

# Testing for Placebo Effects using Data from Medical Trials

Anup Malani\*

October 7, 2003

## Abstract

This paper criticizes the methodology of existing studies that purport to find evidence of placebo effects in medical trials. Because these studies fail to define and properly model placebo effects, it is hard to interpret and determine the strength of their results. The paper addresses this problem by proposing a model of health outcomes that formalizes the dominant medical theory of how placebo effects operate. This theory posits that health outcomes rise in individuals' expectations about their beliefs about the probability that they are getting a beneficial treatment and their beliefs about the efficacy of that treatment. The paper then specifies the conditions under which placebo effects change outcomes in randomized, double-blind, parallel-arm, controlled trials. A blinded trial is a perfect environment to test for placebo effects because it offers an objective and controlled manipulation of subjects' beliefs about the probability and efficacy of treatment. A simple yet accurate method of testing for the existence of placebo effects is to check whether outcomes in trials where a higher fraction of subjects are randomized into active treatment are superior to outcomes in trials with a lower fraction given active treatment. The paper applies this test to data from 150 trials of anti-ulcer medications and finds robust evidence of placebo effects in trials of H<sub>2</sub>-blockers (e.g., Zantac, Tagamet and Pepcid) and of proton-pump inhibitors (e.g., Prilosec, Nexium, and Prevacid).

## 1 Introduction

Placebo effects can roughly be defined as that component of health outcomes that cannot be attributed to the physiological effects of treatment or to the natural progression of disease. There is a lively debate in the medical literature about whether placebo effects actually exist. On one side of the debate there is, e.g., a recent *New England Journal of Medicine* article by Hrobjartsson and Gotzsche [1] that examines 114 studies with both a blinded placebo-control group and an unblinded no-treatment group. The authors find few systematic differences in outcomes between these groups across their sample. Although widely publicized, this result does not disprove the existence of placebo effects. It is consistent with the plausible theory that members of unblinded no-treatment groups seek out alternative medication, which elevates their health outcomes.

On the other side of the debate are, e.g., studies by Kirsch and Sapirstein [2] and Kirsch et al. [3] that point to evidence that members of the placebo-control group in a given

---

\*Associate Professor, University of Virginia Law School. I would like to thank Scott Baker, Gary Becker, Albert Choi, Feifang Hu, Darius Lakdawalla, Ron Michener, Willard Manning, Charles Mullin, Tomas Philipson, Andrew Sellgren and participants in the Public Workshop at the University of Virginia Department of Economics for helpful comments. I am grateful to David Greene for excellent research assistance. All errors are my own.

double-blinded trial manifested substantially improved health outcomes. These findings are weak support for placebo effects because the improvements could be due to the natural progression of disease [4]. Better studies employ a balanced-placebo design wherein subjects are first randomized across treatments and then across instructions about treatment, with one group given each treatment being told they were administered active treatment and the other being told they were administered placebo.<sup>1</sup> Such studies generally find evidence in support of placebo effects [7], although in more recent studies the results are mixed [8]. More importantly, these studies are ethically questionable and perhaps even illegal.

A common weakness of studies on both sides of the debate is that they do not begin with a formal model of placebo effects that can be clearly falsified. It is unclear, therefore, how powerful their evidence on the existence of placebo effects actually is. This paper addresses this shortcoming by proposing a formal model of the dominant medical theory [9] for how placebo effects operate. This so-called expectancy theory posits that the more optimistic a patient is about the efficacy of a treatment, the more positive will be her health response to that treatment. Moreover, if the patient is told she is being administered a treatment she thinks will prove helpful, she will manifest an improved health outcome even if she is never in fact given treatment [10, 11]. This theory is formalized by assuming that health outcomes are a function not just of the treatment, but also the expected value of treatment in the eyes of the patient.

The paper proposes to test this theory in the context of medical trials, which, if blinded, provide an objective and controlled manipulation of beliefs that permits a relatively clean test of the effect of beliefs about treatment on health responses to treatment. The focus is on parallel-group, randomized, controlled trials (RCT). These are modeled as lotteries over a treatment and control. Although informed consent procedures reveal the ex ante probability of obtaining treatment, blinding ensures that subjects do not learn their ultimate assignment. This probability, along with subjects' assessments of the relative efficacy of the treatment, affects subjects' beliefs about the expected value of the trial. If these beliefs affect outcomes, the probability of obtaining treatment will affect trial outcomes.

More specifically, if trial subjects believe that the treatment is superior to the control, the models predict that subjects in the treatment (control) arm of trials with a higher share treated would manifest better health outcomes than subjects in the treatment (control) arm of trials with a lower share treated. If the antecedent condition holds, this prediction provides a clean test of the existence of placebo effects as defined by the expectancy theory. Fortunately, the antecedent condition holds. Because enrollment in trials is voluntary, only those who are more optimistic about treatment than control enroll.

The paper applies this test for placebo effects to data from over 150 RCTs of anti-ulcer medications. The advantage of ulcer trials is that outcomes are objectively measured: ulcer healing is verified by endoscopy. In trials where patients were asked for informed consent and thus had some indication of their odds of obtaining active treatment, a positive correlation is found between the share treated and outcomes in each arm of trials of H<sub>2</sub>-blockers (e.g., Zantac, Tagamet, and Pepcid) and of proton-pump inhibitors (e.g., Prilosec, Nexium, and Prevacid) controlling for available group-level clinical covariates and study-level design covariates. This positive correlation is significant and robust to covariate specification. In trials without informed consent, this correlation vanishes. This result provides strong evidence for the existence of placebo effects.

---

<sup>1</sup>Moreover, Penick & Hinckle [5] and Penick & Fisher [6] have performed related experiments that randomize across treatment and instruction about treatment efficacy (as opposed to about treatment state). These trials yielded mixed results and, like the balanced-placebo design, are ethically questionable.

Section 2 presents a model of treatment strategies, one of which is the clinical trial. Section 3 formalizes the expectancy theory of placebo effects. Section 4 derives testable predictions regarding trial outcomes with and without placebo effects. Section 5 tests these predictions against data from ulcer trials.

## 2 Treatment Strategies

This section presents a model of treatment strategies for individuals who are currently ill. There are two possible, future health states: continued illness  $\bar{y}$  and recovery  $y$ , where  $\bar{y} > y$ .<sup>2</sup> Treatments are lotteries over these two states. For now, assume there exist only two treatments, indexed by subscript  $k$ : no treatment ( $k = 0$ ) and an experimental treatment ( $k = 1$ ). Let  $y_{ki}$  be the random variable that describes individual  $i$ 's health outcome given treatment  $k$ . Define  $p_{ki} = \Pr \{y_{ki} = \bar{y} | \text{no placebo effects}\}$ . In the case of no treatment, this probability is simply a function of the natural progression of an individual's ailment. For the experimental treatment, this probability is also a function of the physiological effects of the experimental treatment. Although a slight abuse of medical terminology, the sum of natural progression and the physiological effects of a given treatment will be referred to as the specific effects of that treatment.<sup>3</sup>

Treatments are to be distinguished from treatment strategies. The latter are indexed by  $s$  and defined to be lotteries over treatments. Because treatments are themselves lotteries over health states, treatment strategies are really compound lotteries over health states. This paper focuses on individuals for whom there are only two feasible treatment strategies: certain consumption of no treatment ( $s = 0$ ) or enrollment in a randomized, placebo-controlled trial (RPCT) that is blinded ( $s = BT$ ). The latter strategy entails a probability  $d$  of receiving the experimental treatment and  $(1-d)$  of receiving no treatment. This narrow set of feasible treatment strategies is appropriate under two conditions. First, individuals cannot consume the experimental treatment for sure because, e.g, the government has not approved the experimental treatment for widespread prescription by doctors.<sup>4</sup> In this case, the only way to obtain the experimental treatment is by enrolling in a trial.<sup>5</sup> Second, individuals are only offered one lottery that includes the experimental treatment and this lottery is blinded, i.e., subjects do not learn which treatment they actually consume. That individuals are offered only one lottery is a common constraint because medical trials are costly to operate. Moreover, investigators prefer blinded trials to unblinded trials because, *inter alia*, the former are less vulnerable to subject attrition from the control group.

Blinded RPCTs are conducted as follows. First, individuals are recruited. Enrollment is voluntary and subject to informed consent. As part of this disclosure, subjects are assumed to be given information about the probability that they will receive the experimental treatment.<sup>6</sup> Individuals who choose to participate are called enrollees or subjects. Second,

---

<sup>2</sup>The analysis can easily be extended to the case of continuous outcome variables.

<sup>3</sup>Traditionally the specific effect of treatment excludes the natural progression of disease, which, along with placebo effects, are called the non-specific effects of treatment.

<sup>4</sup>This also implies that subjects cannot create their own lotteries over no treatment and the experimental treatment.

<sup>5</sup>Many of the studies in the data set of anti-ulcer studies in Section 5 take place after the specific anti-ulcer medication investigated has been approved in the country of the study. This does not imply that the data set does not satisfy the assumption in the text. The reason is that, while certain medication-dosage combinations have been approved in the country of a candidate study, the medication-dosage combination investigated in the candidate study may not have been approved.

<sup>6</sup>This assumption is reasonable. U.S. law requires informed consent before enrollment in experiments,

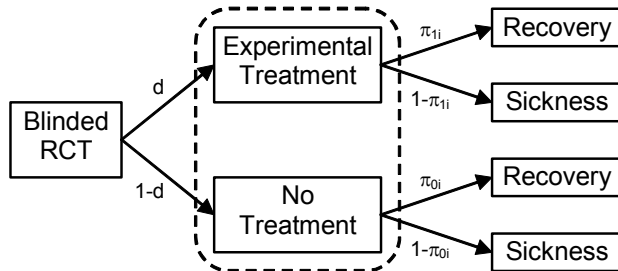


Figure 1: Individual  $i$  views a blinded trial as a compound lottery.

subjects are randomized into the experimental-treatment or a no-treatment control group. Whereas subjects in unblinded trials are told, after randomization, the group to which they are assigned, in blinded trials they are not. Third, subjects in the experimental-treatment group are administered the experimental treatment and subjects in the control group a placebo. In order to focus the analysis on selection *into* medical trials, it is assumed that there is no unblinding in blinded trials. Subjects do not discover their group assignment through, e.g., subject sampling [12] on outcomes or side effects. This controls attrition from blinded trials.

Individual  $i$ 's belief about the probability that she is consuming treatment  $k$  is  $\delta_{ki}$ , where  $\sum_{k=0}^1 \delta_{ki} = 1$ . Her belief about the specific effect of treatment  $k$  is  $\pi_{ki}$ . From the individual's perspective, strategies are defined by the vector of beliefs  $(\delta_{0i}, \delta_{1i})$  about treatment consumption. Outside the context of a trial, the only feasible strategy is consumption of no treatment. Because the individual knows her treatment for sure,  $\delta_{0i} = 1$ . In the context of a blinded trial, the individual knows she will obtain either the experimental treatment or no treatment, but does not know which. Because of informed consent, the individual knows she will receive the experimental treatment with probability  $d$ . Given full compliance, the individual will set her beliefs such that  $\delta_{0i} = (1 - d)$  and  $\delta_{1i} = d$ . Thus, the individual views the no treatment and blinded RPCT strategies as the treatment probability vectors  $(1, 0)$  and  $(1 - d, d)$ , respectively. Figure 1 presents the compound lottery that is a blinded RPCT from the perspective of individual  $i$ .

Individuals are assumed to have preferences that conform to Savage's axioms and thus permit representation in the form of a subjective expected utility function. Individuals draw utility from health and other items, including wealth. For simplicity, it is further assumed that individuals have identical utility functions and that these functions are additively separable in health and other items. Let  $u(y)$  be the utility from health outcome  $y$ , with  $u' > 0$ ,  $u'' < 0$ . The expected utility of strategy  $s$  to individual  $i$  is a weighted sum of the utility from each health outcome, with the weights being her subjective beliefs about the probability of each outcome given the compound lottery  $s$ :  $U_i^s = \pi_i^s u(\bar{y}) + (1 - \pi_i^s) u(\underline{y})$ . Given that an individual knows her treatment status outside the context of a trial, her subjective belief is  $\pi_i^0 = \pi_{0i}$ . Belief about the probability of recovery given the strategy of enrolling in a blinded trial depends on the probability of being given the experimental treatment:  $\pi_i^{BT} = d\pi_{1i} + (1 - d)\pi_{0i}$ .

---

see 21 C.F.R. § 7.3(f) (trials of drugs and medical devices); 45 C.F.R. § 46.101(a) (trials of new medical procedures), and requires that informed consent reveal the probability of receiving the experimental treatment, see, e.g., 21 C.F.R. § 20.25.

In order to determine the sorting of individuals to strategies, one must know the distribution of beliefs about the efficacy of each treatment among the population. Since the object of these beliefs is the actual efficacy of each treatment, let  $\mathbf{g}_p$  give the probability distribution function of  $\mathbf{p}_i = (p_{0i}, p_{1i})$  across the population. Let  $\mathbf{g}_\pi$  give the probability distribution for  $\boldsymbol{\pi}_i = (\pi_{0i}, \pi_{1i})$ . All distributions discussed are assumed to be well-defined. The expectations operator  $E_g(\cdot)$  will be employed when expectations are taken over the joint distribution of  $(\mathbf{p}_i, \boldsymbol{\pi}_i)$ .

### 3 Expectancy Theory of Placebo Effects

According to the expectancy theory of placebo effects, patients manifest changed health outcomes in response to expectations regarding treatment. In particular, the more effective a patient expects a treatment to be, the better her response to it. Moreover, the more likely a patient thinks she is to get a beneficial treatment the better is her health outcome holding constant whether or not she receives treatment.<sup>7</sup>

While a significant number of studies claim to find evidence of these responses, the studies in Table 1 are among the better representatives of this class. These studies, and others like them, provide important insights into the nature of placebo effects. They do not, however, provide compelling evidence to support the existence of placebo effects. First, the studies have small sample sizes, as measured by the number of trials in which placebo effects are observed. Second, many of the trials examine subjective rather than objective criteria. For example, the studies cited by Skovlund and the Pollo et al. study relied upon patient self-reports of pain levels. In contrast, this paper tests for the existence of placebo effects using a large number of trials with an objective measure of outcomes.

#### 3.1 Definition of placebo effects

This section formalizes the expectancy theory of placebo effects, first, with a definition of these effects and, second, with a model of health outcomes that complies with the definition and thus can be said to incorporate placebo effects.

**Definition 1** *Individual  $i$  experiences a positive placebo effect if  $\pi_i^s > q_{ki}$  and  $E[y_{ki}] > \pi_{ki}$ , where expectations are taken over individual  $i$ 's outcomes. She experiences a negative placebo effect if  $\pi_i^s < q_{ki}$  and  $E[y_{ki}] < \pi_{ki}$ .*

Treatment  $k$  is said to be associated with a positive placebo effect if two conditions hold. First, conditional on information the individual has about the probability of consuming each treatment and the efficacy of these treatments, the individual's expectation  $\pi_i^s$  about the probability of recovery given strategy  $s$ , including only the specific effects of that strategy, are greater than some arbitrary cut-off  $q_{ki}$  in state  $k$ . Second, the individual experiences a

---

<sup>7</sup>There are two alternative theories of placebo effects that have received substantial attention in the medical literature. The first – the so-called conditioning theory – suggests that, if the body has previously experienced a specific health response after consumption of a medication, consumption of a substance close in appearance or smell to that medication will trigger a similar non-specific health response [13, 14]. The reaction is at the sub-conscious level. No well-formed beliefs about efficacy are involved. However, this theory has been criticized as merely a specific case of the expectancy theory where beliefs are formed based on prior experience and manifested only at the subconscious level. See **Rescorla, 1988**. The second theory – the so-called motivation theory – posits that individuals with a stronger desire to respond to treatment experience placebo effects when treated [9].

probability of recovery after consumption of treatment that is greater than the probability of recovery given only the specific effects of that treatment. A negative placebo (or the so-called “nocebo”) effect is defined to exist under opposite conditions.

Note that the existence of placebo effects depends on beliefs about only the specific effects of a treatment strategy. It does not depend on beliefs about placebo effects. If it were otherwise, placebo effects would exist only if people believed they existed. If beliefs about placebo effects were endogenous, there would be no way to predict which people would believe in placebo effects and thus experienced placebo effects. Indeed, if individuals had full control over beliefs, it would be in each individual’s interest to believe not just in placebo effects, but in fully curative placebo effects. One does not, however, observe such beliefs in the real world, let alone complete recovery by each individual due to powerful placebo effects.<sup>8</sup> Placebo effects could be defined to depend on beliefs about placebo effects if such beliefs were exogenous. This is not incompatible with the definition of placebo effects above. Surely beliefs about the specific effects of treatment are not immaterial to whether subjects experience placebo effects. In that case, the cutoff  $q_{ki}$ , which can vary across individuals, can be defined to include beliefs about the placebo effects given treatment  $k$ .

The purpose of requiring expectations to exceed a certain tipping point  $q_{ki}$  before placebo effects are said to exist is to permit the existence of nocebo effects and to relate such effects in a simple manner to positive placebo effects. Without a tipping point, it is unclear how one would define nocebo effects to exist. An obvious candidate for the cut-off is the actual specific effects of treatment  $k$ ,  $p_{ki}$ . It seems reasonable to suppose that a positive placebo effect exists if a patient expects more from a treatment strategy than the specific effects of the treatment  $k$  that she consumes ( $\pi_i^s > p_{ki}$ ) and experiences a lottery over health outcomes given consumption of treatment  $k$  that is better than that implied by solely the specific effects of that treatment ( $E[y_{ki}] > p_{ki}$ ).<sup>9</sup>

### 3.2 Model of health outcomes

Suppose that the probability of recovery by individual  $i$  given treatment  $k$  is

$$\Pr \{y_{ki} = \bar{y} | a, p_{ki}, \pi_i^s\} = p_{ki} + a (\pi_i^s - p_{ki}), \quad (1)$$

where  $a \in [0, 1]$ , for all  $s, k$ . This parameterization posits that health outcomes given treatment  $k$  are the sum of the specific effects of treatment  $k$  plus a placebo effect driven by expectations given strategy  $s$  and weighted by the parameter  $a$ . Consistent with the definition of placebo effects under the expectancy theory, there is a positive or negative placebo effect if the individual’s beliefs given her treatment strategy are greater or less, respectively, than the specific effects of her actual treatment.

---

<sup>8</sup>Berthelot et al. [15], e.g., interviewed 300 rheumatology patients and found that that three-quarters did not know about placebo effects.

<sup>9</sup>The difficulty with this cut-off is that  $p_{ki}$  may not be known to the patient. Moreover, there is no proven psycho-physiological reason for why positive placebo effects would only exist when beliefs are greater than specific effects. Nevertheless, the next section presents a model of health outcomes that uses the specific effect of a treatment as the cut-off between positive and negative placebo effects for two reasons. First, the test for the existence of placebo effects recommended in this paper is valid even with an arbitrary cutoff, so long as the correlation between the cutoff and beliefs is not too negative [16]. Second, using specific effects as cutoffs permits modelling observable health outcomes as a weighted sum of the actual and believed specific effects of treatment. This is simple and intuitive assumption about the relationship between treatment, beliefs, and outcomes.

The simple linear formulation of expectancy theory in (1) implies that health outcomes are a weighted average of the specific effects of a treatment and beliefs about treatment strategy:  $\Pr \{y_{ki} = \bar{y} | a, p_{ki}, \pi_i^s\} = (1 - a) p_{ki} + a \pi_i^s$ . The parameter  $a$  indicates the relative importance of beliefs in determining the health outcome given consumption of treatment  $k$ . If  $a = 0$ , beliefs have no effects on outcomes. If  $a = 1$ , outcomes are completely determined by beliefs.<sup>10</sup> The influence of placebo effects is assumed to be invariant across treatment states and across individuals. The former assumption permits that the influence of beliefs may vary across ailments, but requires that, for a given ailment, the influence of beliefs does not vary across treatments. The purpose of the latter assumption is to facilitate application of the test for the existence of placebo effects in section 4 to data from ulcer trials. Those data are aggregated to the level of treatment groups so estimation of a random effects model is not feasible.<sup>11</sup>

## 4 Tests for the Existence of Placebo Effects

This section employs the models of treatment strategy and health outcomes to generate predictions regarding outcomes observed in clinical trials with and without placebo effects. Any difference in predictions can be used to test for the existence of placebo effects.

Assume, for the moment, that all individuals who believe that the experimental treatment is superior to no treatment ( $\pi_{1i} > \pi_{0i}$ ) are sorted into the trial. The remaining are sorted out of the trial. The next section demonstrates that this assumption is a natural result of the process of individual self-selection into blinded RPCTs. Define  $\tilde{\pi}_{ki} = \pi_{ki} - \pi_{0i}$ , the relative benefit of treatment  $k$  over no treatment, so that  $\pi_{1i} > \pi_{0i}$  can conveniently be written  $\tilde{\pi}_{1i} > 0$ .

In the absence of placebo effects, the mean outcome observed in the group that receives treatment  $k$  in a blinded RPCT is

$$E_g [y_k | \tilde{\pi}_1 > 0] = E_g [p_k | \tilde{\pi}_1 > 0].$$

(Henceforth,  $i$  subscripts are dropped to simplify notation whenever their use does not add to the exposition.) In the presence of placebo effects, the mean outcome is

$$E_g [y_k | \tilde{\pi}_1 > 0, d] = (1 - a) E_g [p_k | \tilde{\pi}_1 > 0] + a \{ d E_g [\pi_1 | \tilde{\pi}_1 > 0] + (1 - d) E_g [\pi_0 | \tilde{\pi}_1 > 0] \}.$$

Without placebo effects, outcomes are solely a function of the specific effects of treatment. With placebo effects, outcomes are a weighted average of specific effects of treatment  $k$  and beliefs about specific effects. In a blinded RPCT, the relevant beliefs are those about the

---

<sup>10</sup>It is assumed that  $a_k < 0$  is not possible. Such a parameterization would be consistent with a model of health outcomes where individuals manifest nocebo effects if a drug does not live up to expectations. There does not appear to be any direct empirical support for this sort of regret-based model of outcomes in the literature on placebo effects. However, the literature on the motivation theory of placebo effects may ultimately shed light on this issue.

<sup>11</sup>One implication of this assumption is that the influence of placebo effects is independent of individuals' beliefs about the efficacy of different treatment states (though not of treatment state per se). There seems to be no medical basis for thinking that individuals who are more or less responsive to treatment, who are more or less likely to recover naturally from an ailment, or who are more or less optimistic about the efficacy of treatment, manifest outcomes that are more or less driven by placebo effects. Nor is there any obvious, logical reason why individuals whose outcomes are more or less driven by expectations would believe that treatment is—from a physiological perspective—more or less effective.

specific effects not of treatment  $k$  but of the trial. The reason is that subjects are blinded and therefore do not know that they consume treatment  $k$ . Subjects know only that there is a probability  $d$  of obtaining the experimental treatment. This implies an expected value of  $d\pi_{1i} + (1 - d)\pi_{0i}$  for the trial for individual  $i$ .

In the absence of placebo effects, outcomes do not depend on the share treated. In the presence of placebo effects, because beliefs about the trial depend on the share treated, outcomes depend on the share treated. More specifically, given the assumed sorting of individuals into trials, an increase in  $d$  lifts mean outcomes:

$$\frac{\partial E_g [y_k | \tilde{\pi}_1 > 0, d]}{\partial d} = a E_g [\tilde{\pi}_1 | \tilde{\pi}_1 > 0] > 0.$$

As the share treated rises, individuals in the trial become more optimistic because there is a better chance of getting the experimental treatment, which is thought to be better than no treatment. This optimism translates into better average outcomes when there exist placebo effects. This yields the following test for placebo effects:

**Proposition 2** *Suppose all and only individuals for whom  $\tilde{\pi}_{1i} > 0$  enroll in a blinded RPCT. If trials that have higher treatment shares but are otherwise identical yield higher mean outcomes conditional on treatment, then there exist placebo effects as defined in section 3.*

This is the central theoretical result of this paper. It is robust to the functional form of the relationship between observed outcomes, specific effects of treatment, and beliefs about specific effects.

#### 4.1 Self-selection

This section justifies the assumption in Proposition 2 that only and all individuals who believe that the specific effects of the experimental treatment are superior to those of no treatment enroll in a blinded RPCT. This sorting is a direct implication of individual self-selection into the type of trial described in section 2. In that context, expected utility maximization implies that individual  $i$  will enroll in a trial if and only if the probability of recovery given enrollment in a blinded RCT is greater than the probability of recovery given no treatment. If subjects do not take placebo effects into account when deciding whether to enroll, this condition can be written  $\pi_i^{BT} = d\pi_{1i} + (1 - d)\pi_{0i} > \pi_{0i}$  and is satisfied so long as  $\pi_{1i} > \pi_{0i} \Leftrightarrow \tilde{\pi}_{1i} > 0$ . Therefore, the trial will attract only and all individuals who believe the specific effects of the experimental treatment are superior to no treatment, regardless of the share treated. This yields the following result.

**Proposition 3** *In the model of treatment strategies in section 2, self-selection implies that individuals will enroll in a blinded RPCT if and only if  $\tilde{\pi}_{1i} > 0$ . Therefore, the test for placebo effects in Proposition 2 is valid for this model.*

The test generally remains valid even if subjects take placebo effects into account when deciding whether to enroll in trials.<sup>12</sup> In this case, the individual assesses the value of the blinded trial to be  $\pi_i^{BT} = d[\alpha_i\pi_{1i} + (1 - \alpha_i)\{d\pi_{1i} + (1 - d)\pi_{0i}\}] + (1 - d)[\alpha_i\pi_{0i} + (1 - \alpha_i)\{d\pi_{1i} + (1 - d)\pi_{0i}\}]$ , where  $\alpha_i$  is the individual's belief regarding the influence a

<sup>12</sup>Malani [16] justifies the assumption of an expected utility representation for preferences in this case.

of beliefs on outcomes. Implicit in this formulation is the assumption that beliefs about the influence of beliefs are independent of beliefs about the specific effects of treatments. A little algebra reveals that the condition  $\pi_i^{BT} > \pi_{0i}$  again collapses to  $\tilde{\pi}_{1i} > 0$ . The reason is that blinding ensures that placebo effects have the same impact whether one is randomized into the experimental or no treatment group. That leaves the specific effects of treatment as the only basis upon which to choose between foregoing treatment altogether and enrollment in a blinded RPCT, which is a lottery over experimental treatment and no treatment. A critical assumption behind the conclusion that an individual's knowledge of placebo effects does not change the sorting of individuals to trials is that the influence of placebo effects is identical in regardless of treatment consumed. If this is not the case, proposition 3 is invalid.

Proposition 3 is not valid if investigators offer individuals incentives such as cash or in-kind benefits to participate in a blinded RPCT. Such payments may induce individuals who believe that no treatment is superior to experimental treatment, i.e.,  $\tilde{\pi}_{1i} < 0$ , to enroll. In this case, one effect of an increase in the share treated is to lower expectations and thus outcomes of individuals who believe that  $\tilde{\pi}_{1i} < 0$  yet enroll in a blinded RPCT:  $\partial E_g [y_k | \tilde{\pi}_1 < 0, d] / \partial d = a E_g [\tilde{\pi}_1 | \tilde{\pi}_1 < 0] < 0$ . The overall effect of an increase in the share treated will be ambiguous.<sup>13</sup> This is a more serious problem where the experimental treatment is available outside the trial. In this case, participation incentives are necessary to induce participation even by individuals who believe  $\tilde{\pi}_{1i} > 0$ .

## 4.2 Conventional treatment

Thus far it has been assumed that there are only two treatments: the experimental treatment and no treatment. If individuals have available to them a conventional treatment outside the context of a trial, there is the possibility of self-selection based on the share treated.<sup>14</sup> If there is also a correlation between beliefs – which drive self-selection – and outcomes, then it is not obvious that the test for placebo effects set forth in Proposition 2 will work. Without placebo effects, the share treated may affect outcomes through self-selection. With placebo effects, share treated will continue to affect outcomes due to a direct relationship between beliefs and outcomes, but will also have an effect through self-selection. In the presence of a conventional alternative, self-selection is in effect noise

---

<sup>13</sup>Fortunately, there is an alternative test for placebo effects when investigators employ participation incentives. If placebo effects exist, an increase in the share treated should have an increasingly positive or decreasingly negative effect on mean outcomes conditional on treatment. To see this, divide the population into two groups: one with individuals who believe  $\tilde{\pi}_{1i} > 0$  and another with those who believe  $\tilde{\pi}_{1i} < 0$ . Everyone in the former group will participate in a blinded RPCT regardless of the share treated. Moreover, an increase in the share treated will proportionally raise this groups mean outcomes if there exist placebo effects. Members of the latter group will participate only if the utility of the participation incentive is greater than the disutility from having to risk consumption of the experimental treatment. This risk, and thus the cost of participation, rises with the share treated. So an increase in the share treated will cause some individuals in the second group to not enroll in the trial. Even if those who choose not to enroll are the least pessimistic individuals in the second group, the change in mean outcomes with the share treated, a weighted average of the change in outcomes of participants from the  $\tilde{\pi}_{1i} > 0$  and the  $\tilde{\pi}_{1i} < 0$  groups, will be more positive or less negative. The reason is that everytime the share treated rises some of the individuals that would place negative pressure on the relationship between share treated and mean outcomes no longer participate.

<sup>14</sup>In the presence of a conventional alternative, it is assumed that, although subjects are free to exit a trial at any time, if they remain, they do not consume any treatment other than that which they have been assigned. This stops subjects from obtaining conventional treatment in addition to their assigned treatment.

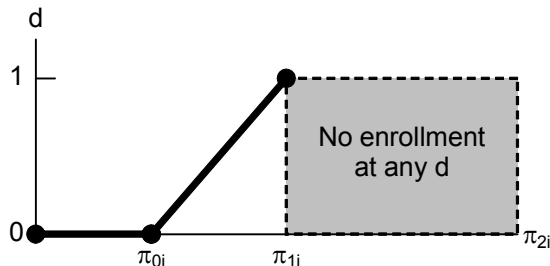


Figure 2: Share treated required to induce enrollment in a blinded RPCT at different levels of  $\pi_{2i}$ . It is assumed  $\pi_{1i} > \pi_{0i}$ .

that may obscure identification of a direct relationship between beliefs and outcomes. One solution is to rely on the fact that placebo effects imply a positive relationship between share treated and outcomes and search for conditions under which self-selection implies a negative relationship between share treated and outcomes. Under these conditions, one could continue to test for placebo effects by searching for a positive relationship between share treated and outcomes. This section sets forth these conditions, which, it turns out, are fairly reasonable.

To see that the presence of a conventional treatment implies self-selection based on share treated, let  $k = 2$  indicate the conventional treatment. An individual will enroll in a blinded RPCT in the presence of a conventional treatment if and only if  $d\pi_{1i} + (1 - d)\pi_{0i} > \max\{\pi_{0i}, \pi_{2i}\}$  or, equivalently,

$$s = BT \Leftrightarrow \tilde{\pi}_{1i} > \max\{0, \tilde{\pi}_{2i}/d\}. \quad (2)$$

In words, an individual will enroll in a trial so long as its lottery over the the experimental treatment and no treatment is better than the both no treatment or the conventional treatment for sure. If either no treatment or the conventional treatment is superior to the experimental treatment, the individual will not enroll in the trial. These superior alternatives are available with certainly outside the trial. If no treatment is superior to the conventional alternative, the subject will enroll if and only if the experimental treatment is better than no treatment – the same selection condition as in the case without a conventional treatment. *If, however, the conventional treatment is better than no treatment, but worse than the experimental treatment, the enrollment decision will depend on the share treated.* If along the continuum from the experimental treatment to no treatment, the conventional treatment is closer to the experimental treatment, it will take a high probability of obtaining the experimental treatment to attract an individual to the trial. If the conventional treatment is closer in efficacy to no treatment, then even a small probability of obtaining the experimental treatment may attract the individual to the trial. Figure 2 presents the share treated required to induce enrollment at different levels of  $\pi_{2i}$  given  $\pi_{1i} > \pi_{0i}$ .

Examining solely the subpopulation for which  $\pi_{1i} > \pi_{2i} > \pi_{0i}$ , i.e., the subpopulation for whom the share treated affects the enrollment decision, the following proposition gives conditions under which an increase in the share treated reduces average beliefs about the specific effects of the experimental treatment among the members of the subpopulation that enroll.

**Proposition 4** *Suppose that  $E_g[\tilde{\pi}_{ki}|\pi_{0i}] = E_g[\tilde{\pi}_{ki}]$ ; that  $\tilde{\pi}_{ki} > 0$  for  $k = 1, 2$ ; and that  $(\ln \tilde{\pi}_{1i}, \ln \tilde{\pi}_{2i})$  have a non-degenerate log-concave or log-convex joint density with mean  $(\mu_1, \mu_2)$  and variance  $\Sigma$ . Define  $u_{ki} = \ln \tilde{\pi}_{ki} - \mu_i$ ,  $W_i = u_{1i} - u_{2i}$ ,  $\sigma = \sigma_{11} + \sigma_{22} - 2\sigma_{12}$ ,  $a_1 = (\sigma_{11} - \sigma_{12})/\sigma$ ,  $a_2 = a_1 - 1$ , and  $V_i = a_1 u_{2i} - a_2 u_{1i}$ . By construction  $u_i = a_i W_i + V_i$ , where  $W_i$  and  $V_i$  are uncorrelated. Suppose further that  $W_i$  and  $V_i$  are actually independent. Define  $\rho_{12} = \text{corr}(\ln \tilde{\pi}_{1i}, \ln \tilde{\pi}_{2i})$ . If  $\sigma_{11} > \sigma_{12}$  or, equivalently,  $\rho_{12} < \sigma_1/\sigma_2$ , then  $\partial E_g[\pi_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d \leq 0$  and  $\partial E_g[\pi_{0i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d = 0$ .*

**Proof.** See Appendix. ■

Note that one can write  $\pi_{ki} = \pi_{0i} + \tilde{\pi}_{ki}$ , for  $k = 1, 2$ . Because natural progression is defined as the probability of recovery without treatment, this equation says that  $\tilde{\pi}_{ki}$  is individual  $i$ 's beliefs about the physiological effects of treatment  $k$ . The assumption that  $\tilde{\pi}_{ki}$  and  $\pi_{0i}$  are independent implies that the physiological effects of treatment  $k$  are independent of the natural progression of disease. The purpose of this assumption is to ensure that selection pressures due to changes in treatment share do not affect the distribution of beliefs about no treatment among enrollees.

The assumption that  $(\ln \tilde{\pi}_{1i}, \ln \tilde{\pi}_{2i})$  is log-concave is not very restrictive. The class of log-concave or log-convex densities is quite large. It includes, e.g., the bivariate normal distribution. In that case, the fact that  $W_i$  and  $V_i$  are (by construction) uncorrelated implies they are also independent. Importantly, truncation of the range of  $\tilde{\pi}_{ki}$  at, e.g., one does not alter the result in Proposition 4. See Proposition 6 in An [17].

Proposition 4 says that, so long as the covariance between (log) beliefs about the physiological effects of the experimental and conventional treatments is less than the variance of (log) beliefs about the experimental treatment, changes in the share treated, if anything, reduce enrollees' expectations regarding the experimental treatment. From the definition of the correlation coefficient, it is obvious that the condition on the covariance is satisfied whenever the variance of beliefs about the physiological effects of the experimental treatment is greater than the variance of beliefs about the conventional treatment. In this light, the condition on the covariance does not appear at all unreasonable. The experimental treatment, by virtue of being new, will be associated with greater uncertainty in beliefs among the patient population than the conventional treatment.<sup>15</sup>

The intuition behind this result begins with the observation that an RPCT only attracts individuals who believe that the experimental treatment is so much better than conventional treatment that, even with the risk of obtaining no treatment at all, enrolling in the trial is a superior strategy to opting for conventional treatment. If one alters a trial to increase the probability of obtaining the experimental treatment, a patient who is marginally not optimistic enough about the experimental treatment to have risked randomization into the placebo-control group before may now be willing to take that risk because it is smaller. Moreover, because trials only attract subjects who believe that the experimental treatment is superior to no treatment, altering a trial to increase the probability of obtaining the experimental treatment also means that the trial will be attractive to individuals who were previously just marginally too pessimistic about no treatment to enroll because there is now less risk of obtaining a placebo.

---

<sup>15</sup>Using a Taylor-series approximation, one can approximate the condition that  $\rho_{12} < \sigma_1/\sigma_2$  for  $(\ln \tilde{\pi}_1, \ln \tilde{\pi}_2)$  with the condition that  $\rho_{12*} < (\mu_{2*}/\mu_{1*})(\sigma_{1*}/\sigma_{2*})$ , where  $(\mu_*, \Sigma_*)$  are the mean and variance of  $(\tilde{\pi}_{1i}, \tilde{\pi}_{2i})$ . Where the mean belief about the physiological effect of experimental treatment is greater than the mean belief about that of the conventional treatment, the condition on the correlation between those beliefs is more constraining.

This logic is valid only if individuals who are optimistic about the experimental treatment aren't too optimistic about conventional treatment as well. If individuals who are more optimistic about the experimental treatment are also (sufficiently) more optimistic about conventional treatment, individuals who are more optimistic about the experimental treatment are not more likely to join a trial at any given level of share treated. Although the value of trial is higher given these individuals' optimism about the experimental treatment, so is their optimism about the conventional alternative. If these individuals are sufficiently optimistic about the alternative, they may prefer it to enrollment in the trial despite their high expectations for the experimental treatment.<sup>16</sup>

Self-selection implies that trials with different treatment shares may have enrollee populations with different beliefs about each treatment. If there is no correlation between individual beliefs and individual outcomes, share treated will not affect outcomes in the absence of placebo effects. Although share treated may affect outcomes in the presence of placebo effects, the direction of the relationship is not obvious. Nevertheless, one can test for placebo effect by checking for any sort of correlation between treatment shares and outcomes.

If, however, individual beliefs and individual outcomes are correlated, treatment shares will affect outcomes even in the absence of placebo effects. It is reasonable to suppose this correlation is positive. If it were negative, that would imply individuals who respond well to treatment estimate that they do not, and those that do not respond well estimate that they do. Although one might suppose individuals over- or under-estimate treatment response, surely individuals do not guess their personal response in the manner suggested by a negative correlation. Therefore, it is assumed that treatment efficacy and beliefs about efficacy are related according to the function

$$p_{ki} = f(\pi_{ki}) + v_{ki}, \quad (3)$$

where  $f$  is everywhere continuously differentiable,  $f' > 0$ , and  $v_{ki}$  is independent of  $\pi_{k'i'}$  for all  $(k', i')$  and of  $v_{k'i'}$  for all  $(k', i')$  except  $(k' = k, i' = i)$ . The error term  $v_{ki}$  reflects error in predictions of specific efficacy by individual  $i$ .

This assumption implies the following proposition.

**Proposition 5** *Suppose  $\partial E_g [\pi_{1i} | \tilde{\pi}_{1i} > \tilde{\pi}_{2i} / d] / \partial d \leq 0$  but  $\partial E_g [\pi_{0i} | \tilde{\pi}_{1i} > \tilde{\pi}_{2i} / d] / \partial d = 0$ . If trials that have higher treatment shares but otherwise are identical yield higher mean outcomes in the experimental treatment or yield different mean outcomes in no treatment groups, then there exist placebo effects as defined in section 3.*

**Proof.** See Appendix. ■

The change in trial outcomes given self-selection due to a change in share treated are a weighted sum of the change in specific effects due to self-selection, the change in beliefs about specific effects due to self-selection, and beliefs about the advantage of the experimental

---

<sup>16</sup>Consider the following numerical example. Suppose that one-half the population has beliefs  $(\pi_{0i}, \pi_{1i}, \pi_{2i}) = (0.5, 0.5, 0.5)$  and one-half has beliefs  $(0.75 - 1.5\varepsilon, 0.75 + \varepsilon, 0.75)$ . A trial with  $d = 0.5$  will attract only the first group of individuals. A trial with  $d = 0.6$ , however, will attract both groups. Yet the average subject in the second trial would be more optimistic than the average person in the first trial with respect to the efficacy of both the experimental treatment and no treatment. The reason is that correlation of  $\pi_{1i}$  and  $\pi_{2i}$  in the population is very high (specifically, greater than  $d = 0.5$ ).

treatment over no treatment:

$$\frac{\partial E_g [y_k | s = BT]}{\partial d} = (1 - a) \frac{\partial E_g [p_k | s = BT]}{\partial d} \tag{4}$$

$$+ a \frac{\partial E_g [\pi^{BT} | s = BT]}{\partial d} \tag{5}$$

$$+ a \tilde{\pi}_1 \tag{6}$$

The last term is a product of placebo effects.

The presence of a conventional treatment means that the change in beliefs about specific effects of the trial (5) may be non-zero. Proposition 4 gives the conditions under which this term is negative in the experimental treatment group and zero in the control group. The fact that beliefs and specific effects are correlated – where the direction of causation runs from specific effects to beliefs – means that the change in specific effects (4) is also non-zero in the experimental treatment group. Proposition 5 says the assumptions that  $f' > 0$  and that  $v_{ki}$  is independent of  $\pi_{k'i'}$  for all  $(k', i')$  mean the direction of changes in specific effects of a treatment are the same the direction of changes in beliefs about that treatment, i.e., are negative and zero in the case of the experimental treatment group and no treatment group, respectively.

If there are no placebo effects, changes in specific effects are all that drive outcomes. A change in the share treated will lower outcomes in the experimental group and not affect outcomes in the control group. If, however, there are placebo effects, the weight assigned to the second (5) and (6) third terms is positive. The second term is negative for reasons just given. Moreover, it is a direct implication of the selection equation (2) that the term itself is positive. The net effect is that the share treated can raise or lower outcomes in either group. *Therefore, if one finds evidence of a positive relationship between share treated and outcomes in the experimental treatment or any non-zero relationship between share treated and outcomes in the control group, there must exist placebo effects.*

Note that the test for the experimental group is subject to significant Type II error. It is possible that placebo effects exist but that the self-selection, which tends to lower outcomes, overwhelms it. If one could estimate the relationship between share treated and outcomes while controlling for the effects of self-selection embodied in (4) and (5), the risk of Type II error would be diminished. This may be possible, e.g., if one has individual level data on enrollees and non-enrollees.

Finally, it should be noted that the results from this section apply to conventional control trials. All that is required is that one switch the subscripts  $k = 0$  and  $k = 2$ .

## 5 Application to Ulcer Trials

This section tests for the existence of placebo effects in clinical trials of anti-ulcer medications that promise to heal ulcers.<sup>17</sup> Ulcer trials are chosen because they offer objective

---

<sup>17</sup>de Craen et al. [18] and Moerman [19] claim to find evidence of placebo effects in ulcer trials. However, their findings do not warrant such a conclusion. The de Craen study finds that outcomes in the placebo group of trials with a four-times-a-day (q.i.d.) regimen of placebo treatment are 6 - 8 percent higher than in placebo groups of trials with a twice-a-day (b.i.d.) regimen. This is highly suggestive of placebo effects. However, the result does not differentiate between types of ulcer medication. When the authors examined only trials of H<sub>2</sub>-receptor antagonists, the difference dropped to 3 percent. (There was no indication whether this estimate is statistically significant.) Moreover, the authors do not relate their results to any specific models of placebo effects. Nor do they formally consider the interaction between the structure of the

measures of health outcomes. Ulcers are erosion of the mucous lining in the stomach or small intestine due to acid buildup. Ulcers can objectively be judged healed via endoscopy, which examines the stomach lining for evidence of damage. Moreover, because ulcers are such a common problem throughout the world, a large number of trials have been conducted.

The majority of ulcer trials examine three types of medication. The first type, H<sub>2</sub>-blocker, was introduced in 1977. The most popular brands are Tagamet (cimetidine), Zantac (ranitidine), and Pepcid (famotidine). H<sub>2</sub>-blockers prevent the production of acid in the stomach.<sup>18</sup> The second type of medication, prostaglandin, was introduced in 1987. The most common prostaglandins are misoprostil and enprostil. These drugs build up and thus repair the mucous lining of the stomach and intestine. The third class, proton-pump inhibitor, and were introduced after prostaglandins. The most popular brands are Prilosec (omeprazole), Nexium (esomeprazole) and Prevacid (lansoprazole). Like H<sub>2</sub>-blockers, these medications prevent the production of acid in the stomach.

## 5.1 Data

The data set includes the published results from over 150 clinical trials studying treatment for pyloric, pre-pyloric and duodenal ulcers.<sup>19</sup> Each of the trials is randomized, parallel-armed, and double-blind, and employs either a placebo, antacid, bismuth subcitrate or conventional control. If conventional controls are employed, they are from either the same or a previous class of medication as the experimental treatment. Importantly, subjects in 110 of the trials were asked for informed consent prior to enrollment. Hence it is reasonable to suppose that subjects in those trials had some indication of their chance of obtaining the experimental treatment.

Data were gathered on the characteristics of trials and of subjects. Data on subjects are aggregated to the arm- or group-level. For example, there are data on the average age of subjects assigned to any given treatment group, but not the age of each subject assigned to that group. Although there are data on subjects in a group as of the date that they are randomized into the group, precise information on how the group changes due to attrition are not available. Table 2 provides summary statistics for the data, which are analyzed at the treatment-group level. Data from groups examining H<sub>2</sub>-blockers, prostaglandins, and proton-pump inhibitors are presented separately.<sup>20</sup>

---

randomized, double-blind, placebo-controlled trials they examine and any theory of placebo effects they purport to test. This is particularly important if, e.g., patients in q.i.d. trials receive the same total dosage as or different treatment than those in b.i.d. trials. (In the sample of ulcer trials I consider, all 12 of the b.i.d. trials examined ranitidine while 29 of 30 q.i.d. trials examined cimetidine. Both are H<sub>2</sub>-blockers, but require substantially different dosages. Moreover, the one q.i.d. trial of ranitidine involved the same total daily dosage as the 12 b.i.d. trials of the drug.) In that case, the result would be difficult to explain even with the model of health outcomes in Section 2.2. Nevertheless, the de Craen study's findings are very interesting, and ought to be subject to the type of analysis in this paper.

Moerman examines the same set of ulcer trials and finds a significant positive correlation (0.49) between outcomes in the placebo group and the treatment group. Unfortunately, Moerman's finding is not very informative: the positive correlation can be explained by self-selection alone. Such pressures tend to depress outcomes in both the new treatment and placebo groups. Moreover, Moerman does not confine his sample to trials with informed consent. It is unclear what information he thinks subjects had or why they responded as they did.

<sup>18</sup>In contrast, antacids are alkali that absorb acid in the stomach. They reduce the amount of acid available to damage the stomach lining, but not enough to permit the healing of damage to that lining.

<sup>19</sup>These are ulcers that arise just before the pyloric tract, in that tract, or in the duodenum, respectively. All are located just before the start of the small intestine.

<sup>20</sup>There are several things to note about the data and Table 2. First, there are more groups given the

## 5.2 Empirical model

The specific effects of treatment  $k$  on individual  $i$  enrolled in trial  $j$  are assumed to be a deterministic, linear function of the vector  $x_{ij}$ , which includes a constant, clinical and demographic variables on individual  $i$ , and structural features of trial  $j$ :  $p_{kij} = \beta'_k x_{ij}$ . This is a strong assumption, but because the ulcer trial data are rather coarse there is little benefit from a more nimble parameterization of  $p_{kij}$ . Beliefs regarding specific effects are assumed to be given by  $\pi_{kij} = \gamma'_k x_{ij} + \varepsilon_{kij}$ ,<sup>21</sup> where  $\varepsilon_{kij}$  is independent of  $x_{ij}$  and is i.i.d. mean-zero across individuals, trials, and treatment states, with mean zero and variance  $\sigma_\varepsilon$ . This condition on  $\varepsilon_{kij}$  implies that individual errors in estimating specific treatment response do not depend on the treatment.

Because trials often take multiple measurements on each individual, a treatment's effect is cast as a hazard rate.<sup>22</sup> Assuming it is constant over time, (1) and the assumed parameterization of  $(p_{kij}, \pi_{ij})$  imply  $-\ln S_{ijk}(t)/t = \theta_{(x)k} x_{ij} + \theta_{(xd)k} d_j x_{ij} + \eta_{ijk}$ ,<sup>23</sup> where  $S_{ijk}(t)$  gives the probability of still having an unhealed ulcer on date  $t$ ,  $\theta_{(x)k} = (1 - a_k) \beta_k + a_k \gamma_0$ ,  $\theta_{(xd)k} = a_k (\gamma_1 - \gamma_0)$ , and  $\eta_{ijk} = a_k d_j \varepsilon_{1ij} + a_k (1 - d_j) \varepsilon_{0ij}$ . Summing over individuals and dividing by  $n_{jk}$ , the number of subjects enrolled in treatment arm  $k$  of trial  $j$ , yields the regression equation  $-\ln \overline{S_{jk}(t)}/t = \theta_{(x)k} \bar{x}_{jk} + \theta_{(xd)k} d_j \bar{x}_{jk} + \bar{\eta}_{jk}$ . The left-hand side is a first-order Taylor approximation around  $\bar{S}_{jk}(t)$ , the average probability of remaining ill at  $t$ .  $\ln \overline{S_{jk}(t)} = \ln(\sum_i S_{ijk}(t)/n_{jk})$  is approximated by  $\ln(\bar{S}_{jk}(t))$ , the log of the observed group survival rate. Approximating  $\bar{S}_{jk}(t)$  is difficult because data on subjects who attrite out of the trials are not available. Therefore,  $\bar{S}_{jk}(t)$  is calculated under three different assumptions: individuals who attrite out heal at the same rate as those who remain, individuals who attrite out do not heal, and these individuals all heal.

---

experimental treatment than given the control. The reason is that each trial typically involves one control arm but multiple experimental treatment arms. Typically these arms will vary the total daily dosage or the daily frequency of medication. Second, the probability of active treatment is estimated by one minus the probability of randomization into the placebo, antacid or bismuth subcitrate group. The reason is that, while individuals may have detectably different beliefs about the efficacy of some active treatment versus no treatment, it is deemed unlikely that individuals have sufficiently refined beliefs about different dosages of a given active treatment that one can identify the responsiveness of outcomes to changes in the probability of randomization into each dosage arm. Thus variation in the number of new-treatment arms across trials generates variation in the share treated in my data. Importantly, trials with same-class controls are assumed to have a probability of new treatment equal to one. Third, 33, 40, and 80 percent of H<sub>2</sub>-blocker, prostaglandin, and proton-pump inhibitor trials, respectively, have lower-class drug conventional controls. Fourth, the antacid-permitted variable is coded from 1 to 5. One indicates that subjects were prohibited from taking antacids, two that subjects were discouraged from taking antacids, three that subjects were permitted to take antacids (or the study did not counsel subjects on antacids), four that antacids were provided, and five that antacids were required. Fifth, the difference in total dosage and dosage frequency between different classes of anti-ulcer medications has little significance. Because the classes have different chemistries and modalities, their recommended dosages are not comparable. Moreover, the total dosage of placebo, antacid or bismuth subcitrate controls is omitted because it has little meaning given that the control is either inert or subjects are typically permitted to take antacid.

<sup>21</sup>If  $\gamma_k = \beta_k$ , then the assumption implies rational expectations. The assumption of linearity limits the scope of cognitive errors that may plague individual projections, but should not otherwise be controversial. The error term measures mistakes in prediction by an individual.

<sup>22</sup>Survival analysis is better able to employ information on the timing of treatment response than, say, a simple qualitative dependent-variable framework. Viewing (1) as a hazard function does not affect selection into the trial so long as the hazard function is time-invariant.

<sup>23</sup>Division by  $t$  should not be problematic because ulcer trials are fairly short (maximum 12 weeks for trials in the data set and 12-18 months in long-term ulcer trials) so  $t$  is small.

The regression equation that is ultimately estimated is

$$-\frac{1}{t} \ln (\bar{S}_{jkt}) = \theta_{(x)k} \bar{x}_j + \theta_{(xd)k} d_j \bar{x}_j + \omega_{jk}, \quad (7)$$

for  $k = 0, 1$ , where  $\bar{S}_{jkt}$  is approximated in one of the three methods discussed in the last paragraph,  $\bar{x}_j = (1/n_{jk}) \sum_i x_{ij}$ ,  $\omega_{jk} = \bar{\eta}_{jk} + u_{jkt} + v_{jkt}$ , and  $u_{jkt}$  is the error from approximating  $\sum_i S_{ijk}(t)$  with some function of  $y_{jkt}$ . The Lagrange remainder from approximating  $\sum_i \ln S_{ijk}(t)$  with  $n_{jk} \ln \sum_i S_{ijk}(t)$  is absorbed into the coefficient on the constant.<sup>24</sup> For simplicity this is left out of the definition of  $\theta_{(x)k}$ . The error term  $v_{jkt}$  captures the variation in the remainder across arms and trials.<sup>25</sup>

Under the conditions set forth in Section 4, a test for placebo effects is whether there is a positive relationship between  $d_j$  and survival. In the presence of a conventional alternative, this test may be complicated by self-selection based on share treated, which implies that  $E(\varepsilon_{kij} d_j | s = BT) \neq 0$ . This problem is addressed in three different ways. The first method, which is labeled the “no-conventional-alternative” approach, involves estimating (7) on a subsample of H<sub>2</sub>-blockers for which it is reasonable to believe there was no conventional alternative and thus no self-selection based on share treated. It is hypothesized that this includes all H<sub>2</sub>-blocker trials before 1987, when prostaglandins are first introduced.<sup>26</sup> Because there is no obvious conventional alternative-free subsample for the other class of drugs, one cannot test for placebo effects employing this first approach to selection. The second method, which is labeled “ $x$ -captures-selection” approach, is to employ the entire sample of trials for each class of anti-ulcer drug but assume that selection pressures are fully captured by observable variables. With the first two approaches, placebo effects are tested for by counting the number of treatment arms for which the predicted  $\hat{\theta}_{(xd)k} \bar{x}_j$  is significantly greater than zero. A third method, which is labeled “additional- $d$ ” approach, is to assume  $E(\varepsilon_{kij} d_j | s = BT) = 0$  but partition  $\bar{x}_j = (\bar{x}_j^o, \bar{x}_j^u)$ , where  $\bar{x}_j^o$  is observable but  $\bar{x}_j^u$  may not be.  $\bar{x}_j^o$  is chosen such that, as a theoretical matter, it ought to be  $\theta_{(xd)k}^o d_j \bar{x}_j^o > 0$ . Given that selection pressures depend on  $d_j$ , it is assumed that  $\theta_{(xd)k}^u d_j \bar{x}_j^u = \phi_{(xd)k} d_j^2 + e_{jk}$ , where  $e_{jk}$  is independent of  $d_j$  and  $\bar{x}_j^o$ . If selection is a problem but there are no placebo effects,  $\theta_{(xd)k}^u$  and thus  $\phi_{(xd)k}$  should be zero. Thus, if the estimate of  $\phi_{(xd)k}$  is significantly different from zero, then there must exist placebo effects or selection is not a problem.

### 5.3 Results

Table 3 presents results employing the  $x$ -captures-selection method of testing for placebo effects in H<sub>2</sub>-blocker trials.<sup>27</sup> The first four columns present coefficient estimates for treatment arms and different specifications of  $\bar{x}_j$ ; the last four do the same for control arms.

<sup>24</sup>This approach bears some resemblance to Amemiya and Nold’s [20] adaptation of the logit model to grouped data.

<sup>25</sup>Because the hazard rate is  $\Pr\{y_{ki} = \bar{y}|s\} \in [0, 1]$ , the parameterization in this paper requires estimation of a linear probability model. While that model has been criticized for, e.g., potentially generating predicted values outside the 0-1 range, it is not wholly inappropriate for the application in this paper. As a theoretical matter, the dependent variable in (7) can range from  $(0, \infty)$ . Moreover, because individuals in an arm are aggregated, the error term is more likely to resemble a normal distribution [21]. Finally, (7) was modified and estimated as a generalized linear model with the dependent variable obeying the binomial distribution and a logistic link function. However, nearly every specification performs worse on Pregibon’s [22] link test than the linear probability model.

<sup>26</sup>Antacids are a poor substitute because they cannot heal an ulcer. Moreover, most trials permit subjects to consume antacids as well as assigned treatment.

<sup>27</sup>Estimation was by feasible GLS. Each measurement on an arm of a trial counts as an observation. Some trials offer multiple measurements on the same arm. Observations were weighted such that each

Below the coefficient estimates for each specification are a panel of F-tests of the joint significance of certain subsets of regressors interacted with  $d_j$ . Their purpose is to test whether placebo effects operate through any of these subsets of variables. Below the F-tests are four further panels. The first gives the mean (and standard deviation of) estimated marginal effect of share treated on outcomes:

$$\partial[-\ln(\bar{S}_{jkt})/t]/\partial d_j = \hat{\theta}_{(xd)k\bar{x}_j} \approx \partial y_j / \partial d_j$$

The estimate of  $\partial y_j / \partial d_j$  for each observation – a measurement on a particular arm of a particular study – depends on the value of covariates for that observation and therefore has its own standard error. The second panel gives the mean t-statistic for each estimate and the standard deviation of the t-statistic across all the measurements. To complete the picture that these panels paint about the distribution of the estimated  $\partial y_j / \partial d_j$  within and across measurements, the third panel counts the number of measurements where the share treated is estimated to have raised outcomes at different levels of confidence (employing a one-sided test). The fourth counts of the number of measurements where the share treated either raised or depressed outcomes (employing a two-sided test).

For reference purposes, the theory in section 4 can be summarized as follows. Assuming (1) is a correct model of health outcomes, if there is no self-selection, either because there is no conventional alternative or selection is controlled by inclusion of appropriate covariates, then a positive relationship between share treated and outcomes is evidence of placebo effects. If there is self-selection, but beliefs about specific effects and specific effects themselves are not correlated, then any non-zero relationship between share treated and outcomes is evidence of placebo effects. Finally, if there is self-selection and beliefs and specific effects are positively correlated, but the physiological effects of treatments are independent of the natural progression of disease, then a positive relationship and non-zero relationship between share treated and outcomes in the experimental treatment and control groups, respectively, are evidence of placebo effects. The experimental group test, however, is subject to type II error: there may exist placebo effects, but it is drowned out by a self-selection effect.

The mean estimated marginal effect of share treated on outcomes in treatment arms is positive across all specifications. All measurements in the first specification manifest a positive relationship at the 95 percent confidence level. Around three-quarters of measurements in other specifications manifest a positive relationship at the 90 percent confidence level; over one-half do so at the 95 percent level. This last finding is illustrated graphically

---

arm makes a contribution to estimates in proportion to the number of subjects in the arm, regardless of the number of measurements made on each subject. The regression model suggests that the variance of error terms depends on the share randomized into each arm. However, only one randomization share per trial – namely the share not given a lower class or non-healing control – is measured. Therefore group-wise heteroskedasticity is permitted at the trial-level, but not at the arm-level. Only estimates where the dependent variable is calculated assuming subjects who attrite out heal at the same rate as those who are evaluated are reported. Results from regressions which assume that those who attrite out either all heal or all do not heal are not materially different.

Four specifications of  $\bar{x}_j$  are estimated. Specification (1) includes a constant and  $d_j$ ; (2) adds trial-level variables (antacid usage, daily frequency of medication, total daily dosage of medication, total daily dosage of the more common drugs in the relevant class of medications) as well as interactions of these trial-level variables with  $d_j$ ; (3) adds subject-level variables (sex, smoker, and age) and their interactions with  $d_j$ ; and (4) removes the trial-level variables and interactions from (3). When the  $x$ -captures-selection method is employed to control for selection,  $d_j^2$  is added to each specification. Each specification is checked against Pregibon’s link test [22]. Those that fail are marked with a dagger (†). The residuals were checked for but did not reveal non-sawtooth patterns.

in Figures 3 and 4. The former plots the estimated  $\partial y_j / \partial d_j$  from specification two in ascending order, along with the 95 percent confidence interval for each estimated  $\partial y_j / \partial d_j$ . The latter plots how estimated marginal effect of share treated varies with one covariate, for specification two and an arbitrary subsample of the data. The covariate is the number of times per day medication was administered in arm  $j$ . The subsample is measurements that take place two weeks after enrollment in a trial. In contrast, the mean estimated marginal effect of share treated on outcomes in control arms is negative. Although, no arms that pass Pregibon’s link (specification) test manifest a positive relationship, significant or otherwise, between one-half and all measurements manifest a negative relationship significant at the 95 percent confidence level.

If the covariates of a specification are able to control for selection, these findings suggest that the treatment arms of H<sub>2</sub>-blocker trials manifest significant evidence of placebo effects, but the treatment arms do not. If the covariates do not fully control for selection, then both arms of these trials manifest significant evidence of placebo effects. The treatment arms do so because they manifest a positive relationship between share treated and outcomes; the control arms do so because they manifest a non-zero relationship. The results of the F-tests do not permit one to conclude that placebo effects fail to operate through any natural subset of variables.

Table 4 presents further results for treatment arms, but for all three classes of anti-ulcer drug, by specification and approach to selection. Table 5 does the same for control arms of trials. The tables present only the mean estimated marginal effect of share treated; counts of arms that manifest a significant positive relationship between share treated and outcomes; and counts of arms that manifest a significant non-zero relationship between share treated and outcomes. Coefficient estimates for the entire model and F-tests are omitted.

Arms treated with H<sub>2</sub>-blockers manifest strong evidence of placebo effects. With the no-conventional-alternative approach approximately one-third of measurements on these arms manifest evidence of a positive relationship at the 95 percent confidence level in specifications one and two. The additional- $d$  approach supports an inference of placebo effects in the first, third and fourth specifications, which return a coefficient on  $d_j^2$  that is significantly different from zero. Arms treated with prostaglandins manifest weak evidence of placebo effects. While no arms manifest evidence of a positive relationship using the  $x$ -captures-selection approach, nearly all manifest a non-zero relationship at the 95 percent confidence level under specifications one and two. The additional- $d$  approach reveals significant evidence of placebo effects in the first two specifications. Arms treated with proton-pump inhibitors manifest moderate evidence of placebo effects. Around a quarter of measurements manifest a significant positive relationship under the  $x$ -captures-selection approach in specifications two to four. At least a quarter, and often nearly all, measurements manifest evidence of a non-zero relationship. However, under the additional- $d$  method, only the second specification manifests placebo effects.

Turning to the control arms of these trials, while virtually no H<sub>2</sub>-blocker trials reveal evidence of a positive relationship under the no-conventional-alternative approach, nearly all trials reveal a non-zero relationship in specifications one and two. Moreover, the additional- $d$  approach reveals evidence of placebo effects in specifications two through four. There is not enough evidence to conclude that there are placebo effects in the control arms of prostaglandin trials using the  $x$ -captures-selection approach. This is likely a result of the small number of observations on prostaglandin trials. Nevertheless, there is evidence of placebo effects in specifications one and two employing the additional- $d$  method. Finally,

around a third of measurements on control arms of proton-pump inhibitor trials manifest evidence of a positive relationship under specifications three and four when the  $x$ -captures-selection method is employed. At least two-third of measurements manifest evidence of a significant non-zero relationship in specifications two through four. Moreover, specification one manifests evidence of placebo effects when the additional- $d$  approach is employed.<sup>28</sup>

Only trials where published reports confirm that subjects were asked for informed consent were included in these regression. The sample also contains, however, 20 trials of H<sub>2</sub>-blockers where it cannot be confirmed that informed consent was requested. Regressions employing only these trials reveal significantly diminished evidence of placebo effects, except in the first specification employing  $x$ -captures-selection approach.<sup>29</sup> (These results are omitted.) This finding is consistent with the assumption that individuals learn the probability of receiving the experimental treatment via informed consent. Without this disclosure, an individual’s beliefs about the probability of treatment are unrelated to the actual share treated. Therefore, outcomes, even if they depend on beliefs, should be invariant to the share treated in a trial.

## 6 Conclusion

This paper provides evidence of placebo effects in blinded, parallel-arm RCTs of H<sub>2</sub>-blockers and proton-pump inhibitors. The findings are summarized in Table 6. The results are notable because they are derived from a model with clear assumptions and falsifiable predictions. These features make interpretation of findings easier and lend credibility to the conclusion that placebo effects affect ulcer healing.

Further investigation is required, however, before placebo effects can be labeled a serious medical phenomenon. These effects may be ailment, drug and context specific. Future research ought to proceed in two directions. One is theoretical. This paper assumes health outcomes follow a simple linear model and placebo effects have the same influence regardless of treatment taken. It would be useful to have a more general test for placebo effects that relaxes these assumptions. This paper also examines only one type of trial. Many other designs are employed in the medical literature. Malani [16] examines how robust the test for placebo effects proposed in this paper is, e.g., to a multiplicative model of outcomes, and proposes tests for placebo effects in unblinded and cross-over trials. However, this work is just a beginning. A second avenue for research is application to other medical contexts, such as trials for other ulcer drugs, especially antibiotics, and trials of drugs for other ailments. Such research would also dispel concern that the findings of placebo effects are an artifact of data from particular ulcer trials.

If placebo effect are found to be a widespread phenomenon, two questions will naturally follow. First, why do beliefs affect outcomes? Are placebo effects a behavioral or physiological phenomenon? Does a higher probability of treatment cause enrollees to take greater

---

<sup>28</sup>The proportional hazard model

$$\ln(-\ln(\bar{S}_{jkt})) = \theta_{(t)k}t + \theta_{(x)k}\bar{x}_j + \theta_{(xd)k}d_j\bar{x}_j + \omega_{jk}$$

was also estimated and produced similar results. The main difference is that there is somewhat stronger evidence of placebo effects. The coefficient on time is positive and significant, suggesting the risk of healing rises over time. The results from the proportional hazard model are not reported because far fewer specifications pass the link test.

<sup>29</sup>A subsample that also included trials conducted outside the U.S., Canada, Western Europe, Australia or New Zealand yield similar results.

care of themselves in a way that improves their outcomes? Or does knowledge of higher probability of treatment trigger a heightened immune response in enrollees? The answer may change our understanding of the connection between cognition and physiology.

A second question is: what are the implications of placebo effects for clinical practice? Malani [23] demonstrates that placebo effects undermine the internal validity of clinical trials. In particular, they cause investigators to underestimate the effect of treatment on the treated. Outside the context of trials, enrollees would know whether they are taking medication. Inside a blinded trial, they do not. Those randomized into the experimental treatment group underperform – relative to knowing consumption of the experimental treatment outside the trial – because they think there is a chance they are getting the less valuable control treatment. Conversely, those randomized into the control group overperform – relative to knowing consumption of no treatment outside a trial – because there is chance they may be getting the more valuable experimental treatment. The difference between mean outcomes in the experimental and no treatment groups of a trial will therefore be less than the difference between outcomes if enrollees consumed experimental treatment and no treatment outside the trial. This negative bias may not only skew prescription practice, but also cause the U.S. FDA not to approve valuable drugs.

An even more intriguing question is whether placebo effects can be used for therapeutic purposes? For example, can a doctor cure a patient by suggesting that a drug is more effective than it really is or by falsely telling the patient that she is receiving treatment? If so, and if fooling patients is less costly than producing drugs, then placebo effects may be able to reduce the costs of health care. The existence of rational expectations may limit the productivity of placebo effects as a long-term cost-cutting measure. It is an open question, however, how substantial the short-run cost savings might be.

## A Proofs

**Proposition 4.** All expectations are taken with respect to  $\mathbf{g}_\pi$ . Independence of  $\tilde{\pi}_{ki}$  and  $\pi_{0i}$  for  $k = 1, 2$  implies that  $\partial E[\pi_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d = 0$  and  $\partial E[\pi_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d = \partial E[\tilde{\pi}_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d$ . Given  $\tilde{\pi}_{ki} > 0$ , for  $k = 1, 2$ ,  $E[\tilde{\pi}_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d] = E[\tilde{\pi}_{1i}|\ln \tilde{\pi}_{1i} - \ln \tilde{\pi}_{2i} > -\ln d]$ . Because  $\ln x$  is monotone increasing in  $x$ ,

$$\text{sign} \left( \frac{\partial E[\tilde{\pi}_{1i}|\ln \tilde{\pi}_{1i} - \ln \tilde{\pi}_{2i} > -\ln d]}{\partial d} \right) = \text{sign} \left( \frac{\partial E[\ln \tilde{\pi}_{1i}|\ln \tilde{\pi}_{1i} - \ln \tilde{\pi}_{2i} > -\ln d]}{\partial d} \right).$$

Because  $W_i$  and  $V_i$  are independent,

$$E[\ln \tilde{\pi}_{1i}|\ln \tilde{\pi}_{1i} - \ln \tilde{\pi}_{2i} > -\ln d] = \mu_1 + a_1 E(W_i|W_i > c(d, \mu)),$$

where  $c(d, \mu) = -(\mu_1 - \mu_2) - \ln d$ .

Log-concavity or log-convexity of  $(\ln \tilde{\pi}_{1i}, \ln \tilde{\pi}_{2i})$  implies log-concavity or log-convexity, respectively, of  $W_i$ , by theorem 5 in Bagnoli and Bergstrom [24] and **[what]**, respectively. Propositions 1 and 2 in Heckman and Honore [25] demonstrate that log-concavity or log-convexity of  $W_i$  implies  $\partial E[W_i|W_i \geq c]/\partial c \geq 0$ . Therefore,

$$\frac{\partial E[\ln \tilde{\pi}_{1i}|\ln \tilde{\pi}_{1i} - \ln \tilde{\pi}_{2i} > -\ln d]}{\partial d} = \frac{\sigma_{11} - \sigma_{12}}{\sigma} \frac{\partial E[W_i|W_i \geq c]}{\partial c} \frac{\partial c(d, \mu)}{\partial d}.$$

Because  $\partial c(d, \mu)/\partial d < 0$ ,  $\partial E[\ln \tilde{\pi}_{1i}|\ln \tilde{\pi}_{1i} - \ln \tilde{\pi}_{2i} > -\ln d]/\partial d \leq 0$  so long as  $\sigma_{11} > \sigma_{12}$ . This condition is the same as  $\rho_{12} < \sigma_1/\sigma_2$ .

**Proposition 5.** All expectations are taken with respect to  $\mathbf{g}_\pi$ . By (1)

$$\begin{aligned} \frac{\partial E[y_{ki}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]}{\partial d} &= (1-a) \frac{\partial E[p_{ki}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]}{\partial d} \\ &+ a \left[ d \frac{\partial E[\tilde{\pi}_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]}{\partial d} + \frac{\partial E[\pi_{0i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]}{\partial d} \right] + a\tilde{\pi}_{1i}. \end{aligned}$$

Working backwards, the selection equation (2) implies  $\tilde{\pi}_{1i} > 0$ . By assumption,  $\partial E[\pi_{0i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d = 0$ . The proof to Proposition 4 demonstrates that

$$\text{sign}(\partial E[\tilde{\pi}_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d) = \text{sign}(\partial E[\pi_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d),$$

so  $\partial E[\tilde{\pi}_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d \leq 0$ . By (3),

$$\frac{\partial E[p_{ki}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]}{\partial d} = \frac{\partial E[f(\pi_{ki})|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]}{\partial d} + \frac{\partial E[v_{ki}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]}{\partial d}.$$

The last term is zero because  $v_{ki}$  is assumed independent of  $\pi_{k'i'}$  for all  $(k', i')$  and of  $v_{k'i'}$  for all  $(k', i')$  except  $(k' = k, i' = i)$ . Since  $f' > 0$ ,

$$\text{sign}(\partial E[f(\pi_{ki})|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]) = \text{sign}(\partial E[\pi_{ki}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]).$$

So  $\partial E[p_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d \leq 0$  and  $\partial E[p_{0i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d = 0$ .

If there are no placebo effects, i.e.,  $a = 0$ , then  $\partial E[y_{1i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d \leq 0$  and  $\partial E[y_{0i}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d = 0$ . If, however, there exist placebo effects, the sign of  $\partial E[y_{ki}|\tilde{\pi}_{1i} > \tilde{\pi}_{2i}/d]/\partial d$  is ambiguous. If one finds that outcomes rise with share treated in the experimental treatment group or are at all correlated with share treated in the control group and the model of health outcomes in (1) is accurate, then it must be that  $a > 0$ , i.e., that placebo effects exist.

## References

- [1] A. Hrobjartsson and P. C. Gotzsche, "Is the placebo powerless? an analysis of clinical trials comparing placebo with no treatment," *New England Journal of Medicine*, vol. 344, no. 21, pp. 1594–1602, 2001.
- [2] I. Kirsch and G. Sapirstein, "Listening to prozac but hearing placebo: A meta-analysis of antidepressant medication," *Prevention and Treatment*, vol. 1, no. 0002a, p. <http://www.journals.apa.org/prevention/volume1/pre0010002a.html>, 1998.
- [3] T. J. M. A. S. Kirsch, Irving and S. S. Nicholls, "The emperor's new drugs: An analysis of antidepressant medication data submitted to the u.s. food and drug administration," *Prevention and Treatment*, vol. 5, no. 0023a, p. <http://www.journals.apa.org/prevention/volume5/pre0050023a.html>, 2002.
- [4] A. Hrobjartsson, "The uncontrollable placebo effect," *European Journal of Clinical Pharmacology*, vol. 50, pp. 345–348, 1996.
- [5] S. Penick and L. E. Hinkle, "The effect of expectation on response to phenmetrazine," *Psychosomatic Medicine*, vol. 26, no. 4, pp. 369–374, 1964.

- [6] S. B. Penick and S. Fisher, “Drug-set interaction: Psychological and physiological effects of epinephrine under differential expectations,” *Psychosomatic Medicine*, vol. 27, no. 2, pp. 177–182, 1965.
- [7] G. A. Marlatt and D. J. Rohsenow, “Cognitive processes in alcohol use: expectancy and the balanced placebo design,” in *Advances in Substance Abuse*, N. Mello, Ed. Greenwich, Conn.: JAI Press, 1980.
- [8] I. Kirsch and L. J. Weixel, “Deouble-blind versus deceptive administration of a placebo,” *Behavioral Neuroscience*, vol. 102, no. 2, pp. 319–323, April 1988.
- [9] M. P. Jensen and P. Karoly, “Motivation and expectancy factors in symptom perception: A laboratory study of the placebo effect,” *Psychosomatic Medicine*, vol. 53, pp. 144–152, 1991.
- [10] M. A. A. C. C. G. M. Pollo, Antonella and F. Benedetti, “Response expectancies in placebo analgesia and their clinical relevance,” *Pain*, vol. 93, pp. 77–84, 2001.
- [11] L. S. M. I. K. A. D. G. H. M. Price, Donald D. and S. S. Nicholls, “An analysis of factors that contribute to the magnitude of placebo analgesia in an experimental paradigm,” *Pain*, vol. 83, pp. 147–156, 1999.
- [12] T. Philipson and J. DeSimone, “Experiments and subject sampling,” *Biometrika*, vol. 84, pp. 221–234, 1997.
- [13] C. L. P. Voudouris, Nicholas J. and G. Coleman, “The role of conditioning and verbal expectancy in the placebo response,” *Pain*, vol. 43, pp. 121–128, 1990.
- [14] —, “Conditioned response models of placebo phenomena: further support,” *Pain*, vol. 38, pp. 109–116, 1989.
- [15] Y. M. M. A. Berthelot, Jean-Marie and A. Prost, “Interindividual variations in beliefs about the placebo effect: a study in 300 rheumatology inpatients and 100 nurses,” *Joint Bone Spine*, vol. 68, pp. 65–70, 2001.
- [16] A. Malani, “Placebo effects, self-selection and the external validity of clinical trials,” Ph.D. dissertation, University of Chicago, Department of Economics, Sept. 2003.
- [17] M. Y. An, “Log-concave probability distributions: Theory and statistical testing,” June 1996, centre for Labour Market and Social Research, University of Aarhus, Working Paper 96-01.
- [18] D. E. M. S. H. H. G. N. T. J. G. T. de Craen, Anton J. M. and J. Kleijnen, “Placebo effect in the treatment of duodenal ulcer,” *British Journal of Clinical Pharmacology*, vol. 48, pp. 853–860, 1999.
- [19] D. E. Moerman, “Cultural variations in the placebo effect: ulcers, anxiety, and blood pressure,” *Medical Anthropology Quarterly*, vol. 14, pp. 51–72, 2000.
- [20] T. Amemiya and F. Nold, “A modified logit model,” *Review of Economics and Statistics*, vol. 57, no. 2, pp. 255–257, May 1975.
- [21] G. Maddala, *Limited Dependent and Qualitative Variables in Econometrics*, ser. Econometric Society Monographs. New York: Cambridge University Press, 1983.

- [22] D. Pregibon, “Goodness of link tests for generalized linear models,” *Applied Statistics*, vol. 29, pp. 15–24, 1980.
- [23] A. Malani, “The internal and external validity of clinical trials in the presence of placebo effects,” 2003, university of Virginia.
- [24] M. Bagnoli and T. Bergstrom, “Log-concave probability and its applications,” 1989, university of Michigan.
- [25] J. Heckman and B. E. Honore, “The empirical content of the roy model,” *Econometrica*, vol. 58, no. 5, pp. 1121–1149, September 1990.
- [26] E. Skovlund, “Should we tell trial patients that they might receive placebo?” *Lancet*, vol. 337, p. 1041, 1991.
- [27] G. F. H. L. Skovlund, E. and B.-I. Nesheim, “Comparison of postpartum pain treatments using a sequential trial design i: paracetamol versus placebo,” *European Journal of Clinical Pharmacology*, vol. 40, no. 4, pp. 343–347, 1991.
- [28] —, “Comparison of postpartum pain treatments using a sequential trial design ii: naproxen versus paracetamol,” *European Journal of Clinical Pharmacology*, vol. 40, no. 6, pp. 539–542, 1991.

Source	Findings
Skovlund (1991)[26]	Noted that patients in the treatment group of a trial which compared paracetamol to naproxen (conventional control) [27] responded better than patients in the treatment group of a trial which compared paracetamol to placebo [28]. Likewise, patients in the control group of the second trial responded better than patients in the control group of the first trial. Skovlund suggests that the patients in the naproxen-control trial may have responded better because they knew that, no matter what, they would get an active drug, whereas the patients in the placebo-control trial were skeptical because they knew they might receive a placebo. Because a better control group option improves expected outcomes, these trials suggest the importance of beliefs about the efficacy of the control-treatment to outcomes. Because the effect was recorded in both treatment and control groups, the trials underscore the relevance of beliefs about treatment state to outcomes in both states.
Pollo et al. (2001)[10]	Studied the placebo effect among thoracotomized patients, i.e., patients with their chests surgically opened for, e.g., heart operation). They divided patients into three groups and gave each group the pain killer buprenorphine on demand as well as an intravenous infusion of saline solution. The first group (natural history) was told nothing about the saline drip. The second group (double blind) was told the saline drip might be a powerful painkiller or it might be a placebo. The third group (deception) was told the saline drip was a potent painkiller. The authors found that the natural-history group asked for buprenorphine more often than the double-blind group, which asked for buprenorphine more often than the deception group. This study suggests that increasing expectations about the probability of being treated with a powerful painkiller (the saline solution) produced better health outcomes as measured by the level of demand for a second pain killer (buprenorphine).
Marlatt and Rohsenow (1980)[7]	Examined the impact of alcohol on cognition employing a balanced-placebo design wherein subjects are first randomized across treatment states and then across instructions about treatment states, with one group being told that they were given active treatment and the other being told they were given placebo. They find superior outcomes among those told they were administered active treatment, controlling for actual treatment state. This suggests that beliefs about treatment state impact outcomes.

Table 1: Studies supporting the expectancy theory of placebo effects.

	Trials of								
	H2-Blockers			Prostaglandins			PPIs		
	Obs.	Mean	SD	Obs.	Mean	SD	Obs.	Mean	SD
<b>Treatment arm</b>									
Date trial results published	225	1987	4.62	46	1988	3.62	97	1991	3.97
Share of subjects not given control	225	0.88	0.20	46	0.56	0.08	97	0.61	0.16
Share of arms in placebo-control trials	225	0.22	0.41	46	0.47	0.50	97	0.08	0.27
Share in antacid-control trials	225	0.06	0.23	46	0.00	0.00	97	0.00	0.00
Share in lower-class drug control trials	225	0.01	0.09	46	0.53	0.50	97	0.84	0.36
Share same-class drug control trials	225	0.72	0.45	46	0.00	0.00	97	0.08	0.27
Average number of measurements	225	1.42	0.58	46	1.34	0.52	97	1.60	0.65
Number enrolled	225	198	147	46	123	110	97	110	40
Number evaluated (per protocol)	225	181	133	46	108	96	97	102	37
Share of subjs. not healed (method 1)	225	0.22	0.17	46	0.36	0.18	97	0.22	0.18
Share not healed (meth. 2)	225	0.29	0.17	46	0.42	0.17	97	0.28	0.17
Share not healed (meth. 3)	225	0.21	0.16	46	0.32	0.17	97	0.21	0.18
- ln(share not healed)/t (meth. 1)	213	0.048	0.01	46	0.04	0.02	82	0.087	0.03
Treatment response of subjs. (meth. 1)	225	0.07	0.15	46	-0.09	0.28	97	-0.13	0.14
- (ln(S1) - ln(S0))/t (meth. 1)	212	0.007	0.01	45	0.00	0.02	82	0.03	0.03
Antacids permitted in trial (1-5)?	225	3.66	0.64	46	3.49	0.61	97	3.10	1.15
Frequency of dosage (times/day)	225	1.74	1.00	46	2.75	0.98	97	1.19	0.66
Total daily dosage (mg)	225	0.44	0.36	46	0.09	0.23	97	0.03	0.02
Share male	212	0.72	0.07	41	0.73	0.09	95	0.70	0.09
Share that smoke	181	0.57	0.09	43	0.54	0.13	91	0.49	0.11
Average age (years) of subjs.	207	46	4	41	44	5	87	46	5
<b>Control arm</b>									
Date trial results published	86	1986	5.04	40	1989	3.79	67	1991	4.26
Share of subjects not given control	86	0.56	0.11	40	0.53	0.07	67	0.53	0.08
Share of arms in placebo-control trials	86	0.75	0.43	40	0.39	0.49	67	0.04	0.19
Share in antacid-control trials	86	0.32	0.47	40	0.00	0.00	67	0.00	0.00
Share in lower-class drug control trials	86	0.00	0.00	40	0.61	0.49	67	0.96	0.19
Share same-class drug control trials	86	0.00	0.00	40	0.00	0.00	67	0.00	0.00
Average number of measurements	86	1.29	0.53	40	1.35	0.53	67	1.60	0.64
Number enrolled	86	89	85	40	130	116	67	113	39
Number evaluated (per protocol)	86	79	76	40	117	105	67	104	38
Share of subjs. not healed (method 1)	86	0.56	0.23	39	0.36	0.28	67	0.36	0.22
Share not healed (meth. 2)	86	0.60	0.20	39	0.41	0.26	67	0.42	0.20
Share not healed (meth. 3)	86	0.50	0.21	39	0.32	0.25	67	0.33	0.21
- ln(share not healed)/t (meth. 1)	85	0.03	0.02	38	0.04	0.02	66	0.05	0.02
Antacids permitted in trial (1-5)?	86	3.31	1.04	40	3.52	0.60	67	3.16	1.15
Share male	77	0.76	0.10	35	0.77	0.09	65	0.70	0.09
Share that smoke	59	0.63	0.12	37	0.52	0.11	61	0.49	0.14
Average age (years) of subjs.	68	45	5	35	44	5	59	46	6

Notes. Each observation represents a measurement on the indicated arm of indicated trial. Each trial may have more than test-treatment arm and more than one measurement on each arm. Therefore, there are more observations than trials. However, each trial has only one control arm. While all trials are controlled, not all are placebo controlled. Means and standard deviations are calculated weighting each arm in proportion to the number of subjects evaluated per protocol, regardless of the number of measurements on the arm. Frequency of medication and total dosage are not provided for control arms because such variables are meaningless for placebo arms. Share of subjects not given control is very high for H<sub>2</sub>-blocker trials because same-class control trials are assumed to have a share treated of one. There is assumed to be no control arm in such trials. The antacid permitted variable takes a value of 1 if antacids were prohibited, 2 if discouraged, 3 if permitted or not discussed, 4 if antacids were provided, and 5 if antacids were required.

Table 2: Summary statistics on data from ulcer trials, by drug type and arm.

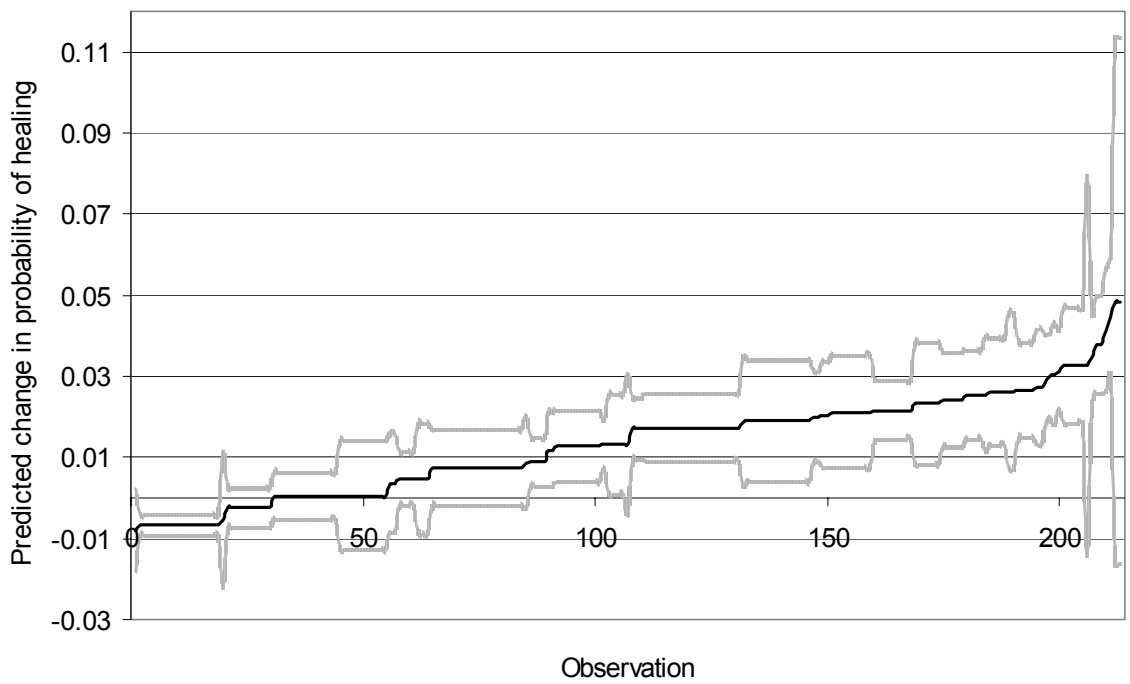
Arm Specification	Treatment				Control			
	1	2	3	4†	1	2	3†	4†
Constant	0.038 ***	0.026 ***	0.085	0.249 ***	0.034 ***	0.043 ***	-2.31 ***	-2.4 ***
Std. err.	0.001	0.012	0.096	0.108	0.0005	0.01	0.62	0.535
Share treated (d)	0.012 ***	0.028 ***	0.004	-0.14	-0.015 ***	-0.03 **	4.304 ***	4.389 ***
	0.001	0.013	0.145	0.137	0.001	0.015	1.227	1.052
Antacid role (1-5)		0.002	-0.01 ***			-0	-0.01 ***	
		0.002	0.002			0.003	0.005	
Daily freq. of treatment of treatment		-0.007 ***	0.003					
		0.002	0.003					
Total daily dosage (mg)		0.002	0.097 ***					
		0.036	0.042					
Cimetidine dosage (mg)		0.025	-0.06 *					
		0.032	0.036					
Ranitidine dosage (mg)		0.097 ***	0.082 ***					
		0.023	0.029					
Antacid role * d		-0.004 ***	0.009 ***			0.004	0.02 ***	
		0.002	0.003			0.005	0.009	
Freq. * d		0.007 ***	-0					
		0.003	0.004					
Dosage * d		0.009	-0.1 ***					
		0.037	0.043					
Cimetidine dosage * d		-0.031	0.056 *					
		0.033	0.036					
Ranitidine dosage * d		-0.093 ***	-0.08 ***					
		0.024	0.032					
Share male			-0.08 ***	-0.1 ***			-0.12	-0.19
			0.033	0.026			0.177	0.175
Share smokers			0.015	0.021			0.292 ***	0.325 ***
			0.016	0.017			0.125	0.123
Log age (mean yrs)			-0.01	-0.04 *			0.607 ***	0.625 ***
			0.024	0.026			0.141	0.121
Male * d			0.069 *	0.097 ***			0.181	0.301
			0.047	0.034			0.349	0.339
Smoker * d			0.012	0.001			-0.54 ***	-0.62 ***
			0.025	0.023			0.246	0.244
Log age * d			-0.01	0.021			-1.11 ***	-1.12 ***
			0.037	0.035			0.276	0.237
<u>F-tests of joint significance of coefficients on</u>								
d	406917 ***	4.88 **	0.00	1.04	187.86 ***	3.63 *	12.30 ***	17.40 ***
P-value	0.00	0.03	0.98	0.31	0.00	0.06	0.00	0.00
Trial variables * d		83.06 ***	44.32 ***			0.52	4.88 **	
		0.00	0.00			0.47	0.03	
Subject variables * d			2.25	10.41 **			27.71 ***	36.92 ***
			0.52	0.02			0.00	0.00
Trial and subject vars. * d			45.65 ***				28.42 ***	
			0.00				0.00	
d and all interactions		85.79 ***	59.03 ***	20.27 ***		11.09 ***	64.76 ***	65.25 ***
		0.00	0.00	0.00		0.00	0.00	0.00

Table 3: Evidence of placebo effects from H2-blocker trials with informed consent employing x-captures-selection approach to self-selection.

Table 3, continued.

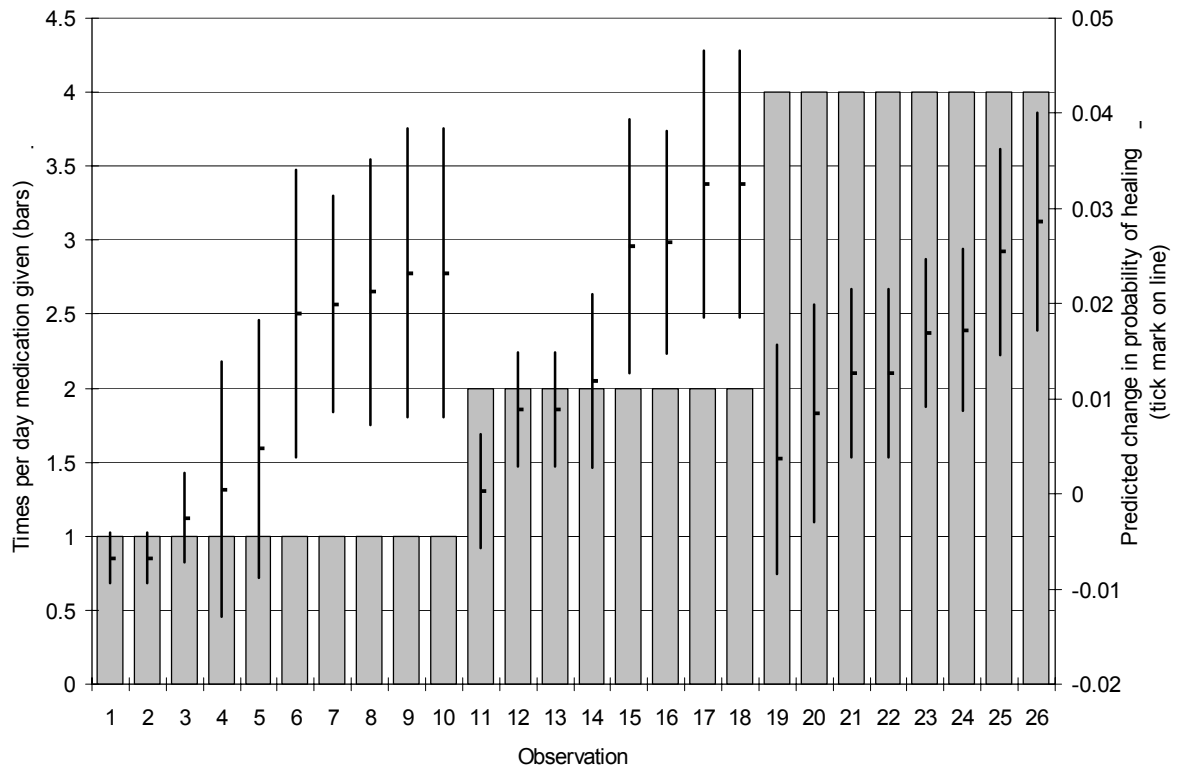
Arm Specification	Treatment				Control			
	1	2	3	4†	1	2	3†	4†
Studies	58	58	37	37	58	58	37	37
Arms	133	133	98	98	58	58	37	37
Obs./Meas.	213	213	163	163	85	85	55	55
<u>Estimated dy/dd</u>								
Mean	0.012	0.012	0.02	0.01	-0.015	-0.02	-0.06	-0.05
Std. err.	0	0.013	0.02	0.01	0	0	0.12	0.13
Min.	0.012	-0.01	-0.06	-0.01	-0.015	-0.02	-0.4	-0.42
Max	0.012	0.049	0.06	0.04	-0.015	-0.01	0.12	0.1
<u>T-value of estimate</u>								
Mean	12	1.338	2.57	2.28	-15	-1.91	-1.43	-1.15
Std. err.	0	3.006	2.21	1.32	0	0.7	2.61	2.94
Min.	12	-5.01	-2.86	-1.06	-15	-2.71	-6.36	-6.36
Max	12	6.303	6.25	4.39	-15	-0.62	3.33	3.51
<u>Measurements manifesting positive relationship between d and outcomes (by conf. level)</u>								
At all	213	183	130	151	0	0	12	17
At 85%	213	153	107	127	0	0	7	12
At 90%	213	153	103	122	0	0	4	9
At 95%	213	124	87	114	0	0	4	8
<u>Measurements manifesting a non-zero overall relationship between d and outcomes (by conf. level)</u>								
At 85%	213	167	111	121	85	35	31	36
At 90%	213	142	92	114	85	35	31	36
At 95%	213	142	85	107	85	35	30	34

Notes. Estimation was by feasible GLS. The dependent variable is the  $-\ln(S(t))/t$ , where  $S(t)$  is the fraction that remain ill after  $t$  days. Each measurement on an arm of a trial counts as an observation. Some trials offer multiple measurements on the same arm. I weighted observations such that each arm makes a contribution to estimates in proportion to the number of subjects in the arm, regardless of the number of measurements made on each subject. I permit group-wise heteroskedasticity at the trial-level, but not at the arm-level. The dependent variable is calculated assuming subjects who attrite out heal at the same rate as those who are evaluated. Standard errors are reported below coefficients. Coefficients significant at the 1%/5%/10% level marked with \*\*/\*\*/\* . P-values for F-tests for the joint significance of indicated groups of variables are reported below the F-test statistics. Specifications that fail Pregibon's link test are marked with a dagger. Counts are generated in two steps. First, for each observation or, equivalently, measurement, the vector of variables interacted with  $d$  (including  $d$  itself) is multiplied by the estimated coefficients on that vector and divided by  $d$ . Second, standard errors for this product are calculated based on the variance-covariance matrix for the estimated coefficients. The two panels of counts present the number of observations for which the first step calculation yields positive or non-zero results, respectively, at different confidence levels given the standard errors computed in the second step. The first panel employs a one-sided test and the second a two-sided test.



Notes. Black line gives estimated  $\partial y_j / \partial d_j$ . Gray lines give 95% confidence intervals for predictions.

Figure 3: Predicted change in outcomes due to a change in share treated across all observations on treatment arm of H2-blocker trials, employing x-captures-selection approach to self-selection and specification 2.



Notes. Tick mark and line give mean and 95% confidence interval, respectively, for predictions.

Figure 4: Predicted change in outcomes due to a change in share treated across observations at week 2 endoscopy in H2-blocker trials, employing x-captures-selection approach to self-selection and specification 2.

Treatment	H2-Blocker				Prostagladin				PPI			
	1	2	3	4	1	2	3	4	1	2	3	4
<u>No-conventional-altern. method</u>			†	†								
Obs./meas.	174	174	133	133								
Studies	38	38	22	22								
<u>Estimated dy/dd</u>												
Mean	0.01	0.02	-0.03	0.004								
Std. err.	0.001	0.09	0.03	0.003								
<u>Obs. manifesting positive relationship between share treated and outcomes (by confidence level, one-sided test)</u>												
At 90%	174	58	5	43								
At 95%	174	56	1	37								
<u>Obs. manifesting a non-zero overall relationship between share treated and outcomes (by confidence level, two-sided test)</u>												
At 90%	174	88	52	37								
At 95%	174	88	34	30								
<u>X-captures-selection method</u>												
				†			†					
Obs./meas.	213	213	163	163	46	46	41	41	82	82	70	70
Studies	58	58	37	37	21	21	17	17	29	29	24	24
<u>Estimated dy/dd</u>												
Mean	0.01	0.01	0.02	0.01	-0.04	-28	-118	0.13	-0.02	0.01	-0.06	-0.03
Std. err.	0	0.01	0.02	0.01	0	84.5	343	0.11	0	0.10	0.27	0.16
<u>Obs. manifesting positive relationship between share treated and outcomes (by confidence level, one-sided test)</u>												
At 90%	213	153	103	122	0	0	3	8	0	16	21	25
At 95%	213	124	87	114	0	0	0	6	0	16	21	25
<u>Obs. manifesting a non-zero overall relationship between share treated and outcomes (by confidence level, two-sided test)</u>												
At 90%	213	142	92	114	46	40	6	10	82	24	45	59
At 95%	213	142	85	107	46	40	1	6	82	18	45	59
<u>Additional-d method</u>												
		†										
Obs./meas.	213	213	163	163	46	46	41	41	82	82	70	70
Studies	58	58	37	37	21	21	17	17	29	29	24	24
d <sup>2</sup>	0.14 ***	0.15 ***	0.15 ***	0.18 ***	1.29 **	-1.86 ***	-3.02 **	0.51	0.001	-0.39 ***	-0.03	-0.04
	0.01	0.02	0.03	0.02	0.78	0.80	1.55	1.25	0.09	0.11	0.18	0.08

Notes. Estimation was by feasible GLS. The dependent variable is the  $-\ln(S(t))/t$ , where  $S(t)$  is the fraction that remain ill after  $t$  days. Each measurement on an arm of a trial counts as an observation. Some trials offer multiple measurements on the same arm. I weighted observations such that each arm makes a contribution to estimates in proportion to the number of subjects in the arm, regardless of the number of measurements made on each subject. I permit group-wise heteroskedasticity at the trial-level, but not at the arm-level. The dependent variable is calculated assuming subjects who attrite out heal at the same rate as those who are evaluated. Specification (1) includes a constant and  $d$ ; (2) adds trial-level variables (antacid usage, daily frequency of medication, total daily dosage of medication, total daily dosage of the more common drugs in the relevant class of medications) as well as interactions of these trial-level variables with  $d$ ; (3) adds subject-level variables (sex, smoker, and age) and their interactions with  $d$ ; and (4) removes the trial-level variables and interactions from (3). When I employ the instrument approach to test for placebo effects, I add  $d^2$  to each specification. Coefficients significant at the 1%/5%/10% level marked with \*\*\*/\*\*/\* . Specifications that fail Pregibon's link test are marked with a dagger. The standard error of predicted  $dy/dd$  is standard error across observations, not standard error of the prediction for a given observation. The former is zero for the first specification since the observations have identical values for the only covariate (the constant) other than  $d$ .

Table 4: Evidence of placebo effects in treatment arms of ulcer trials with informed consent, by drug type, specification, and approach to self-selection.

Treatment	H2-Blocker				Prostagladin				PPI			
	1	2	3	4	1	2	3	4	1	2	3	4
<u>No-conventional-altern. method</u>												
Obs./meas.	51	51	30	30								
Studies	38	38	22	22								
<u>Estimated dy/dd</u>												
Mean	-0.04	-0.04	0.03	0.02								
Std. err.	0.01	0.01	0.07	0.07								
<u>Obs. manifesting positive relationship between share treated and outcomes (by confidence level, one-sided test)</u>												
At 90%	0	0	3	2								
At 95%	0	0	3	2								
<u>Obs. manifesting a non-zero overall relationship between share treated and outcomes (by confidence level, two-sided test)</u>												
At 90%	51	46	4	3								
At 95%	51	46	2	2								
<u>X-captures-selection method</u>												
			†	†								
Obs./meas.	85	85	55	55	38	38	33	33	66	66	56	56
Studies	58	58	37	37	21	21	17	17	29	29	24	24
<u>Estimated dy/dd</u>												
Mean	-0.02	-0.02	-0.06	-0.05	-0.02			0.39	0.01	0.01	-0.02	-0.02
Std. err.	0	0.004	0.12	0.13	0			0.30	0	0.08	0.28	0.29
<u>Obs. manifesting positive relationship between share treated and outcomes (by confidence level, one-sided test)</u>												
At 90%	0	0	4	9	38			21	0	10	20	22
At 95%	0	0	4	8	38			21	0	10	18	18
<u>Obs. manifesting a non-zero overall relationship between share treated and outcomes (by confidence level, two-sided test)</u>												
At 90%	85	35	31	36	38			27	0	52	42	42
At 95%	85	35	30	34	38			27	0	52	39	42
<u>Additional-d method</u>												
							†					
Obs./meas.	85	85	55	55	38	38	33	33	66	66	56	56
Studies	58	58	37	37	21	21	17	17	29	29	24	24
d <sup>2</sup>	0.06	0.11 **	-0.56 ***	-0.63 ***	-1.91 ***	-4.65 ***		-0.57	-0.81 ***	-0.37 **	0.62	0.55
	0.05	0.07	0.09	0.10	0.67	0.88		0.49	0.18	0.20	1.01	1.00

Notes. Estimation was by feasible GLS. The dependent variable is the  $-\ln(S(t))/t$ , where  $S(t)$  is the fraction that remain ill after  $t$  days. Each measurement on an arm of a trial counts as an observation. Some trials offer multiple measurements on the same arm. I weighted observations such that each arm makes a contribution to estimates in proportion to the number of subjects in the arm, regardless of the number of measurements made on each subject. I permit group-wise heteroskedasticity at the trial-level, but not at the arm-level. The dependent variable is calculated assuming subjects who attrite out heal at the same rate as those who are evaluated. Specification (1) includes a constant and  $d$ ; (2) adds trial-level variables (antacid usage, daily frequency of medication, total daily dosage of medication, total daily dosage of the more common drugs in the relevant class of medications) as well as interactions of these trial-level variables with  $d$ ; (3) adds subject-level variables (sex, smoker, and age) and their interactions with  $d$ ; and (4) removes the trial-level variables and interactions from (3). When I employ the instrument approach to test for placebo effects, I add  $d^2$  to each specification. Coefficients significant at the 1%/5%/10% level marked with \*\*\*/\*\*/\*\*. Specifications that fail Pregibon's link test are marked with a dagger. The standard error of predicted  $dy/dd$  is standard error across observations, not standard error of the prediction for a given observation. The former is zero for the first specification since the observations have identical values for the only covariate (the constant) other than  $d$ .

Table 5: Evidence of placebo effects in control arms of ulcer trials with informed consent, by drug type, specification, and approach to self-selection.

Class of Drug Tested in Trial	Strength of Evidence of Placebo Effects in	
	Treatment Arm	Control Arm
H <sub>2</sub> -Blockers	Strong	Moderate
Prostaglandins	Weak	Weak
Proton-Pump Inhibitors	Moderate	Moderate

Table 6: Summary of results.