

PREDICTING PRINCIPAL STRATUM MEMBERSHIP IN RANDOMIZED ENVIRONMENTAL INTERVENTIONS

by
Katherine Freeland

A thesis submitted to Johns Hopkins University in conformity with the requirements
for the degree of Master of Science

Baltimore, Maryland

April 2015

Abstract

Environmental interventions targeted at reducing indoor allergens and pollutants have shown promise as a method for improving respiratory health outcomes in children by reducing exposure in the home. However, in these interventions, it is difficult to determine the effect of reduced exposure, a post-randomization variable, on the respiratory outcomes. Using principal stratification, a framework for calculating principal effects (i.e. effects within a stratum), we are able to measure the effect of reduced exposure on respiratory outcomes. These principal effects allow for the comparison of treatment effects for those who would and would not have seen a reduction in allergen (or pollutant) levels. With the exposure reduction variable, we can identify principal strata membership for some individuals in the control and treatment groups. However, the observed data only allow us partial identification of strata for other individuals. We develop a resampling based estimator that incorporates the uncertainty from fitting a model (of which the ‘true’ form is unknown) to predict likely stratum membership, classifying individuals based on this estimated probability, as well as finite sampling uncertainty. This estimation procedure allows the model form to change to best fit the resampled data at hand, reflecting the true uncertainty in this process.

Advisor: Dr. Roger D. Peng

Additional Reader: Dr. Elizabeth A. Stuart

Contents

1	Introduction	1
2	Literature Review	2
2.1	Asthma & the Environment	2
2.1.1	The Inner City Asthma Study	3
2.1.2	CCAUE Asthma Intervention Trial	4
2.1.3	Particulate Reduction Education in City Homes Study	5
2.1.4	Challenges	5
2.2	Adjustment for Post-Treatment Variables	6
2.2.1	Defining Estimands After Stratification	8
2.2.2	Estimating the Effect After Stratification	10
3	Compliance and Causal Inference	12
3.1	Assumptions	13
4	Statistical Method	13
4.1	Quantifying Uncertainty	17
5	Case Studies	19
5.1	Butz, <i>et al.</i>	19
5.2	Morgan, <i>et al.</i>	23
5.2.1	Imputation	25
5.2.2	Bla g1 Reduction in Entire Sample	25
5.2.3	Bla g1 Reduction in Sensitized Sample	28
5.2.4	Allergen Exposure Score	31
5.3	Impact of Model Selection	34
5.4	Discussion	35
6	Conclusion	37
7	References	39
8	Curriculum Vitae	42

List of Tables

1	PREACH Principal Effect Estimates	22
2	Bla g1 Reduction in Entire Sample: Principal Effect Estimates	27
3	Bla g1 Reduction in Sensitized Sample: Principal Effect Estimates	30
4	Allergen Exposure Score: Principal Effect Estimates	33

List of Figures

1	Determining Compliance with a Post-Treatment Measure	14
2	Transformation for Calculated Probabilities	17
3	PREACH Compliance Stratification	20
4	Bootstrap Distribution of PREACH Principal Effects	21
5	Proportion of Individuals per Stratum in PREACH Analysis	23
6	Bla g1 Reduction in Entire Sample: Responder Stratification	26
7	Bla g1 Reduction in Entire Sample: Bootstrap Distribution of Principal Effects	27
8	Bla g1 Reduction in Entire Sample: Proportion of Individuals per Stratum	28
9	Bla g1 Reduction in Sensitized Sample: Responder Stratification	29
10	Bla g1 Reduction in Sensitized Sample: Bootstrap Distribution of Principal Effects	30
11	Bla g1 Reduction in Sensitized Sample: Proportion of Individuals per Stratum	31
12	Allergen Exposure Score: Responder Stratification	32
13	Allergen Exposure Score: Bootstrap Distribution of Principal Effects	33
14	Proportion of Individuals per Stratum in Allergen Exposure Score Analysis	34
15	Estimates With and Without Model Form Uncertainty	35

1 Introduction

The *Guidelines for the Diagnosis and Management of Asthma* published by the National Heart, Lung, and Blood Institute have long recognized the important role the environment plays in asthma management. These guidelines recommend reducing exposure to allergens/pollutants a patient is sensitized and exposed to through the use of a "multifaceted, comprehensive approach" [5]. Their recommendations include the use of HEPA filters, allergen-impermeable encasings for pillows and mattresses, and pest control. Despite these general guidelines, understanding the efficacy of such an approach and its role in mediating asthma morbidity is still under study.

In recent years several clinical trials have provided evidence that reducing exposure to indoor pollutants/allergens is associated with improved health outcomes among inner-city children with asthma. However, investigators are unable to manipulate the indoor environment directly, and instead must randomly assign participants to receive an environmental intervention program or a control condition. The actual level of exposure occurs after randomization and is thus confounded in similar ways to treatments in observational studies. To better understand the efficacy of these interventions, we need to estimate the effect of the intervention for the subset of patients which actually experience a significant reduction in exposure.

Several strategies exist in the literature for stratifying on a post-treatment variable and estimating causal effects. The method proposed here, based on the principal stratification framework, randomly assigns individuals to latent classes, defined by their potential change in exposure under the treatment and control conditions, for which they are most likely to belong. Effects in each of these strata are then estimated using a resampling procedure that incorporates the uncertainty introduced at each step of the estimation. In contrast to other methods, the one proposed here incorporates the uncertainty associated with the form of the model used to predict likely stratum membership. The estimator allows this model to change form with each resampled data set, reflecting the true ambiguity around the choice of baseline covariates. We believe that this estimation procedure provides important information about the efficacy of these inter-

ventions and will help optimize future environmental remediation programs to achieve the largest health impact with the most efficient use of resources.

2 Literature Review

2.1 Asthma & the Environment

The Clean Air Act, passed in 1963, was the first federal legislation pertaining to the control of air pollution levels in the United States. This act specifies that National Ambient Air Quality standards must be set to protect the most sensitive members of the population. Children with asthma are one such subpopulation [21]. Several studies have been done showing that exposure to ambient air pollution is associated with increases in asthma morbidity in this population [19, 21]. However, interventions designed to reduce exposure to outdoor pollutants must be implemented at the national or state level and can be extremely costly and complex. Additionally, these type of interventions can take a long time to implement and their benefits on public health outcomes may only be realized long after the start of implementation.

Due to the delayed effects of this type of intervention, studies have been done to explore the associations between indoor air pollutant/allergen levels and asthma morbidity in children. Studies done by Rosenstreich et al. [16], Strachan and Cook [18], and Mortimer et al. [13] have all shown that exposure to indoor environmental allergens and air pollutants is an important contributor to asthma morbidity among inner-city children.¹ Therefore, it is hypothesized that reducing exposure to these indoor agents, which is relatively inexpensive, should reduce asthma morbidity among this population. Although asthma management guidelines stress the importance of environmental control measures, there has been limited evidence of their efficacy. Several clinical trials in the past decade have been conducted in attempt to demonstrate the efficacy of these types of environmental interventions. We will now review selected trials whose goal

¹Inner-city children have been shown to have the highest prevalence and mortality rates for asthma in the United States [3].

was to reduce asthma morbidity in this population through environmental remediation.

2.1.1 The Inner City Asthma Study

The Inner City Asthma Study (ICAS) was a seven-site study designed to evaluate the effectiveness of both an environmental intervention as well as a physician feedback intervention using a two-by-two factorial design. We will focus here only on the environmental intervention. Previous studies that looked at environmental interventions focused on a single allergen (or pollutant). However, Morgan et al. notes that inner-city children encounter multiple exposures each day and argue that focus should be shifted to improving the indoor environment as a whole. Consequently, the environmental intervention they implemented was multifaceted and tailored to each child's allergen sensitization. By addressing a child's entire "environmental risk profile" they hypothesized they could improve asthma symptoms and decrease the use of health care services [12].

Morgan et al. [12] found there was a greater reduction in asthma-related symptoms in the intervention group compared to the control group. Additionally, significant reductions were seen in the number of disruptions to a caretaker's plans, number of days with lost sleep (for the child and the caretaker), and number of school days missed. Furthermore, reductions were seen in the number of unscheduled visits to the emergency room and clinic as compared to the control group. They also found that levels of cockroach and dust mite allergen were decreased in the homes of both the intervention and control group participants, although this decrease was larger in the intervention group. Overall, they found the intervention decreased asthma symptoms among inner city children. This decrease translated to a 34 fewer symptom days reported among the intervention group over the 2-year study period. Another important contribution of this study was the demonstration of sustained change (i.e. the greatest reduction in symptom days was achieved two months after randomization and was maintained throughout the two year study period). They concluded that environmental remediation can be used to produce sustained reductions in allergen levels and improvements in asthma-associated morbidity.

2.1.2 CCAUE Asthma Intervention Trial

The Johns Hopkins' Center for Childhood Asthma in the Urban Environment (CCAUE) conducted a randomized controlled clinical trial to test the effectiveness of home-based interventions on reducing allergen and particulate exposure and improving the health of the asthmatic children living there [4]. Similar to ICAS, the intervention under study consisted of two components, physical (reducing allergen/pollutant exposure) and behavioral (visits from an environmental educator). Each family received a HEPA filter, allergen-proof mattress and pillow encasings, professional pest control (for families with evidence of an infestation or with a child sensitized to cockroach allergen), and mouse extermination. They enrolled 100 families, with half allocated to receive the intervention (those in the control group received the intervention at the end of the study).

They concluded that their intervention substantially reduced exposure to particulate matter and cockroach allergen among inner-city children. Additionally, they saw modest decreases in asthma symptoms among children in the intervention group. Among the treatment group, 58% of children reported daytime asthma symptoms in a two-week period at baseline. This changed to 59%, 50%, 38%, and 55% at 3, 6, 9, and 12 months, respectively. The 6, 9, and 12 month changes represent a statistically significant reduction from baseline. In the control group 50%, 55%, 66%, 60%, and 59% of children experienced daytime symptoms at baseline, 3, 6, 9, and 12 months, respectively. Significant intervention effects (comparing percentages between the treatment and control groups) were observed at 6, 9, and 12 months. Eggleston et al. hypothesized many explanations for this "less than striking" result (as compared with ICAS results published a year earlier). Even with only modest results in regards to asthma symptoms (significant decreases in number of days with symptoms but no change in lung function) Eggleston et al. concluded that the intervention was effective and felt it would be an important component in larger public health strategy to reduce symptoms among inner-city children with asthma.

2.1.3 Particulate Reduction Education in City Homes Study

Unlike the previous two trials reviewed, the Particulate Reduction Education in City Homes (PREACH) study focused on reducing the exposure of children to second hand smoke and consequently particulate matter (PM) (instead of multiple allergens *and* PM)². Like the trials that came before, the proposed intervention had both a physical and behavioral component. They provided participants in one of the intervention arms with air cleaners (to reduce PM) and behavioral coaching promoting home smoking bans. The other intervention arm received air cleaners only. These groups were compared to each other and to a control group.

The PREACH study found a significant increase of 1.36 (SD: 4.2) symptom free days in a two-week period among the children with air cleaners in the home compared to a decrease of 0.24 (SD: 3.04) symptom free days in the control group. The children in the intervention arms also saw a significant reduction in PM at 6 months. However, even after this significant reduction, PM levels in these homes were still higher than those of children who do not reside with a smoker (and higher than the 24-hour EPA standard for outdoor air quality). They found that the addition of a health coach did not provide any additional reduction in PM to just providing an air cleaner. This suggests that an improvement in health effects is associated with a reduction in PM (due to the air cleaner rather than a change in household smoking behavior).

2.1.4 Challenges

The studies reviewed all show that environmental interventions targeted at reducing indoor pollutants and allergens can lead to improvements in asthma morbidity among inner-city children. However, these studies do not allow for us to conclude that an improvement in the indoor environment *caused* the reduction in symptoms. In these studies, the reduction in pollutants/allergens occurs after randomization and is not directly under investigator control. This allows for correlations to be observed but nothing can be directly concluded about the causal nature of the relationship. This observed corre-

²Smoking is dominant contributor of PM [2]

lation could imply a causal relationship between reducing allergens/pollutants and improved health outcomes but this is subjected to the same kinds of confounding found in observational studies. To begin to understand the possible causal relationship between these variables, further statistical analysis needs to be undertaken.

Completing this additional analysis will better our understanding of the causal role that indoor allergens/pollutants play in the reduction of asthma morbidity and the factors that mediate the effects of environmental interventions. With a greater understanding of how these interventions effect health outcomes we can better optimize and tailor future interventions to achieve the maximum health benefit and to utilize resources most efficiently. The following sections introduce existing statistical methods for estimating causal effects after adjusting for post-treatment/randomization variables and propose a new method for this estimation.

2.2 Adjustment for Post-Treatment Variables

The need to adjust for a variable measured after treatment assignment (or randomization, in the case of randomized controlled studies) when that variable does not represent the outcome of interest is a common problem facing public health and medical researchers and has been addressed previously in the literature. Before beginning to review this literature, we will review the potential outcomes framework to illustrate and help define the causal estimands of interest.

We consider a simple scenario in which a group of units, $i = 1, 2, \dots, n$, can either be assigned to a treatment ($X = 1$) or the control ($X = 0$). The outcome of interest will be denoted by Y and is measured at a specific time after treatment assignment for all units. Let $Y_i(x)$ be the value of Y if unit i is assigned to treatment x . The individual-level causal effect of X on Y is the comparison between $Y_i(0)$ and $Y_i(1)$ (e.g. $Y_i(1) - Y_i(0)$). However, these potential outcomes are not completely observable since we can only observe $Y_i(0)$ or $Y_i(1)$ for each i (unit).

Since we cannot estimate individual-level causal effects, we estimate the average

treatment effect in a given population [9]. We denote this as,

$$\delta = \mathbb{E}[Y_i(1) - Y_i(0)]$$

Treatment assignment is strongly ignorable (i.e. treatment assignment is independent of the potential outcomes and each individual has a non-zero probability of being assigned to either treatment) when units are randomized to receive treatment. In randomized controlled trials, this effect is equivalent to,

$$\delta = \mathbb{E}[Y|X = 1] - \mathbb{E}[Y|X = 0]$$

Commonly, we wish to define this treatment effect for a subset of the population defined by an additional covariate of interest S_i measured on each unit after treatment assignment (e.g. change in level of particulate matter). However, S_i generally contains both information about the unit as well as the treatment of interest. In these instances, we are interested in the effect of X on Y after adjusting for S . We wish to estimate this in such a way that the resulting estimate is a causal effect.

Rosenbaum [15] discussed the consequences of adjusting for a post-treatment variable that has been affected by the treatment. With S_i affected by the treatment, we must again consider the potential outcomes framework. For each unit, let $S_i(x)$ be the value of S (the post-treatment variable of interest) for unit i if assigned to treatment x . Once again, we can only observe $S_i(0)$ or $S_i(1)$ for each individual. Rosenbaum goes on to examine the expected difference in $Y_i(1)$ and $Y_i(0)$ after adjusting for the observed value of $S_i(x)$. This is expressed as,

$$\tilde{\delta} = \mathbb{E}[Y_i(1)|X = 1, S_i(1) = s] - \mathbb{E}[Y_i(0)|X = 0, S_i(0) = s]$$

and is called the adjusted treatment difference. The quantity explored by Rosenbaum is the net treatment difference, defined as $\tilde{\delta} - \delta$, which is commonly discussed in attempt to highlight the mechanism by which the treatment produces its effect. However, this

net treatment difference is not a causal effect. The two groups of subjects, those with a post-treatment value of s under control ($S_i(0) = s$) and those with a post-treatment value of s under treatment ($S_i(1) = s$) may not be a comparable group of subjects. By conditioning on observed values of S we introduce post-treatment selection bias in the estimation of causal effects.

2.2.1 Defining Estimands After Stratification

To estimate causal effects, after adjusting for values of S , Frangakis and Rubin introduce the principal stratification framework. Their adjustment for S always results in causal effects because it compares potential outcomes for a common set of individuals. Principal stratification P is defined with respect to S of $i = 1, \dots, n$ units such that within any set of P , all units have the same vector of potential outcomes, $(S_i(0), S_i(1))$. Let, S_i^P indicate the stratum of P which the i^{th} unit belongs to. Thus, the estimate of a principal effect is the comparison between $\{Y_i(1) : S_i^P = s\}$ and $\{Y_i(0) : S_i^P = s\}$. By conditioning on the value of S under the treatment and control condition we are able to estimate unbiased treatment effects. These principal effects are critical for the understanding of how treatments act on units. If these strata were known, investigators could estimate the average principal causal effect in each stratum. However, these strata are never completely known in practice and must be estimated.

Joffe, Small, and Hsu [11] compare and contrast several methods for both defining and estimating effects among groups defined by a post-treatment variable (also termed “auxiliary outcome”). In addition to the principal stratification method proposed by Frangakis and Rubin [6], Joffe, Small, and Hsu consider stratification on a single potential outcome, an observed auxiliary variable, an expected auxiliary variable, and multiple expected auxiliary variables. Lastly, they consider a conventional approach where subgroup membership is based solely on pretreatment covariates. For simplicity, we will assume that S is a binary variable taking on values of 0 or 1.

For stratification on a single potential outcome we are interested in the comparison of individuals belonging to two groups, say $\{Y_i(1)|S_i(1) = s\}$ and $\{Y_i(0)|S_i(1) = s\}$.

This is the comparison of outcomes under treatment and control conditions for individuals who have the same value of s under treatment.³ Membership in this particular stratum is only partially observed (i.e. we cannot observe both $S_i(1)$ and $Y_i(0)$ for a single unit i) and thus complicates causal effect estimation. Methods used for principal effect estimation can be used here, as well as techniques for observed auxiliary variable stratification discussed next.

When stratifying on an observed auxiliary variable, we compare $\{Y_i(1)|S = s, X = x\}$ and $\{Y_i(0)|S = s, X = x\}$. These are the potential outcomes under treatment and control for individuals with the same value of S and X . Here we are conditioning on observed values (rather than the potential outcomes) of S and X so the subgroups are fully observable. However, these values are not able to be observed at the time of treatment (like principal stratification and single potential outcome stratification). Thus, effects of this nature are more explanatory (e.g. to explain differences between treatment groups).

Another approach, expected auxiliary variable stratification, incorporates baseline covariate information collected at the time of treatment assignment. We define the effect for a group of individuals for whom $S = 1$ with a certain probability, conditioned on baseline covariates \mathbf{Z} and treatment assignment. We denote this probability as $\mu^x(\mathbf{Z}) = \mathbb{E}(S_i(x)|\mathbf{Z})$. This value, $\mu^x(\mathbf{Z})$, has been referred to as the “principal score”. We then make comparisons between potential outcomes for individuals with the same principal score (or with a principal score above/below some cutoff). This can be written as, $\mathbb{E}[Y_i(1)|\mu^x(\mathbf{Z})] - \mathbb{E}[Y_i(0)|\mu^x(\mathbf{Z})]$. This can be extended to the case of expected multiple auxiliary stratification where groups are defined by both $\mu^x(\mathbf{Z})$ and $\mu^{x'}(\mathbf{Z})$ where $x \neq x'$.

Lastly, Joffe, Small, and Hsu [11] consider the case where strata are fully determined by baseline covariate values. The different estimands explored previously have value in different scenarios. In practice, clinicians must prescribe the treatment they

³We could also compare $\{Y_i(1)|S_i(0) = s\}$ and $\{Y_i(0)|S_i(0) = s\}$. This is the comparison of outcomes under treatment and control conditions for individuals who have the same value of s under control.

believe will provide the best possible result. This decision is typically made with only baseline covariate data in hand. Consequently, estimates stratifying on baseline covariate data *only* may be most beneficial in these situations. However, in another scenario, stratifying on the expected value of a post-treatment variable may be extremely useful. Joffe, Small, and Hsu considers the situation in which a diagnostic test can very accurately predict which patients will and will not develop a post-treatment auxiliary outcome (S). Here, the treatment effect stratified by observed values of S is extremely relevant. Another situation in which estimating the causal effect after stratifying by a post-treatment variable is beneficial is when treatment is administered over an extended time period and the post-treatment variable is observed relatively early on. If this variable is observed relatively quickly, treatment can be changed based on this new information.

2.2.2 Estimating the Effect After Stratification

Joffe, Small, and Hsu [11] also reviewed estimation methods for the different estimands described above. When stratifying on expected auxiliary variables, a simple way of estimating this effect is to 1) estimate the expected auxiliary variable $\mu^x(\mathbf{Z})$ and 2) estimate the effects by level of this expected variable. The first step can be done using a regression model (regressing S on \mathbf{Z} , stratifying by X). The second step can be accomplished using standard methods under the ignorable treatment assignment assumption. For stratification on an observed auxiliary variable, estimation is much more complex. Joffe, et al. sketch an estimation procedure similar to G-estimation in structural nested models when Y is continuous. The approach they outline is semiparametric and is valid for structural distribution models. However, if Y is binary, mean models are required and the usual logit link is needed to formulate the model.

Principal strata (PS), as defined by Frangakis and Rubin [6], are not determined completely by observed data. Thus estimation for principal effects will be more dependent on additional assumptions and restrictions. We note that the estimation techniques used for PS can also be used for stratification on an observed auxiliary variable or a sin-

gle potential outcome. Since these stratification strategies focus on certain strata (rather than all strata), we can marginalize over the unobserved levels of S after estimating principal effects. In estimating principal effects, we generally must impose various types of restrictions on the parameters. These are discussed in detail by Joffe, Small, and Hsu. A Bayesian approach [10] which combines prior information with the likelihood to produce a posterior and approaches based on deriving bounds on causal effects [17] do not require such restrictions.

The principal stratification framework introduced by Frangakis and Rubin was used by Peng et al. [14] to better understand and identify the causal role that indoor pollutants play in asthma morbidity among inner city children. The treatment variable X is defined as the assignment to an environmental intervention ($X = 1$) or control ($X = 0$). The post-treatment variable is the experience of at least a $20 \mu\text{g}/\text{m}^3$ reduction in $\text{PM}_{2.5}$ at 6 months ($S = 1$) or not ($S = 0$). Peng et al. make the monotonicity assumption [1], excluding the group commonly referred to as the ‘defiers’. None of the exclusion restrictions were made (see Section 3.1 for a description of these assumptions and restrictions).

Peng et al. took a Bayesian approach to estimating principal effects that allowed the incorporation of existing knowledge about the effects of environmental interventions on asthma morbidity. They modeled the outcome Y_i with a normal distribution with mean μ_{px} and σ_p^2 where p denotes the stratum specific parameter. Here, the outcome, is the difference in symptom free days between baseline and 6 month follow-up. A complete-data posterior distribution of the model parameters was constructed. Using a Gibbs sampler, samples were drawn from this distribution to construct the full conditional distributions of the parameters. The parameters of interest were the differences in the stratum specific means, $\theta_p = \mu_{p1} - \mu_{p0}$. This approach allowed Peng et al. to estimate the effect of the intervention for the subset of children who would experience a reduction of at least $20 \mu\text{g}/\text{m}^3$ under treatment. They found that this principal effect was much greater than the overall effect found in the original study.

3 Compliance and Causal Inference

The clinical trials reviewed previously share a common controlled intervention: the randomization of individuals to receive an environmental intervention or not. They also all share a common estimand: the causal effect of *randomization* to a particular treatment (i.e. the intention-to-treat effect, assignment to treatment vs. control) on asthma symptoms after a given time. Another important estimand in these studies is the effect of the *reduction in exposure levels* on asthma symptoms at a given time. This estimand, taking into account the change in an individual’s allergen/pollutant exposure, provides information about the efficacy of the treatment (i.e. the effect of reducing allergen exposure). To begin to estimate this effect, we must adjust for a post-treatment measure of compliance.

Of arguably most interest is the complier average causal effect, where compliers are individuals for whom the intervention changes their actual treatment (i.e. who take the treatment when assigned to the treatment arm and do not take the treatment when assigned to the control arm). In the framework of environmental interventions, “taking the treatment” could be defined, for example, as experiencing a certain reduction in allergen exposure. When administering an environmental intervention there is the possibility of two-sided noncompliance. These two sides exist because individuals in the treatment group can *not* experience a reduction in exposure and, alternatively, individuals in the control group *can* experience the reduction in exposure. These noncompliers can be divided into three distinct categories based on noncompliance behavior. Always-takers are individuals who *always* receive the treatment regardless of assignment. In contrast, never-takers are those who *never* receive the treatment regardless of assignment. Defiers are those who experience the opposite of their treatment assignment, i.e., they receive the treatment if in the control group and do not receive the treatment if in the treatment group.

3.1 Assumptions

Several common assumptions are made when estimating the complier average casual effect. First, regarding the defiers, we make the monotonicity assumption [1]. This assumption posits that assignment to the treatment arm can only *increase* the probability of receiving the treatment and thus, eliminates this specific type of noncompliance. This appears to be a reasonable assumption in the context of these types of environmental interventions. Additionally we make the stable unit treatment value assumption (SUTVA) which assumes no interference between units (e.g. the reduction of exposure of one individual does not affect the number of symptom free days of another individual) [17]. This assumption also states that only one version of the treatment exists. Third, we assume that the treatment assignment mechanism is ignorable. This is a valid assumption in the context of randomized trials being considered here. Commonly, exclusion restrictions are made as well. These restrictions state that randomization to a treatment does not affect the outcome if it does not affect the intervention actually received [1, 10]. This assumption translates into a null effect for the always-takers and the never-takers described previously. We do not make this restriction here because there is also scientific interest in the effect of these interventions that is not associated with reduced exposure. Sizeable effect estimates in these populations would suggest that the interventions under investigation are impacting the outcome through other mechanisms and may spur further investigation.

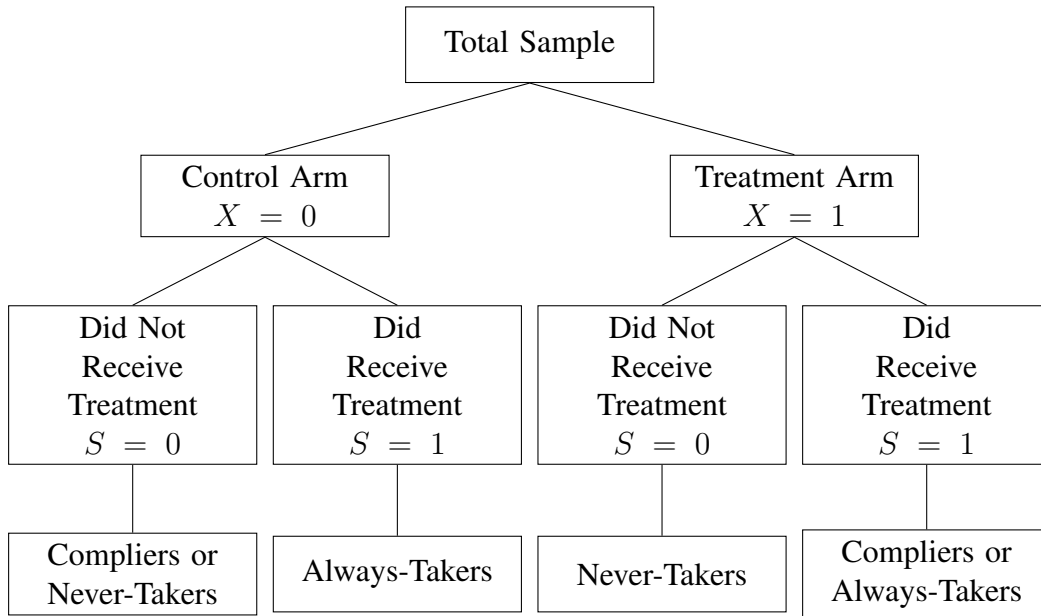
4 Statistical Method

To compare treatments after adjusting for a post-treatment measure of compliance we use the principal stratification framework developed by Frangakis and Rubin [6]. We use S to denote the post-treatment measure of compliance. For simplicity we make S binary, taking a value of 1 if the intervention was received and 0 if not. The strata defined by S are called principal strata and the effects of treatment after adjusting for S are called the principal effects. Since S is a measure of compliance, the three strata rep-

resent the compliers, always-takers, and never-takers (recall, we make the monotonicity assumption described in Section 3.1).

To determine the value of S for a given individual we look at the change in allergen exposure (or alternatively, change in $\text{PM}_{2.5}$) from baseline to end of study. For example, an individual received the treatment, $S = 1$, if he/she experienced a 50% or greater reduction in allergen exposure and that he/she did not receive the treatment, $S = 0$, if he/she experienced a reduction less than 50%.⁴ The following tree in Figure 1 illustrates this stratification.

Figure 1: *Determining Compliance with a Post-Treatment Measure*



As seen in Figure 1, the compliance measure, S , does not completely determine strata membership (because of possible two-sided noncompliance). Some individuals in both the treatment and control groups have stratum membership that is only *partially identified* (e.g. an individual in the control group where $S = 0$ is either never-taker or a complier). To estimate the principal effects, we must determine stratum membership for these individuals.

To estimate the principal effect among compliers we must determine which of the individuals in the control group for whom $S = 0$, and which of the individuals in

⁴This definition can be altered to reflect what is felt to be a clinically meaningful definition of “receiving treatment”

the treatment group for whom $S = 1$, are most likely to be compliers. To do this we calculate the probability, given S , the treatment assignment ($X = 1$ for treatment group, 0 otherwise), and a vector of covariates \mathbf{Z} , that an individual is a complier.

All individuals in the control group for whom $S = 0$ are either never-takers or compliers. This can be written as,

$$1 = P(N|X = 0, S = 0, \mathbf{Z}) + P(C|X = 0, S = 0, \mathbf{Z})$$

Thus, by determining one of these probabilities, we can completely determine the other. Using the rules of conditional probability we have,

$$\begin{aligned} P(N|X = 0, S = 0, \mathbf{Z}) &= \frac{P(N, X = 0, S = 0|\mathbf{Z})}{P(X = 0, S = 0|\mathbf{Z})} \\ &= \frac{P(N, X = 0|\mathbf{Z})}{1 - P(A|X = 0|\mathbf{Z})} \\ &= \frac{P(N|X = 0|\mathbf{Z})P(X = 0)}{1 - P(A|X = 0|\mathbf{Z})} \end{aligned}$$

To estimate $P(N|X = 0|\mathbf{Z})$ for each individual in the control group, we will build a logistic regression model using the the individuals in the treatment group (where status as never-taker or not never-taker is identified). In building this model we make the assumption that $P(N|X = 1, \mathbf{Z}) = P(N|X = 0, \mathbf{Z})$. This assumption holds when treatment is randomly assigned. This regression model is constructed using baseline covariate data only. $P(X = 0)$ is the marginal probability of being in the control group (e.g., 0.5 if there is equal allocation between treatment groups). The computation of $P(A|X = 0, \mathbf{Z})$ is discussed below. With the probability of being a never-taker in the control group determined, the probability of being a complier in the control group is,

$$P(C|X = 0, S = 0, \mathbf{Z}) = 1 - P(N|X = 0, S = 0, \mathbf{Z})$$

Similarly, we can determine the probability of being an always-taker or a complier in the treatment group. Again, all individuals in the treatment group for whom $S = 1$

belong to one of these two groups, expressed as,

$$1 = P(A|X = 1, S = 1, \mathbf{Z}) + P(C|X = 1, S = 1, \mathbf{Z})$$

Similarly, we have

$$\begin{aligned} P(A|X = 1, S = 1, \mathbf{Z}) &= \frac{P(A, X = 1, S = 1|\mathbf{Z})}{P(X = 1, S = 1|\mathbf{Z})} \\ &= \frac{P(A, X = 1|\mathbf{Z})}{1 - P(N|X = 1, \mathbf{Z})} \\ &= \frac{P(A|X = 1|\mathbf{Z})P(X = 1)}{1 - P(N|X = 1, \mathbf{Z})} \end{aligned}$$

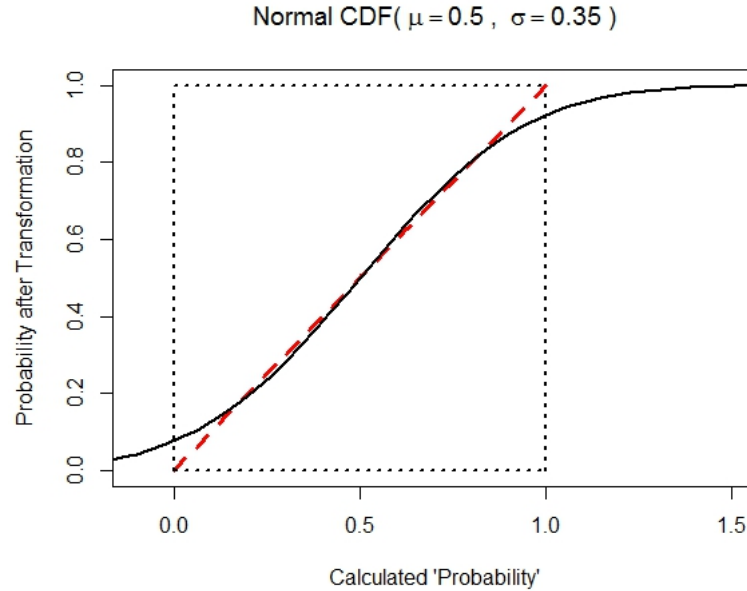
To estimate $P(A|X = 1, \mathbf{Z})$ for each individual in the treatment group, we will build a logistic regression model using the individuals in the control group (where status as always-taker or not always-taker is identified). Again we make the assumption, because of the randomization of treatment assignment, that $P(A|X = 0, \mathbf{Z}) = P(A|X = 1, \mathbf{Z})$. This second model will be constructed using only baseline covariate data, as well. $P(X = 1)$ is the marginal probability of being in the treatment group and $P(N|X = 1, \mathbf{Z})$ is calculated using the model described previously. Consequently, the probability of being a complier in the treatment group is,

$$P(C|X = 1, S = 1) = 1 - P(A|X = 1, S = 1)$$

With the logistic regression models estimating $P(A|X, \mathbf{Z})$ and $P(N|X, \mathbf{Z})$, we compute $P(C|X, S, \mathbf{Z})$ and $P(A|X, S, \mathbf{Z})$ or $P(N|X, S, \mathbf{Z})$ for the partially identified individuals using the formulas above. The models used to estimate these probabilities are only approximations (since we do not know the true model) and only provide a ‘best guess’ at the probability of principal stratum membership. Additionally, although the control and the treatment group are “statistically” identical (i.e. they only randomly differ), the models built to predict always-taker and never-taker status are built using two different populations. Consequently, we may encounter scenarios when $P(A|X = 1, S = 1) > 1$ or $P(N|X = 0, S = 0) > 1$. To avoid assigning proba-

bilities greater than 1, we will transform the probabilities calculated using the formulas above using a Normal cumulative density function with a mean of 0.5 and a standard deviation of 0.35 as seen in Figure 2.

Figure 2: *Transformation for Calculated Probabilities*



This particular transformation was chosen because it preserves most probability values in the allowable range while still pulling in values above this range. With these probabilities in hand, we can assign membership to the most likely principal strata and calculate the principal effects by comparing outcomes, within strata, between the treatment and control group.

4.1 Quantifying Uncertainty

If the stratum membership were known for each individual we could simply estimate the principal effect by looking at the difference in symptom free days (the outcome) in the treatment and control group within each stratum. However, we are estimating the membership of the partially identified individuals and thus have added uncertainty in our estimate that must be accounted for.

To quantify the uncertainty associated with an effect estimated in this way, a resampling-based estimator is proposed. The following steps outline the resampling procedure:

1. Resample, with replacement, from the original data. This is done separately in the treatment and control group to preserve the marginal probability of being assigned to treatment
2. Use Lasso regression models, built with baseline covariates, to estimate the probability of being an always-taker (using the individuals in the control group) or a never-taker (using the individuals in the treatment group). Note: The L1 penalty parameter acts as a sort of variable subset selector, allowing variables to be excluded from the model based on the resampled data set [20].
3. Compute the probability p_i of being a complier for each partially identified individual.
4. Randomly assign principal stratum membership using a Bernoulli(p_i) random variable
5. Calculate principal effects based on these strata assignments
6. Repeat steps 1-5 B times ($B = 10,000$ is recommended)
7. Calculate the mean and desired percentiles of the bootstrapped distributions of principal effects

By resampling from the original data and using that data to determine the ‘best’ model, the uncertainty about the form of the model as well as finite sampling uncertainty is incorporated in the procedure. Assigning stratum membership with a Bernoulli random variable incorporates the additional uncertainty about stratum membership for those partially identified individuals. Thus, the estimated mean and percentiles of the bootstrapped distribution of effects represent plausible values for the true effect and the uncertainty in the estimation process. Next, this method will be applied to data collected in two randomized controlled trials of environmental interventions.

5 Case Studies

5.1 Butz, *et al.*

The Particulate Reduction and Education in City Homes (PREACH) study was a three-armed randomized controlled trial. The goal of the study was to assess the effect of introducing air cleaners into childrens' homes, along with a behavioral intervention promoting home smoking bans, on asthma morbidity among inner city children. To be eligible children must have physician diagnosed asthma and live, for at least 4 days a week, with a smoker who smokes 5 or more cigarettes a day. The 3 arms of the trial were 1) air cleaner only, 2) air cleaner plus health coach, and 3) control. All groups received asthma education from a nurse during four home visits.

The pollutant targeted by the intervention was $PM_{2.5}$, which has been shown to be associated with an increase in asthma symptoms. The original study concluded that the addition of a health coach was not associated with either a reduction in $PM_{2.5}$ or improvement in asthma symptoms. Therefore, we combine the air cleaner plus health coach group with the air cleaner only group to create a single "treated" arm for this secondary analysis.

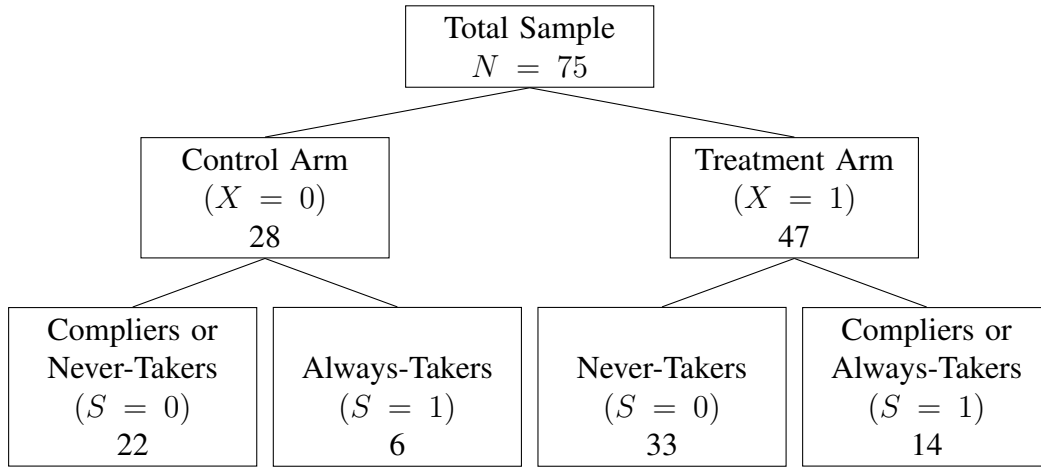
The outcome of interest is symptom free days (SFD) which is defined as the number of days, in a two week period, in which a child experiences no asthma symptoms. Butz et al. [2] concluded that subjects in the air cleaner groups experienced larger increases in SFDs between baseline and 6-month follow-up compared to the control group. The air cleaner groups also saw larger decreases in $PM_{2.5}$ levels from baseline to 6 months compared to the control group. This may suggest that the treatment affected the outcome through the lowering the $PM_{2.5}$. The new method presented in Section 4 was applied to this data set to quantify the effect of lowering $PM_{2.5}$ on SFDs.

The PREACH study collected information on 126 children. However, 48 children were missing information on $PM_{2.5}$ levels at either baseline or 6 months. Without these measurements we are unable to calculate the reduction in $PM_{2.5}$ and consequently S , the intermediate compliance variable. These 48 children were excluded from the following

analysis. Additionally, 3 children were missing information on the baseline covariates that are used to construct the models predicting stratum membership. These children were also excluded, leaving 75 subjects available for analysis.

We are interested in the effect of the treatment for the compliers, the group of individuals who experience a significant decrease in $PM_{2.5}$ while assigned to treatment and do not when assigned to the control. We define a significant decrease to be $20 \mu g/m^3$ and denote this as $S = 1$. We begin by stratifying individuals based on their values of S . For this data set, this can be seen in Figure 3.

Figure 3: *PREACH Compliance Stratification*



In Figure 3 individuals in the control group who experienced a significant reduction in $PM_{2.5}$ are always-takers (6). Individuals in the treatment group who did *not* see a significant reduction are never-takers (33). The remaining individuals are partially identified. We used the method described in Section 4 to classify these individuals and calculate the corresponding principal effects.

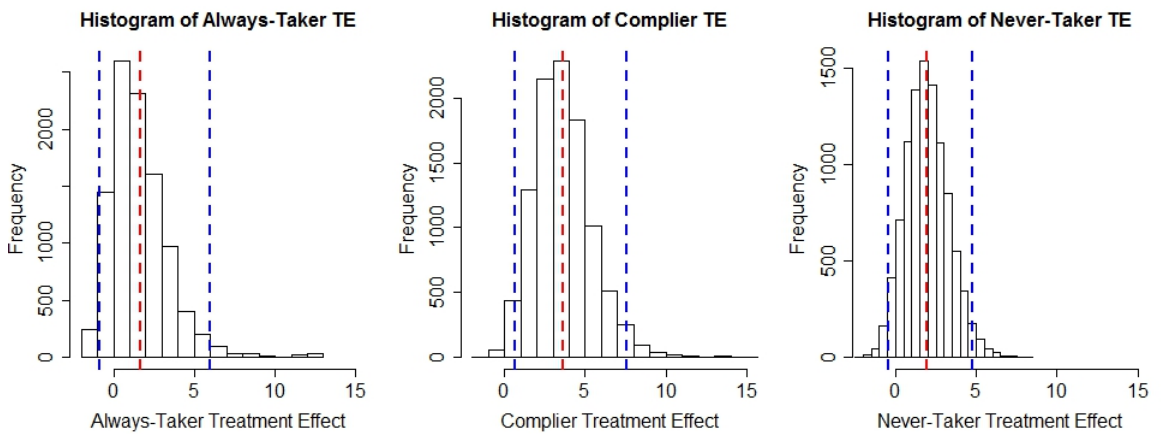
The data was resampled, with replacement, 10,000 times. The control and treatment arms were resampled separately to maintain the marginal probability of being assigned to either group. Additionally, we required that at least 1 always-taker and 1 never-taker to be in the control and treatment groups, respectively. Using the 47 individuals in the treatment group, a lasso regression model (with a constraint of 0.5) was fit to predict never-taker status. With the 28 individuals in the control group another lasso regression model (with a constraint of 0.5) was fit to predict always-taker status. This was repeated

in each of the 10,000 resampled data sets. The baseline covariates used in both of these models were age, a dichotomous measure of asthma severity, and baseline levels of $\text{PM}_{2.5}$ and NO_2 .

Using the formulas and transformation introduced in Section 4 the probability of being a complier, p_i , was computed for the 22 individuals in the control group and the 14 individuals in the treatment group that were only partially identified. Membership to this group was assigned using a $\text{Bernoulli}(p_i)$ random variable. The outcome is change in SFDs from baseline to 6-month follow-up. The treatment effect is the difference in average change in SFDs between the treatment and control group within each stratum. We note that one may wish to calculate the principal effects by fitting a regression model in each stratum. This was not done for this data set because in some iterations certain stratum (especially the always-takers, see Figure 5) were extremely small and fitting this regression model would be inappropriate. However, this was done in the other case study presented in Section 5.2.

The 10,000 principal effect estimates obtained from applying this procedure to each resampled data set formed the bootstrap distributions of principal effects seen in Figure 4.

Figure 4: *Bootstrap Distribution of PREACH Principal Effects*



The red dashed line represents the mean of the distribution while the blue dashed lines denote the 2.5% and 97.5% percentiles. Table 1 compares these estimates to the overall treatment effect calculated by Butz et al. [2]. For the 75 individuals included in

this analysis, the air cleaner groups experienced a 2.1 (95% CI: [0.3, 3.8]) day increase in SFDs compared to the control group. The principal effect among the compliers (those who would experience a reduction in $PM_{2.5}$ under treatment but not under control) is more than 1.5 times larger than this overall estimate. The principal effects among the always-takers and the never-takers are similar in magnitude to each other (and smaller than the complier principal effect) and have bootstrap intervals that contain 0 (indicating a non-significant effect).

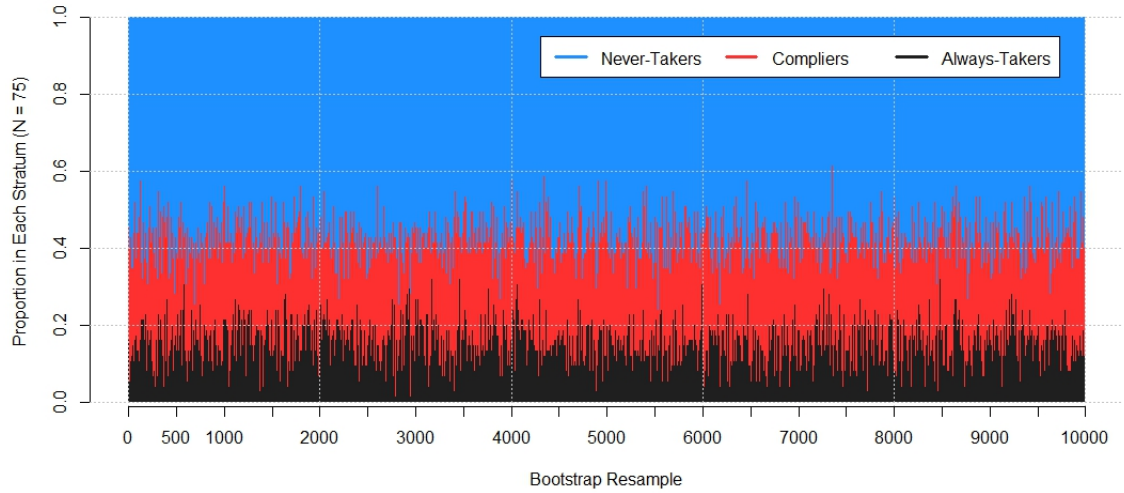
Table 1: PREACH Principal Effect Estimates

	Always-Takers	Compliers	Never-Takers	Overall
Original Analysis				2.1 _(0.3,3.8)
Principal Effects Analysis	1.64 _(-0.86,5.97)	3.63 _(0.64,7.60)	1.95 _(-0.45,4.75)	

Although we posit that the strong effect found amongst the complier group is the result of a reduction in $PM_{2.5}$, it is possible that this effect may be explained by some other covariate. Secondly, we note that the decision to dichotomize reduction in $PM_{2.5}$ to create the binary compliance variable, S , is not without its limitations and drawbacks. Overall, the complier average principal effect allowed us to quantify the effect of the intervention on the sub-population who experienced a significant reduction in $PM_{2.5}$, demonstrating that these interventions produce clinically relevant effects and should be studied further.

There was also interest in examining how stable the stratum size was for the 10,000 resampled data sets. The proportion in each stratum was plotted, by iteration, in Figure 5. From this plot we can see that the each resampled data set contained a large proportion of never-takers (between 35% and 77%). The next largest group, on average, was the compliers (between 5% and 53%). The smallest group was the always-takers, which comprised between only 2% to 37% for any given sample. Given the limited sample size ($N = 75$), the proportion of the compliers in each iteration of the estimation procedure may be problematic when estimating an effect in this stratum.

Figure 5: *Proportion of Individuals per Stratum in PREACH Analysis*



5.2 Morgan, *et al.*

The Inner-City Asthma Study used a two-by-two factorial design to explore the effect of a multifaceted, individually tailored environmental intervention and a physician-feedback intervention on asthma symptoms and the use of asthma-related health care services. Morgan et al. [12] found that there was no interaction between these two interventions and thus their effects are considered separately. To be eligible children must have been between 5 and 11 years of age, have physician diagnosed asthma, be resident of a census tract in which at least 20% of households are below the poverty line, have at least one hospitalization or two unscheduled clinic visits that were asthma related in the last 6 months, and have a positive skin test for at least 1 of 11 indoor allergens. This secondary analysis focused on the results of the environmental intervention (as did Morgan et al. [12]).

Participants were randomized to either the environmental intervention arm or the control arm. Participants in the control arm were evaluated at home visits at 6-month intervals. The goal of the environmental intervention arm was to provide caretakers with the information, supplies, and skills needed to implement a comprehensive environmental remediation. This intervention was organized into 6 modules that focused on reducing exposure to pet, dust mite, cockroach, mold, and rodent allergens. These

interventions were tailored to the results of each participants skin test. During the 1 year intervention period, investigators conducted 5 home visits.

The original outcome of interest was maximal number of symptom days in a two week period. The study investigators found that the maximal number of symptom days was lower in the intervention group by 0.82 days per 2-week period. They also found that levels of cockroach allergen (Bla g1) and dust mite allergens (Der f1 and Der p1) were decreased in both groups over the course of the study, although greater reductions were realized in the intervention group. The original outcome was modified and expressed as number of symptom free days (i.e. $14 - \text{Max No. of Symptom Days}$) to better align with the results of the PREACH analysis. The new method presented in Section 4 was applied to this data set to quantify the effect of lowering allergen exposure on SFDs.

The Inner-City Asthma Study enrolled 937 children at 7 possible sites (Bronx, Boston, Chicago, Dallas, New York City, Seattle/Tacoma, and Tuscon). However, only 869 participants had at least one follow-up assessment where allergen exposure levels and symptom information was collected. Three different sub-analyses were undertaken. The first looked at the effect of a 50% reduction in Bla g1 for all participants. The second looked at the effect of a 50% reduction in Bla g1 for the subset of participants that had a positive skin test for Bla g1. Finally, the last analysis looked at the effect of a 50% reduction in an allergen exposure score.

For each of these different analyses, we were interested in the effect for the compliers, the group of individuals who experienced a significant decrease in the allergen of interest when assigned to treatment and who do not when assigned to the control. We defined a significant decrease to be a 50% reduction in the allergen (or allergen score) of interest and denoted this as $S = 1$. For each analysis only individuals with an allergen measurement at both baseline and 1 year were included. A comparison between the included and excluded individuals was undertaken in each of the three analyses.

5.2.1 Imputation

An additional imputation step was added to the procedure described in Section 4. The Inner-City Asthma Study investigators collected information on many baseline covariates that could be used to build the always-taker and never-taker models. However, not every individual had observed measurements for all of these covariates. Covariates with more than 40% missingness were discarded, but for the remaining covariates multiple imputation (MI) was undertaken. Additionally, a small amount of missingness ($<10\%$) was present in the outcome variable. These outcome values were imputed as well. MI was performed with the multivariate imputation chained equations (MICE) procedure in R, version 3.1.0. Fifty complete data sets were generated using the MICE procedure (imputation was done separately in the control and treatment arms). 10,000 resampled data sets were created by drawing 200 bootstrap samples from each of the 50 MI datasets. This procedure is thought to propagate a comparable amount of uncertainty to resampling from the incomplete data set and imputing the missing values 10,000 times [8].

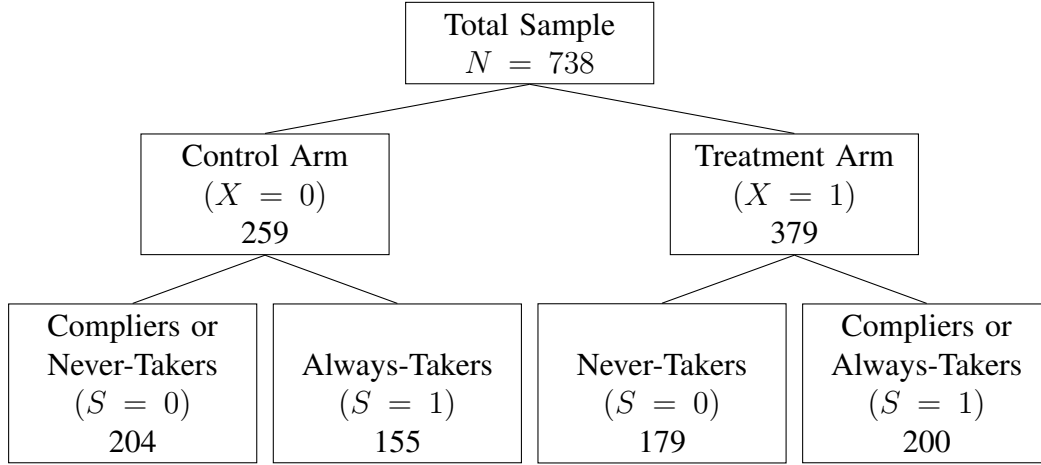
5.2.2 Bla g1 Reduction in Entire Sample

For the 869 individuals with at least one follow-up visit, we wished to estimate the effect of a 50% reduction in cockroach (Bla g1) allergen (as measured on the floor) between baseline and 1 year. Of these 869 individuals, 131 individuals were missing a Bla g1 measurement at either baseline, 1 year, or both. These individuals were excluded from the analysis. The main demographic characteristics did not differ greatly between the 738 individuals included and the 131 excluded. However, the included individuals had a slightly higher proportion of males (64% vs. 52%) and had fewer individuals from the New York City site (13% vs. 28%).

We begin by stratifying individuals based on their value of S . For this subset, this can be seen in Figure 6.

In Figure 6 individuals in the control group who experienced a significant reduction in Bla g1 are always-takers (155). Individuals in the treatment group who did not see

Figure 6: *Blinding Reduction in Entire Sample: Responder Stratification*



a significant reduction are never-takers (179). The remaining individuals are partially identified. The missing data was imputed and resampled as described in Section 5.2.1. Using the 379 individuals in the treatment group and the 359 individuals in the the control group, a lasso regression model (with a constraint of 0.5) was fit to predict never-taker status and always-taker status, respectively. This was repeated in each of the resampled data sets. The baseline covariates eligible for inclusion were gender, age, race, caretaker characteristics, allergen sensitivities, asthma related symptoms, and allergen exposure levels.

Using the formulas and transformation introduced in Section 4, the probability of being a complier, p_i was computed for the 204 individuals in the control group and the 200 individuals in the treatment group. Membership was assigned using a Bernoulli(p_i) random variable. Principal effects were estimated by fitting a linear regression model in each stratum. This model, predicting change in SFDs, included all of the covariates used to predict stratum membership and a treatment indicator variable. These estimates, the coefficients of the treatment indicator variable, formed the bootstrap distributions seen in Figure 7

The red dashed line represents the mean of the distribution while the blue dashed lines denote the 2.5% and 97.5% percentiles. Table 2 compares these estimates to the overall treatment effect calculated by Morgan et al. [12].

The principal effect among the compliers is 0.80 SFDs which is similar to the over-

Figure 7: *Bla g1 Reduction in Entire Sample: Bootstrap Distribution of Principal Effects*

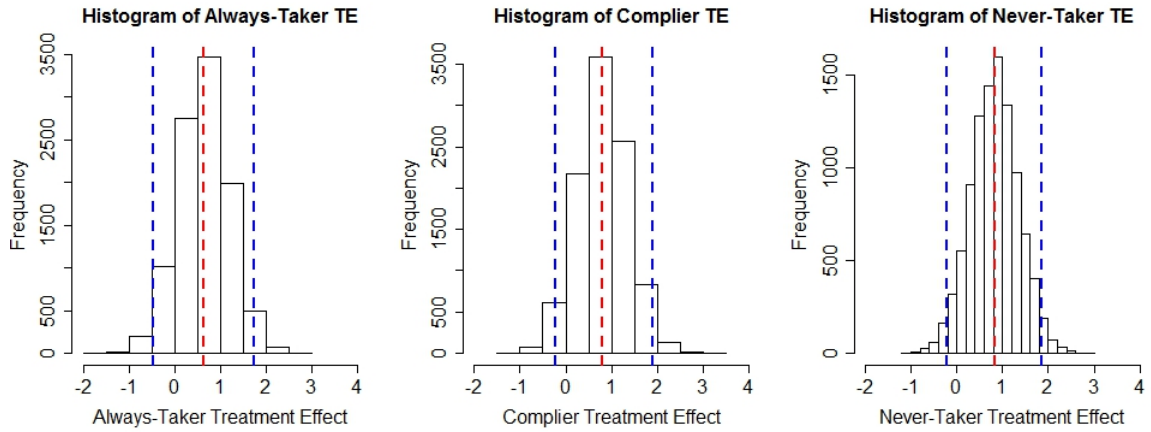


Table 2: *Bla g1 Reduction in Entire Sample: Principal Effect Estimates*

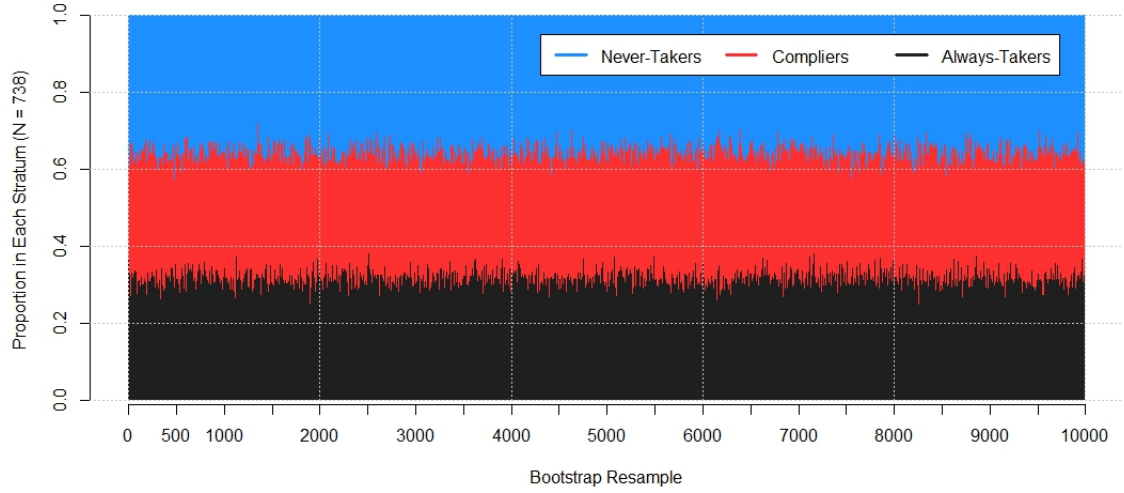
	Always-Takers	Compliers	Never-Takers	Overall
Original Analysis				0.82 _{p<0.001}
Principal Effects Analysis	0.64 _(-0.47,1.72)	0.80 _(-0.23,1.89)	0.83 _(-0.20,1.85)	
Site				
NY/Bos/Chi	0.62 _(-0.67,1.87)	0.81 _(-0.46,2.10)	1.01 _(-0.30,3.27)	
Seattle	-0.05 _(-3.36,3.09)	0.97 _(-2.24,4.06)	-0.54 _(-3.32,2.20)	
Southwest	1.62 _(-1.73,4.63)	0.62 _(-1.96,3.12)	1.30 _(-0.46,3.12)	

all effect found by the original investigators. However, the confidence interval of the estimate (as well as the intervals for the always-takers and never-takers) obtained here is much wider. Possible explanations for these wider intervals will be discussed later. The magnitude of effect is similar among all three strata. This may imply that our definition of “receiving the treatment” may not be appropriate. Principal effects were also examined by site (sites were grouped such that participants from New York City, the Bronx, Boston and Chicago were in one group, participants from Seattle/Tacoma were in another, and participants from Dallas and Tuscon made up the final group). These were calculated with a regression model similar to the one described above, but with the addition of a site by treatment assignment interaction term. This method was chosen to prevent the fitting of a very large and complex model on small subgroup (e.g. never-takers from Seattle). There appears to be heterogeneity of effects at these different locations.

Again, we were interested in examining the stability of stratum size for the 10,000

resample data sets. This information is plotted in Figure 8.

Figure 8: *Bla g1 Reduction in Entire Sample: Proportion of Individuals per Stratum*



It appears the strata stayed approximately the same size in each iteration of the estimation procedure. The width (indicating the proportion in a given sample) seems to be approximately equal for the never-takers, compliers, and always-takers. Combined with the large sample size ($N = 738$), this would suggest there was a sufficient number of individuals in each stratum to estimate principal effects.

5.2.3 Bla g1 Reduction in Sensitized Sample

The Inner-City Asthma Study tailored the interventions such that each participant received a remediation plan specific to the allergens he/she was sensitized to (as indicated by a positive skin test). Therefore, we wished to estimate the effect of a 50% reduction in Bla g1 allergen (as measured on the floor) for the subset of individuals that were sensitized to this allergen.⁵ Of the 869 individuals with sensitization information, 599 were sensitized to Bla g1.⁶ This subset is further reduced to 494 when the 105 individuals with missing Bla g1 measurements at either baseline, 1 year, or both were removed. A similar analysis was undertaken to look for differences between the included and ex-

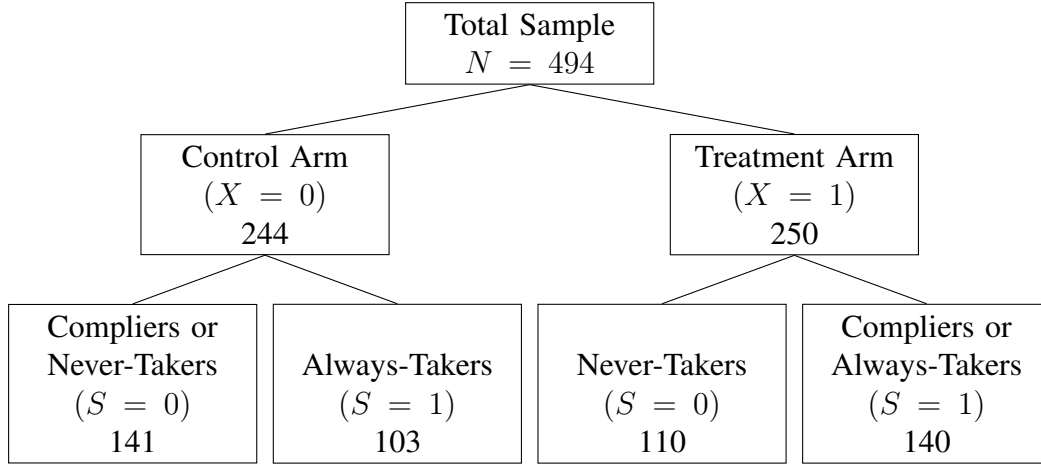
⁵Children sensitized to Bla g1 received professional pest control (Terminix) as part of their personalized intervention.

⁶Only 1 child had a missing value for Bla g1 sensitization. This child was discarded.

cluded individuals. Generally, these two groups appeared similar, however, the included individuals had a slightly higher proportion of males (63% vs. 53%).

Again, we stratified individuals based on their value of S . For this subset, this can be seen in Figure 9.

Figure 9: *Bla g1 Reduction in Sensitized Sample: Responder Stratification*



The individuals in the control group who experienced a significant reduction in Bla g1 are always-takers (103). Individuals in the treatment group who did not see a significant reduction are never-takers (110). The remaining individuals are partially identified. Again, the missing data was imputed and resampled as described in Section 5.2.1. The 244 control individuals and the 250 treatment individuals were used to build models to predict always-taker and never taker status, respectively. The same baseline covariates eligible for inclusion in the first Bla g1 analysis were used here as well.

The formulas and transformation introduced in Section 4 were used to calculate the probability of being a complier, p_i , and used to classify the 281 partially identified individuals. The principal effects were estimated by fitting a linear regression model, predicting change in SFDs, with all of the covariates used to predict stratum membership and a treatment indicator variable. This was done separately in each stratum. These estimates formed the bootstrap distributions seen in Figure 10. The red dashed line represents the mean of the distribution while the blue dashed lines denote the 2.5% and 97.5% percentiles. Table 3 compares these estimates to the overall treatment effect calculated by Morgan et al. [12].

Figure 10: *Bla g1 Reduction in Sensitized Sample: Bootstrap Distribution of Principal Effects*

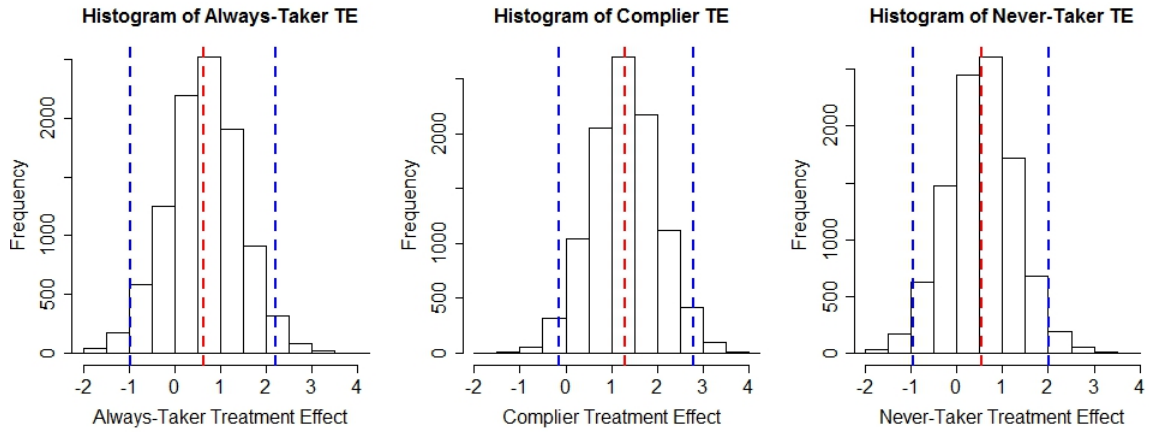


Table 3: *Bla g1 Reduction in Sensitized Sample: Principal Effect Estimates*

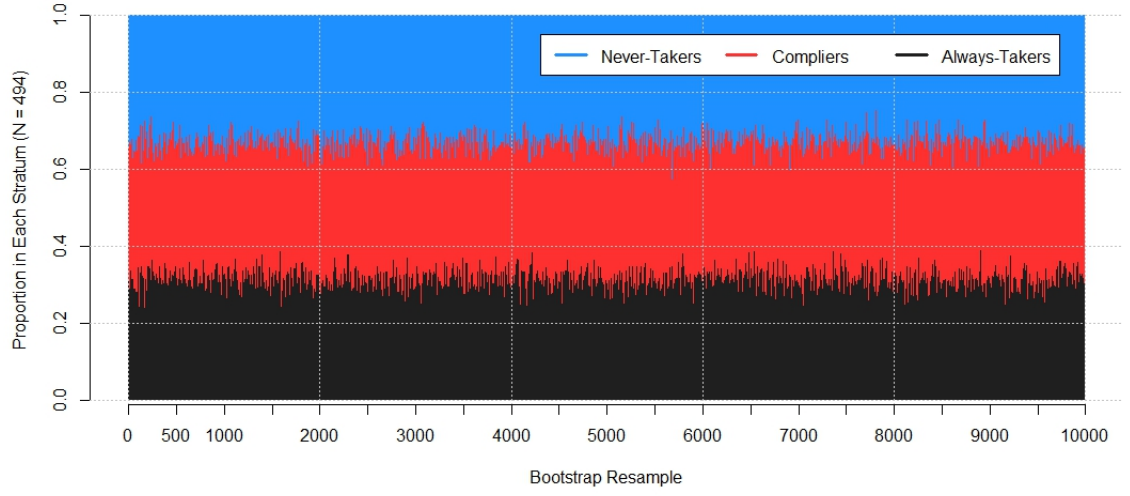
	Always-Takers	Compliers	Never-Takers	Overall
Original Analysis				0.82 _{p<0.001}
Principal Effects Analysis	0.63 _(-0.96,2.20)	1.28 _(-0.14,2.78)	0.54 _(-0.94,2.00)	
Site				
NY/Bos/Chi	1.03 _(-0.73,2.75)	1.51 _(-0.09,3.21)	0.61 _(-1.20,2.51)	
Seattle	-2.13 _(-9.40,5.00)	-0.12 _(-7.14,7.51)	-0.25 _(-5.24,5.77)	
Southwest	-0.17 _(-5.38,4.71)	0.55 _(-3.65,4.60)	-0.43 _(-2.25,3.28)	

The principal effect among compliers is 1.28 SFDs which is higher than the overall estimate found by the original investigators. This seems to be intuitive, since the post-treatment variable that was adjusted for here was targeted by the intervention in all individuals in this subset. Again, heterogeneity of effects across sites was very evident. (Note: Site-specific principal effects were calculated in the same manner as described in Section 5.2.2.)

As seen in the first Bla g1 analysis, it appears the strata stayed approximately the same size in each iteration of the estimation procedure. The width (indicating the proportion in a given sample) seems to be approximately equal for the never-takers, compliers, and always-takers (see Figure 11).

Although this sample was smaller, with only 494 individuals, it would appear there was a sufficient number of individuals (between approximately 100 and 210) in each stratum to estimate principal effects.

Figure 11: *Bla g1 Reduction in Sensitized Sample: Proportion of Individuals per Stratum*



5.2.4 Allergen Exposure Score

In this last analysis, instead of focusing on the reduction of a single allergen, a composite allergen exposure score was calculated. At baseline, this score is the sum of the sensitization indicators multiplied by the floor allergen measurements of the corresponding allergens. The following allergens were considered: dog (Can f1), dust mite (Der f1 and Der p1), cat (Fel d1), and cockroach mix (Bla g1). For the i^{th} individual, this can be expressed as,

$$\text{Score}_{0i} = \sum_{j=1}^5 s_{ij} \cdot m_{0ij}$$

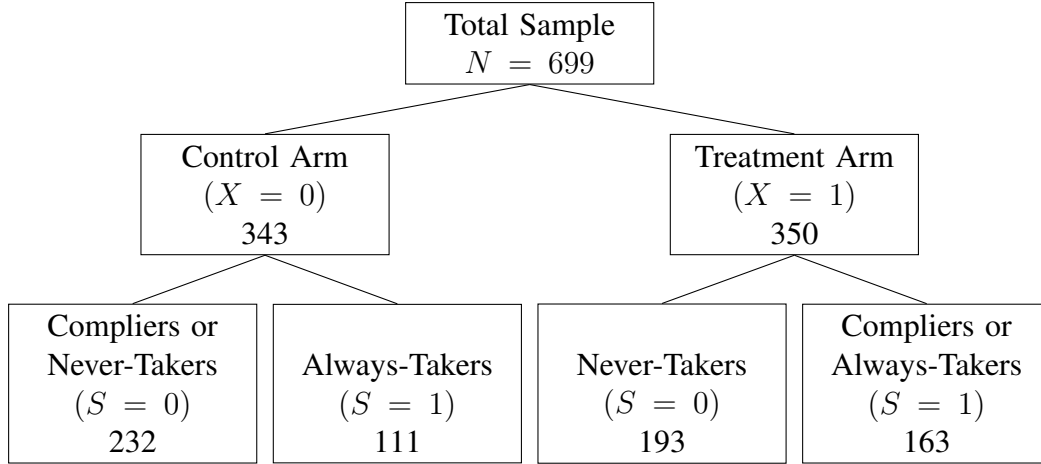
where s_{ij} is the 0/1 indicator of sensitization to the j^{th} allergen and m_{0ij} is the baseline floor measurement of the j^{th} allergen all for the i^{th} individual. At 1 year this score will be,

$$\text{Score}_{1i} = \sum_{j=1}^5 s_{ij} \cdot m_{1ij}$$

A significant reduction, denoted $S = 1$, was considered to be a 50% reduction in score between baseline and 1 year. Again, only included individuals who have an allergen

exposure score at both baseline and 1 year. This excluded 170 (of 839) individuals. An exploratory analysis was done to look for differences between the included and excluded individuals. These individuals looked very similar overall. After stratifying individuals by their value of S , we have the following strata seen in Figure 12,

Figure 12: Allergen Exposure Score: Responder Stratification

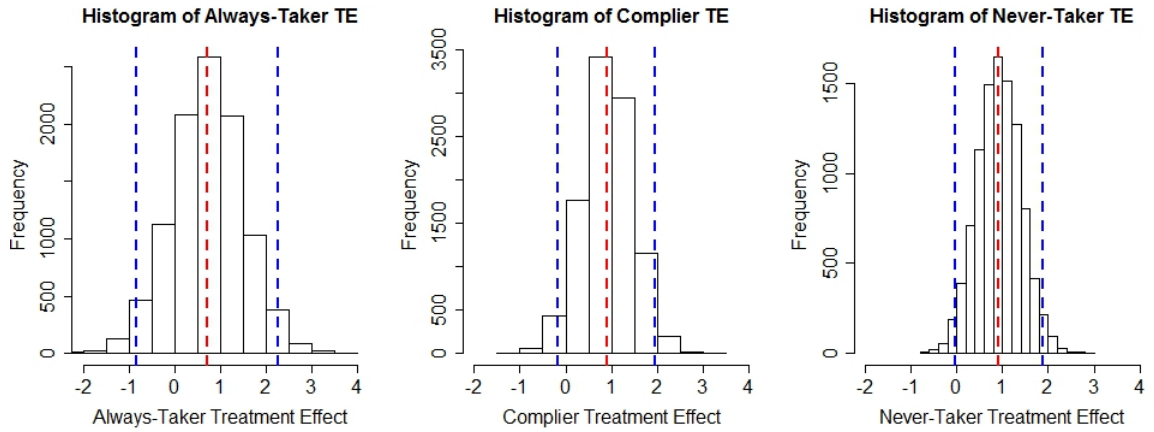


Again, the control group individuals who experienced at least a 50% reduction in allergen exposure score are always-takers (111). Those in the treatment group that did not experience this reduction are never-takers (193). The remaining individuals are partially identified. Missing data was imputed and resampled in the manner described in Section 5.2.1. Models were built to predict always-taker and never-taker status with the control and treatment groups, respectively. The same baseline covariates used in the two previous analyses were used here as well.

The method introduced in Section 4 was used to calculate the probability of being a complier, p_i , and was used to classify the 395 partially identified individuals. After classification using a Bernoulli(p_i) random variable, principal effects were calculated using a linear regression model including all baseline covariates used to predict stratum membership and a treatment assignment indicator. The following bootstrap distribution of principal effects (i.e. the coefficients of the treatment indicator variable in the stratum specific regression models) were estimated and plotted in Figure 13.

Once again, the red dashed line represents the mean of the distribution while the blue dashed lines denote the 2.5% and 97.5% percentiles. Table 4 compares these estimates

Figure 13: Allergen Exposure Score: Bootstrap Distribution of Principal Effects



to the overall treatment effect calculated by Morgan et al. [12].

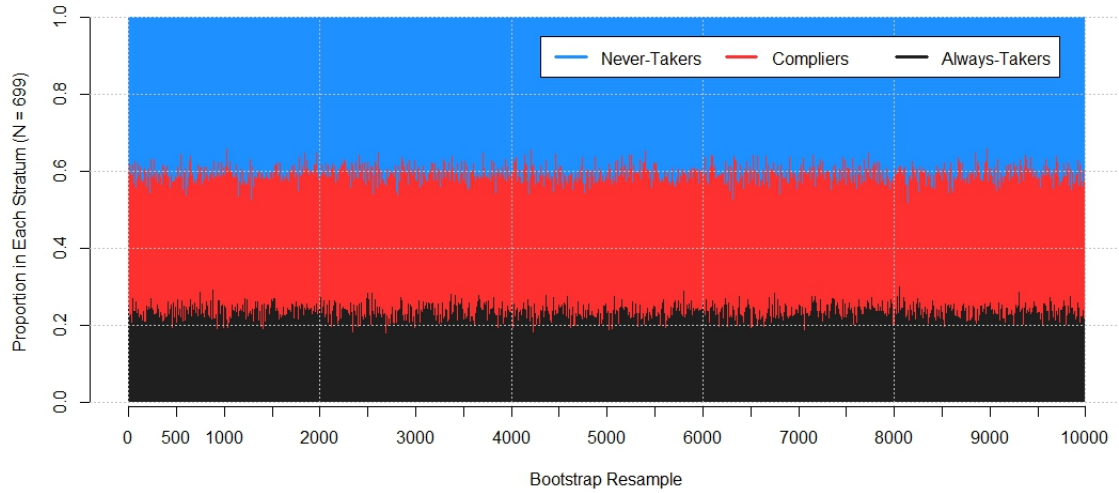
Table 4: Allergen Exposure Score: Principal Effect Estimates

	Always-Takers	Compliers	Never-Takers	Overall
Original Analysis				0.82 _{p<0.001}
Principal Effects Analysis	0.72 _(-0.83,2.25)	0.90 _(-0.17,1.96)	0.92 _(-0.46,2.39)	
Site				
NY/Bos/Chi	1.46 _(-0.68,3.49)	1.20 _(-0.24,2.63)	0.96 _(-0.46,2.39)	
Seattle	-1.31 _(-6.79,3.77)	0.23 _(-3.84,4.15)	0.44 _(-1.98,2.91)	
Southwest	0.08 _(-3.03,3.23)	0.71 _(-0.87,2.26)	1.14 _(-0.25,2.61)	

The principal effect among compliers is estimated to 0.90 SFDs. This estimate attempts to reflect the multifaceted nature of the intervention. However, the small effect among compliers may point to the imperfect collapse of a multidimensional measurement. The intervention acted on the specific allergens a participant was sensitized to, and thus looking for a reduction in a composite measure of those allergens is an appropriate manner to gauge receipt of the intervention. As seen in the previous two analyses, it appears the strata stayed approximately the same size in each iteration of the estimation procedure. The width (indicating the proportion in a given sample) seems to be approximately equal for the never-takers and compliers, while being slightly narrower for the always-takers (see Figure 14).

Such a large proportion of compliers in each resampled dataset would indicate that adequate data was available to calculate the effect in that stratum. Next, the result of the

Figure 14: *Proportion of Individuals per Stratum in Allergen Exposure Score Analysis*



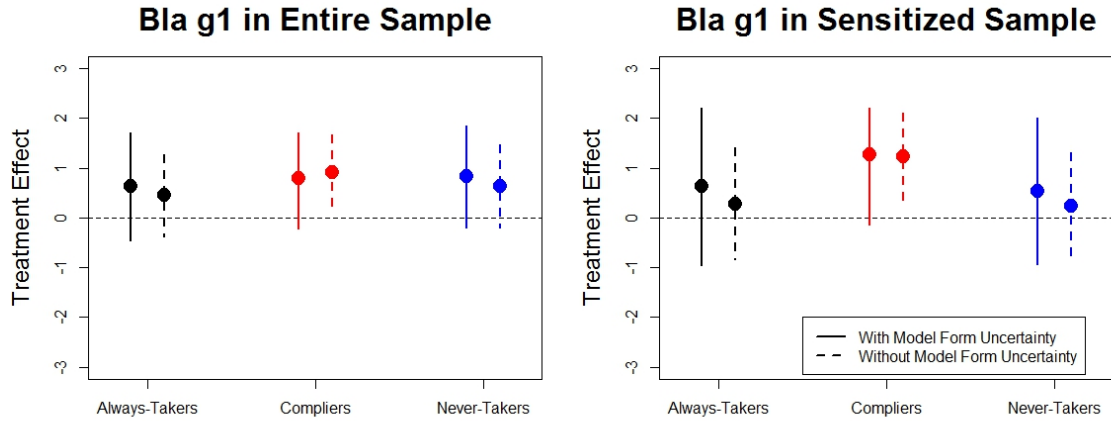
Inner-City Asthma analyses and the PREACH analysis will be discussed in more detail.

5.3 Impact of Model Selection

We found that resampling the data, which allowed the form of the model predicting never-taker and always-taker status to vary, introduced the greatest amount of uncertainty (compared to randomly assigning stratum), especially in the case of the Inner City Asthma Study data. For this data set, the analysis was done in two ways. First, 10,000 complete data sets were imputed and steps 2-5 described in Section 4 were repeated for each data set. This method only incorporated the uncertainty associated with the imputation of the missing values. The second, (presented in Sections 5.2.2- 5.2.3) resampled 200 data sets from 50 imputed data sets and applied steps 2-5 to each one. This method incorporates the uncertainty from the imputation and allows the form of the always-taker and never-taker model to vary from iteration to iteration. For the Blag1 analyses in the entire sample and the sensitized sample, a comparison of these two procedures can be seen in Figure 15.

Both methods result in similar point estimates, however, the first method (seen as the dashed lines in Figure 15) resulted in much narrower confidence intervals. This suggests that if these models were known, much less uncertainty would surround the

Figure 15: *Estimates With and Without Model Form Uncertainty*



estimate of these effects (even in the presence of missing data). Griffin, McCaffrey, and Morral [7] examined the sensitivity of model misspecification in their approach to estimating principal effects and found that their results were extremely sensitive when distributions were very heavy-tailed or extremely skewed. Results were less affected when data was skewed only moderately and tails were only ‘slightly’ heavy. This seems to support the incorporation of additional uncertainty for the form of this model. If we assume this is known but then misspecify the model, the resulting estimate can be very biased and the confidence intervals will be overly optimistic.

5.4 Discussion

In the previous two cases studies our new method for predicting principal strata membership was applied to estimate the benefit of reducing indoor allergens in the homes of asthmatic children residing in the inner-city. In the subset of children for whom allergen/pollutant exposure would be reduced (by at least $20 \mu\text{g}/\text{m}^3$ in the case of $\text{PM}_{2.5}$ or by at least 50% for Bla g1/Allergen Exposure Score), an increase in symptom free days was found which was greater than the overall effect of the intervention.

In conducting both of these case studies several points should be mentioned. First, the selection of a single cutoff point to determine a ‘significant’ decrease in exposure may have affected the outcome estimates in all three strata, especially for the never-

takers. These individuals in particular may have experienced a non-trivial reduction but were nonetheless determined to be never-takers because this reduction was below this threshold. The dichotomization of S allowed for a much simpler and straight-forward analysis. However, complexities may arise in interpreting the principal effects.

The PREACH study focused on a single pollutant, $PM_{2.5}$, and implemented an intervention (air cleaners with HEPA filters) that targeted this pollutant specifically. This made the definition of S much simpler. Due to the multifaceted nature of the Inner City Asthma Study intervention, determining how to define a significant reduction was especially challenging. Morgan et al. [12] found that Bla g1 (cockroach allergen) was reduced in both the treatment and control arms over the course of the study, however, only children that were sensitized to Bla g1 received (or were eligible to receive) a specific module targeted at that allergen. The analyses undertaken in Section 5.2 examine different ways to explore this effect. The first (Section 5.2.2) takes a very broad, public health approach. This analysis estimates the effect of reducing Bla g1 by 50% for all children, both those that are sensitized and those that are not. The second (Section 5.2.3) takes a more targeted approach, looking at the effect of reducing Bla g1 by 50% for the subset of individuals who are sensitized to this allergen.

Both of these approaches have limitations. Very few children enrolled in this data set were sensitized to *only* Bla g1 and thus a child who saw a significant reduction in this particular allergen may still be exposed to considerable quantities of other allergens that trigger asthma-related symptoms. This led to the development of the “allergen exposure score” in Section 5.2.4. This score attempts to collapse a multidimensional intervention into a single number. Looking at the difference in this score between baseline and 1 year attempts to capture the reduction in all allergens that a child is exposed and sensitized to in the home and accurately quantify whether a child has ‘received’ the treatment or not.

In all 3 analyses done with the Inner-City Asthma Study data, the proportion of individuals in each stratum seemed stable across the 10,000 resampled data sets (see Figures 8, 11, 14). Although the number of individuals in each stratum stayed approxi-

mately the same, the individuals in these strata may have changed from one iteration to the next. However, knowing that the principal effects were estimated with a larger group of individuals provides some support for the appropriateness of the resulting estimates. This was less true in the PREACH analysis but may be attributable to the smaller size of the original sample.

Another consideration that must be taken into account is the choice of the lasso regression constraint (applied to the L1 norm of the parameters). When this constraint is small, more of the coefficients (as estimated by the least squares procedure) are shrunk towards 0. When this constraint is large, fewer coefficients are shrunk towards 0, resulting in a larger model. For each of the analyses undertaken constraints of 0.1, 0.5, and 0.8 were used. For the data sets considered, all three constraints produced similar results.

Lastly, each of the case studies presented discarded individuals that did not have exposure levels at baseline and follow-up (6 months for PREACH, 1 year for the Inner City Asthma Study). The resulting treatment and control groups remained balanced, which would imply that the principal effects estimated are still valid. However, it may affect the generalizability of the results as some differences existed between the included and excluded individuals. Sensitivity analyses could be undertaken to assess how these effects would differ in a more representative population.

6 Conclusion

We have introduced a new method for estimating causal effects defined by a post-treatment variable. This method uses baseline covariate data to predict stratum membership for individuals in the population whose stratum is only partially defined. Once all individuals have been assigned to a stratum, we are able to calculate within-stratum (“principal”) effects using simple statistical techniques. We applied this method to data collected in two clinical trials looking at the effectiveness of environmental interventions on lowering asthma morbidity among inner-city children. For the subset

of children who would have seen a significant reduction in exposure as a result of the intervention, we found an increase in symptom free days beyond that of being ‘randomized to the intervention’ (the ITT effect estimated in the original papers). This provides evidence (with the limitations acknowledged earlier) that reducing indoor allergen/pollutant exposure has a causal relationship with improving asthma symptoms in children.

The method developed in Section 4 provides an additional tool for investigators looking to identify the effect of manipulating a single allergen or a specific group of allergens and can possibly provide information on how to best optimize environmental remediation plans in the future. Developing a better understanding of how these factors contribute to the effect of applying an environmental intervention will aid in the development of treatment guidelines and public health policy aimed at improving asthma symptoms among inner-city children.

The estimator developed here allows for the incorporation of uncertainty at each step of the estimation process. It also allows for easy inclusion of a multiple imputation procedure to maximize use of baseline covariates to build the models needed to predict stratum membership. The resulting principal effect distributions represent plausible values for these effects. By taking the mean and 2.5% and 97.5% percentiles we are able to report estimates and confidence intervals that capture the effect of post-treatment variable on the outcome of interest and the associated uncertainty.

7 References

- [1] Joshua D. Angrist, Guido W. Imbens, and Donald B. Rubin. “Identification of causal effects using instrumental variables”. In: *Journal of the American Statistical Association* 91.434 (1996), pp. 444–455.
- [2] Arlene M Butz et al. “A randomized trial of air cleaners and a health coach to improve indoor air quality for inner-city children with asthma and secondhand smoke exposure”. In: *Archives of pediatrics & adolescent medicine* 165.8 (2011), pp. 741–748.
- [3] Robert S Call et al. “Risk factors for asthma in inner city children”. In: *The Journal of pediatrics* 121.6 (1992), pp. 862–866.
- [4] Peyton A Eggleston et al. “Home environmental intervention in inner-city asthma: a randomized controlled clinical trial”. In: *Annals of Allergy, Asthma & Immunology* 95.6 (2005), pp. 518–524.
- [5] “Expert Panel Report 3 (EPR-3): Guidelines for the Diagnosis and Management of Asthma-Summary Report 2007.” In: *The Journal of allergy and clinical immunology* 120.5 Suppl (2007), S94.
- [6] Constantine E Frangakis and Donald B. Rubin. “Principal stratification in causal inference”. In: *Biometrics* 58.1 (2002), pp. 21–29.
- [7] Beth Ann Griffin, Daniel F McCaffrey, and Andrew R Morral. “An application of principal stratification to control for institutionalization at follow-up in studies of substance abuse treatment programs”. In: *The annals of applied statistics* 2.3 (2008), p. 1034.
- [8] Martijn W Heymans et al. “Variable selection under multiple imputation using the bootstrap in a prognostic study”. In: *BMC Medical Research Methodology* 7.1 (2007), p. 33.
- [9] Paul W Holland. “Statistics and causal inference”. In: *Journal of the American Statistical Association* 81.396 (1986), pp. 945–960.

- [10] Guido W. Imbens and Donald B. Rubin. “Bayesian inference for causal effects in randomized experiments with noncompliance”. In: *The Annals of Statistics* (1997), pp. 305–327.
- [11] Marshall M. Joffe, Dylan Small, and Chi-Yuan Hsu. “Defining and estimating intervention effects for groups that will develop an auxiliary outcome”. In: *Journal of the Royal Statistical Society* 178 (2007), pp. 425–443.
- [12] Wayne J Morgan et al. “Results of a home-based environmental intervention among urban children with asthma”. In: *New England Journal of Medicine* 351.11 (2004), pp. 1068–1080.
- [13] KM Mortimer et al. “The effect of air pollution on inner-city children with asthma”. In: *European respiratory journal* 19.4 (2002), pp. 699–705.
- [14] Roger D. Peng et al. “Estimating the health benefit of reducing indoor air pollution in a randomized environmental intervention”. In: *Journal of the Royal Statistical Society* 178 (2015), pp. 425–443.
- [15] Paul R. Rosenbaum. “The consequences of adjustment for a concomitant variable that has been affected by the treatment”. In: *Journal of the Royal Statistical Society* (1984), pp. 656–666.
- [16] David L Rosenstreich et al. “The role of cockroach allergy and exposure to cockroach allergen in causing morbidity among inner-city children with asthma”. In: *New England Journal of Medicine* 336.19 (1997), pp. 1356–1363.
- [17] Donald B. Rubin. “Causal inference using potential outcomes”. In: *Journal of the American Statistical Association* 100.469 (2005).
- [18] David P Strachan and Derek G Cook. “Parental smoking and childhood asthma: longitudinal and case-control studies”. In: *Thorax* 53.3 (1998), pp. 204–212.
- [19] Matthew J Strickland et al. “Short-term associations between ambient air pollutants and pediatric asthma emergency department visits”. In: *American journal of respiratory and critical care medicine* 182.3 (2010), pp. 307–316.

- [20] Robert Tibshirani. “Regression shrinkage and selection via the lasso”. In: *Journal of the Royal Statistical Society. Series B (Methodological)* (1996), pp. 267–288.
- [21] Onchee Yu et al. “Effects of ambient air pollution on symptoms of asthma in Seattle-area children enrolled in the CAMP study.” In: *Environmental Health Perspectives* 108.12 (2000), p. 1209.

8 Curriculum Vitae

Katherine Freeland

EDUCATION

The Johns Hopkins University, Baltimore, MD

Sc.M. in Biostatistics, May 2015

Thesis: “Predicting Principal Strata in Environmental Interventions”

Advisor: Roger Peng, PhD

The University of New Mexico, Albuquerque, NM

B.S. in Statistics, minor in Spanish, May 2013

B.A. in Dance, concentration in Flamenco, May 2013

Thesis: “Exploring Cancer Reoccurrence and Mortality Trends with the SEER Data Set”

Advisor: Michael Sonksen, PhD

PROFESSIONAL EXPERIENCE

Technical Intern, May 2012 - present

Sandia National Laboratories, Albuquerque, NM

Statistics and Surveillance Assessment (May 2013 - present)

International Biological Threat Reduction (May 2012 - May 2013)

Research Assistant, September 2014 - present

Johns Hopkins School of Medicine/Bloomberg School of Public Health, Baltimore, MD

Department of Surgery (September 2014 - present)

Center for Adolescent Health (September 2014 - December 2014)

TEACHING/TUTORING EXPERIENCE

Teaching Assistant, September 2014 - present

Johns Hopkins Bloomberg School of Public Health, Baltimore, MD

Biostatistics 664: Causal Inference in Medicine and Public Health (Spring 2015)

Biostatistics 621/622: Statistical Methods in Public Health I & II (Fall 2014)

POSTERS & PRESENTATIONS

“Predicting Principal Stratum Membership in Environmental Interventions” (poster)

ENAR 2015 Spring Meeting: March 15, 2015 (poster)

“A Margin Based Approach to Determining Sample Size Via Tolerance Bounds”

ASA’s Conference on Statistical Practices: February 21, 2014 (poster)

“Exploring Cancer Reoccurrence and Mortality Trends with the SEER Data Set”

UNM Dept. of Mathematics and Statistics Mini Conference: November 30, 2012