l/Δ is close to zero and the upper confidence bound on θ_u/Δ is close to one for all SNP's. Even if we were to ignore that the confidence intervals are unadjusted, it would be impossible to assess the importance of the positive monotonic effects, if they exist.

For each SNP, we record the number of bootstrap replicates for which the monotonicity assumptions are violated. The monotonicity assumptions are violated if the lower bound exceeds the upper bound, that is if $\theta_l > \theta_u$. For 31 SNP's, the monotonicity assumptions are violated for less than 5% of the bootstrap replicates. This suggests that there are some SNP's that may reduce the chance for lithium treatment to succeed in preventing bipolar episodes.

Figure 4.3 shows the (sorted) proportion of bootstrap replicates for which the monotonicity assumption was violated for each SNP. The proportion of violations is quite large for most of the SNP's, suggesting that most SNP's do interact with lithium treatment to prevent bipolar episodes for some patients.

TABLE 4.3: The SNP's for which lower confidence bound on the lower bound on θ is larger than zero. By $\frac{\theta}{\Delta}$ we denote the proportion of the causal effect that are due to (causal) interaction between the SNP's and lithium treatment

SNP	rs1962292	rs17707219	rs485220	rs7149001
Prop. violated bootstrap repl.	0.024	0.024	0.022	0.022
Estimated lower bound on θ	0.056	0.049	0.077	0.065
Estimated upper bound on θ	0.25	0.24	0.25	0.26
CI lower bound on θ	(0.00092,0.19)	(0.00013,0.18)	(0.0015,0.20)	(0.00045,0.20)
CI upper bound on θ	(0.059,0.33)	(0.064,0.33)	(0.110,0.37)	(0.086,0.34)
Estimated lower bound on $\frac{\theta}{\Delta}$	0.21	0.20	0.24	0.25
Estimated upper bound on $\frac{\theta}{\Delta}$	0.93	0.99	0.78	0.99
CI lower bound on $\frac{\theta}{\Delta}$	(0.00320,0.62)	(0.00054,0.63)	(0.0047,0.52)	(0.0018,0.62)
CI upper bound on $\frac{\theta}{\Delta}$	(0.19,0.99)	(0.18,1.00)	(0.20,0.99)	(0.20,0.99)

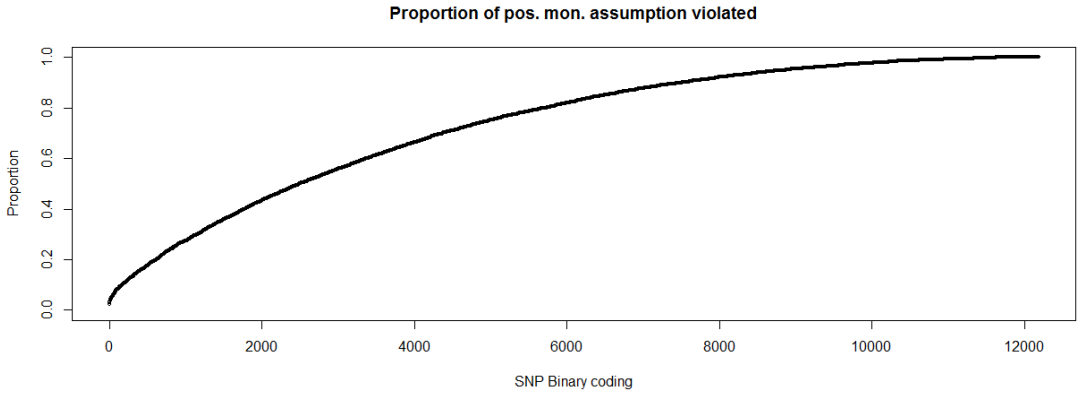

FIGURE 4.3: The proportion of bootstrap replicates for which the assumption of positive monotonic effects were violated for each SNP

4.2.2.2 Ternary coding of genetic markers

With ternary coding of the SNP's and under the assumption of positive monotonic effects of the SNP's, both the lower confidence bound for θ_l and the lower confidence bound for θ_u attains zero for all SNP's. The proportion of bootstrap replicates for which the monotonicity assumptions are violated (Figure 4.4) is at least 0.054 for all SNP's.

Figure 4.4 shows the proportion of bootstrap replicates for which the monotonicity assumptions are violated for each SNP.

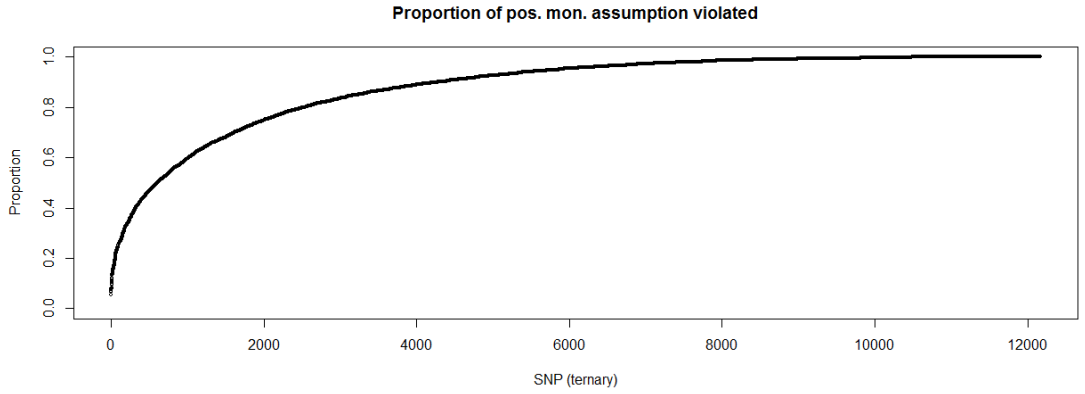

FIGURE 4.4: The sorted proportion of bootstrap replicates for which the assumption of positive monotonic effects are violated ($\theta_l > \theta_u$) for each SNP

4.2.3 Bounds under negative monotonicity assumption

Negative monotonic effects of lithium treatment are assumed as in the previous chapter. Negative monotonic effects of the genetic markers are assumed. In other words, we assume that the occurrence of the SNP in question can cause the patient to respond to lithium treatment and in doing so prevent bipolar episodes but never block the effect of lithium treatment.

4.2.3.1 Binary coding of genetic markers

With binary coding of the SNP's and under the assumption of negative monotonic effects of the SNP's, the lower confidence bound on the lower bound on θ and the lower confidence bound on the upper bound on θ attain zero for all SNP's. That is, even if we were to ignore that the confidence intervals are unadjusted, no conclusions on the presence of negative monotonic effects can be drawn. For each SNP, we record the number of bootstrap replicates for which the monotonicity assumptions are violated. The monotonicity assumptions are violated if the lower bound exceeds the upper bound, that is if $\theta_l > \theta_u$. For the four SNP's tabulated in Table 4.4, the monotonicity assumptions are violated for less than 5% of the bootstrap replicates.

TABLE 4.4: The SNP's for which the monotonicity assumptions are violated ($\theta_l > \theta_u$) for less than 5% of the bootstrap replicates. By $\frac{\theta}{\Delta}$ we denote the proportion of the causal effect that are due to (causal) interaction between the SNP's and lithium treatment

SNP	rs7781044	rs7150844	rs8090471	rs139662
Prop. viol	0.042	0.046	0.030	0.042
Estimated lower bound on θ	0.050	0.024	0.047	0.089
Estimated upper bound on θ	0.23	0.23	0.23	0.25
CI lower bound on θ	(0,0.18)	(0,0.14)	(0,0.18)	(0,0.20)
CI upper bound on θ	(0,0.30)	(0,0.30)	(0,0.32)	(0,0.32)
Estimated lower bound on $\frac{\theta}{\Delta}$	0.22	0.10	0.19	0.36
Estimated upper bound on $\frac{\theta}{\Delta}$	0.99	0.95	0.94	0.99
CI lower bound on $\frac{\theta}{\Delta}$	(0,0.67)	(0,0.54)	(0,0.59)	(0,0.72)
CI upper bound on $\frac{\theta}{\Delta}$	(0,0.99)	(0,0.99)	(0,0.99)	(0,1.00)

Figure 4.5 shows the (sorted) proportion of bootstrap replicates for which the monotonicity assumptions are violated for each SNP. As under the assumption of positive monotonic effects, the proportion of violations is high for most SNP's. This indicates that number of investigated SNP's that can not block the effect of lithium treatment is small.

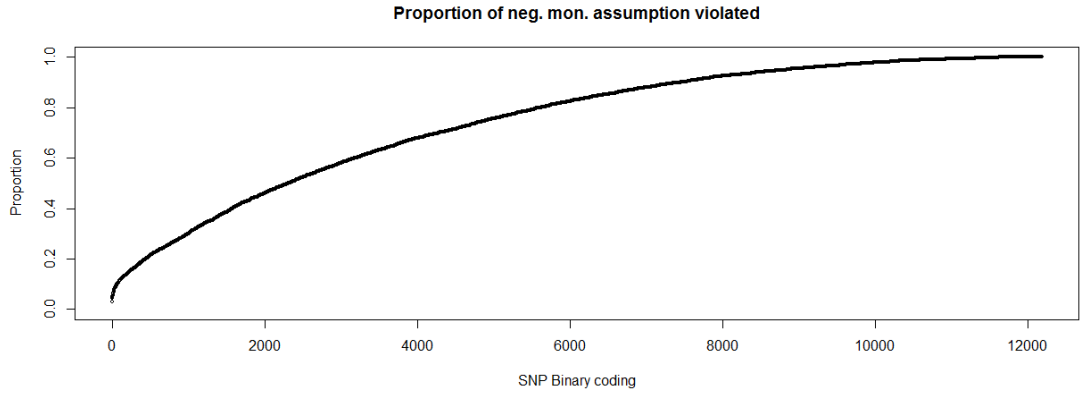

FIGURE 4.5: The proportion of bootstrap replicates for which the assumption of negative monotonic effects were violated for each SNP

4.2.3.2 Ternary coding of genetic markers

With ternary coding of the SNP's and under the assumption of negative monotonic effects of the SNP's, both the lower confidence bound on θ_l and the lower confidence bound on θ_u attains zero for all SNP's. The proportion of bootstrap replicates for which

the monotonicity assumptions are violated (Figure 4.6) is at least 0.27 for all SNP's and considerably larger for most SNP's.

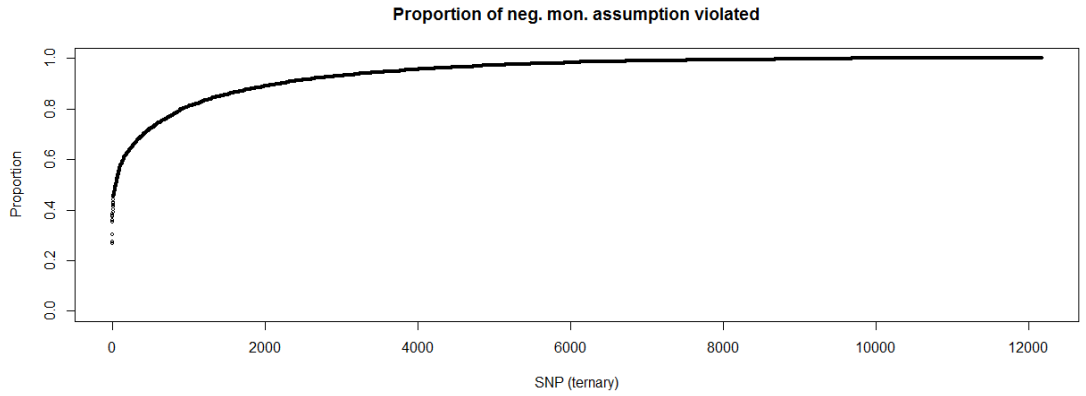

FIGURE 4.6: The sorted proportion of bootstrap replicates for which the assumption of negative monotonic effects were violated for each SNP

4.3 Conclusions

The results suggest that there are genetic markers that, to some extent, influence how some individuals respond to lithium treatment. For most of the genetic markers, the results do not indicate that they always affect lithium treatment response in the same direction. With binary coding (see Chapter 3.4.6) of the genetic markers we find 35 particularly interesting markers that potentially affect lithium treatment response in the same direction. Some of these genetic markers are tabulated in Table 4.3 and in Table 4.4. As shown in these tables, both the lower confidence bound on the lower bound on θ (the proportion of bipolar patients for which the genetic marker affects treatment response) and the lower confidence bound on the lower bound on θ/Δ (the proportion of causal effect that is due to causal interaction between lithium treatment and the genetic marker in question) are close to zero for all genetic markers. The upper confidence bound on the upper bound on θ and the upper confidence bound on the upper bound on θ/Δ are close to one for all genetic markers. That is, the evidence is not strong enough to assess to what extent the genetic markers influence lithium treatment response based on the results.

Chapter 5

Discussion

The aim of this project was to find genetic markers that affect how bipolar patients respond to lithium treatment. We were primarily interested in genetic markers that, to the extent they affect treatment response, always act in the same direction. Knowledge about such genetic markers allows for personalized treatment.

Out of 12186 genetic markers, we found around 35 interesting candidates. Under the assumption that the genetic marker under consideration always act in the same direction, the lower confidence bound on the lower bound on the proportion of individuals for which genetic markers affect lithium treatment response is very close to zero for all interesting candidates. The upper confidence bound on the upper bound is close to one for all interesting candidates. Even more important, the corresponding bounds on the proportion of the causal effects inflicted by lithium and/or the genetic marker under consideration that is due to interaction between the genetic marker and lithium treatment range from a value close to zero to a value close to one for all genetic markers. For this reason, we can not draw any conclusions about how important the role of these genetic markers is, if they do affect lithium treatment response. It is important to emphasize that the result can not be used to establish statistical significance, since we do not adjust for simultaneous comparisons.

Twelve thousand genetic markers is, from a genetic perspective, a very small number. A more extensive genotype data material is available, and this project could easily be extended to include a larger number of genetic markers. Another possible extension of this project is to estimate bounds on the sufficient-cause interaction between genetic markers and lithium treatment rather than the causal interaction, as described by Sjölander et al. (2014b). Sufficient-cause interaction is present if causal interaction is present and, in addition, the interacting exposures are necessary for the outcome to occur. The finding of genetic markers that participate in sufficient-cause interaction with lithium treatment would be even more interesting than the finding of genetic markers that participate in causal interaction.

Appendix A

List of variables

Number	Variable	Description
1	li	Whether participant currently assigned to lithium treatment
2	bipol_type	Type of bipolar disorder: Type I, Type II, bipolar type schizoaffective disorder or Not specified
3	epi_depr_bf_reg	Whether depressive episodes resulting in hospitalization had occurred before registration
4	epi_hypo_bf_reg	Whether hypomanic episodes resulting in hospitalization had occurred before registration
5	epi_manic_bf_reg	Whether manic episodes resulting in hospitalization had occurred before registration
6	epi_mixed_bf_reg	Whether mixed episodes resulting in hospitalization had occurred before registration
7	epi_total_bf_reg	Whether bipolar episodes resulting in hospitalization had occurred before registration
8	info_day	Date of follow up
9	age_at_reg	Participants age at registration
10	current_anti_depr	Whether participant currently assigned to antidepressive medicine
11	current_neuro	Whether the patient is currently assigned to antipsychotic medication
12	current_mood	Whether the patient is currently assigned to mood stabilizers
13	epi_yr	Whether patient experienced episodes resulting in hospitalization during current year
14	epi_last_yr	Whether participant was hospitilized due to episodes the previous year
15	li_last_yr	Whether participant was assigned to lithium the previous year
16	sex	Man/Woman

Appendix B

Models for the IPTW's

```
Call:
glm(formula = li ~ epi_manic_bf_reg + epi_total_bf_reg + bipol_type +
    epi_depr_bf_reg + epi_last_yr + li_last_yr + sex + current_neuro +
    epi_hypo_bf_reg + epi_mixed_bf_reg + age_at_reg + current_anti_depr,
    family = "binomial", data = dataMY)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.7858   0.2109   0.2392   0.2654   2.3394

Coefficients:
                                        Estimate Std. Error z value Pr(>|z|)
(Intercept)                           -2.938476   0.831598  -3.534  0.00041 ***
epi_manic_bf_reg                       0.208634   0.328079   0.636  0.52482
epi_total_bf_reg                      -0.008678   0.288479  -0.030  0.97600
bipol_typeTyp I                        0.625160   0.594229   1.052  0.29278
bipol_typeTyp II                       0.709135   0.651981   1.088  0.27674
bipol_typeUNS (utan narmare specifikation)  0.546787   0.676875   0.808  0.41920
epi_depr_bf_reg                        0.122693   0.413823   0.296  0.76686
epi_last_yr                           -0.260637   0.179871  -1.449  0.14733
li_last_yr                             5.603749   0.176511  31.747  < 2e-16 ***
sexMan                                -0.180257   0.176041  -1.024  0.30586
current_neuro                         -0.204158   0.197167  -1.035  0.30046
epi_hypo_bf_reg                       -0.025438   0.239751  -0.106  0.91550
epi_mixed_bf_reg                       0.058407   0.185663   0.315  0.75308
age_at_reg                             0.003144   0.006572   0.478  0.63232
current_anti_depr                     -0.036133   0.180100  -0.201  0.84099
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3343.1  on 3229  degrees of freedom
Residual deviance: 1142.8  on 3215  degrees of freedom
AIC: 1172.8

Number of Fisher Scoring iterations: 6
```

Listing B.1: Initial model for the IPTW's

```
Call:
glm(formula = li ~ epi_last_yr + li_last_yr, family = "binomial",
    data = dataMY)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.6774   0.2373   0.2373   0.2739   2.2077

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -2.0534     0.1583 -12.972   <2e-16 ***
epi_last_yr  -0.2921     0.1705  -1.713   0.0867 .
li_last_yr    5.6097     0.1721  32.593   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 3343.1  on 3229  degrees of freedom
Residual deviance: 1147.4  on 3227  degrees of freedom
AIC: 1153.4

Number of Fisher Scoring iterations: 6
```

Listing B.2: Choosen model for the IPTW's

Appendix C

Proof that using IPTW preserves the $\mathbf{p_{xz}}$'s

Let Y denote a binary outcome (0 or 1) and let X and Z denote two categorical explanatory variables with levels $0, \ldots, K_X$ and $0, \ldots, K_Z$ respectively. Let C denote a categorical multivariate covariate and assume that C is sufficient for confounding control, that is $P(Y_{xz} = 1|C) = P(Y = 1|(X, Z) = (x, z), C)$ and that (X, Z) has no causal effects on C. Now suppose that we create a pseudo-population (as described in Chapter 3.3) by using the IPTW's

$$\frac{1}{P(Z = z|C = c)} \cdot \frac{1}{P(X = x|Z = z, C = c)} = \frac{1}{P(X = x, Z = z|C = c)}$$

and let $\hat{Y}$ denote the outcome in the pseudo-population created by using the IPTW's. Similarly, let the $(\hat{X}, \hat{Z})$ denote the treatments in the pseudo-population and let $\hat{C}$ denote the covariate in the pseudo-population.

Then

$$P(\hat{Y} = 1|(\hat{X}, \hat{Z}) = (x, z)) =$$
$$= \sum_c P(\hat{Y} = 1|(\hat{X}, \hat{Z}) = (x, z), \hat{C} = c) P(\hat{C} = c|(\hat{X}, \hat{Z}) = (x, z)) =$$
$$= \sum_c P(Y = 1|(X, Z) = (x, z), C = c) P(\hat{C} = c|(\hat{X}, \hat{Z}) = (x, z))$$

where the last step follows from the fact that within strata $(X, Z) = (x, z)$ and $C = c$ the proportion of subjects for which $Y = 1$ is intact after using the IPTW's. Furthermore,

$$P(\hat{C} = c|(\hat{X}, \hat{Z}) = (x, z)) =$$
$$= \frac{P(\hat{C} = c, (\hat{X}, \hat{Z}) = (x, z))}{P((\hat{X}, \hat{Z}) = (x, z))}$$

$$= k\frac{P(C = c, (X, Z) = (x, z))}{P((\hat{X}, \hat{Z}) = (x, z))P((X, Z) = (x, z)|C = c)}$$

$$= k\frac{P(C = c)}{P((\hat{X}, \hat{Z}) = (x, z))}$$

Since $\sum_c P(\hat{C} = c|(\hat{X}, \hat{Z}) = (x, z)) = 1$ it follows that $k = P((\hat{X}, \hat{Z}) = (x, z))$ and

$$P(\hat{C} = c|(\hat{X}, \hat{Z}) = (x, z)) = P(C = c).$$

Hence

$$P(\hat{Y} = 1|(\hat{X}, \hat{Z}) = (x, z)) = \sum_c P(Y = 1|(X, Z) = (x, z), C = c)P(C = c)$$

and since C sufficient for confounding control this equals

$$\sum_c P(Y_{xz} = 1|C = c)P(C = c) = p_{xz}$$

where the last step follows from the assumption that (X, Z) has no causal effects on C. Thus

$$P(\hat{Y} = 1|(\hat{X}, \hat{Z}) = (x, z)) = p_{xz}.$$

References

Agresti, A. (2002). *Categorical Data Analysis*. John Wiley & Sons, Inc.

Alm, S. E. and Britton, T. (2008). *Stokastik - Sannolikhetsteori och statistikteori med tillämpningar*. Liber.

Greenland, S. and Brumback, B. (2002). An overview of relations among causal modelling methods. *International Journal of Epidemiology*, 31(5).

Karantia, A., Bobeckb, C., Ostermanb, M., Kardella, M., Tidemalmc, D., Runesonc, B., Lichtensteind, P., and Landén, M. (2015). Gender differences in the treatment of patients with bipolar disorder: A study of 7354 patients. *Journal of Affective Disorders*, 174.

Liero, H. and Zwanzig, S. (2011). *Introduction to the Theory of Statistical Inference*. Johns Hopkins University Press Chapman Hall/CRC.

Mondimore, F. M. (2014). *Bipolar Disorder: A Guide for Patients and Families*. Johns Hopkins University Press.

R Core Team (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Registercentrum Västra Götaland (2015). BipoläR - the swedish national quality register for bipolar disorder. http://www.psykiatriregister.se/content/bipolLast accessed 2015-04-15.

Robins, J. M., Ángel Hernán, M., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11.

Sjölander, A., Lee, W., Källberg, H., and Pawitan, Y. (2014a). Bounds on causal interactions for binary outcomes. *Biometrics*, 70(3).

Sjölander, A., Lee, W., Källberg, H., and Pawitan, Y. (2014b). Bounds on sufficient-cause interaction. *European Journal of Epidemiology*, 29(11).

Tidemalm, D., Haglund, A., Karanti, A., Landén, M., and Runeson, B. (2014). Attempted suicide in bipolar disorder: Risk factors in a cohort of 6086 patients. *PLoS ONE*, 9(4).

Xu, J. (2014). *Next-generation Sequencing : Current Technologies and Applications*. Caister Academic Press.